

RESEARCH

Open Access



# Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes

Alexandre Heurteau<sup>1</sup>, Charlène Perrois<sup>1</sup>, David Depierre<sup>1</sup>, Olivier Fosseprez<sup>1</sup>, Jonathan Humbert<sup>1,2</sup>, Stéphane Schaak<sup>1</sup> and Olivier Cuvier<sup>1\*</sup>

\* Correspondence: [olivier.cuvier@inserm.fr](mailto:olivier.cuvier@inserm.fr)

<sup>1</sup>Chromatin Dynamics and Cell Proliferation, Center of Integrative Biology (CBI), Laboratoire de Biologie Moléculaire Eucaryote (LBME), CNRS, Université Fédérale Paul Sabatier de Toulouse (UPS), F-31000 Toulouse, France  
Full list of author information is available at the end of the article

## Abstract

**Background:** Chromosomes are subdivided spatially to delimit long-range interactions into topologically associating domains (TADs). TADs are often flanked by chromatin insulators and transcription units that may participate in such demarcation. Remarkably, single-cell *Drosophila* TAD units correspond to dynamic heterochromatin nano-compartments that can self-assemble. The influence of insulators on such dynamic compartmentalization remains unclear. Moreover, to what extent heterochromatin domains are fully compartmentalized away from active genes remains unclear from *Drosophila* to human.

**Results:** Here, we identify H3K27me3 micro-domains genome-wide in *Drosophila*, which are attributed to the three-dimensional spreading of heterochromatin marks into euchromatin. Whereas depletion of insulator proteins increases H3K27me3 spreading locally, across heterochromatin borders, it concomitantly decreases H3K27me3 levels at distant micro-domains discrete sites. Quantifying long-range interactions suggests that random interactions between heterochromatin TADs and neighbor euchromatin cannot predict the presence of micro-domains, arguing against the hypothesis that they reflect defects in self-folding or in insulating repressive TADs. Rather, micro-domains are predicted by specific long-range interactions with the TAD borders bound by insulator proteins and co-factors required for looping. Accordingly, H3K27me3 spreading to distant sites is impaired by insulator mutants that compromise recruitment of looping co-factors. Both depletions and insulator mutants significantly reduce H3K27me3 micro-domains, deregulating the flanking genes.

**Conclusions:** Our data highlight a new regulatory mode of H3K27me3 by insulator-based long-range interactions controlling distant euchromatic genes.

**Keywords:** Insulators, Higher-order chromatin folding, Topologically associating domains, Chromosome compartmentalization into euchromatin and heterochromatin domains



## Introduction

Chromatin folding in 3D has been revealed through microscopy [1, 2] and genome-wide chromosome conformation capture methodologies (3C/Hi-C [3]); [4], which eventually highlighted how chromosomes fold into topologically associating domains (TADs) and sub-TADs [5–12] [13–15]. TADs notably promote specific long-range contacts between distant sites and regulatory elements that localize within the same topological unit. At higher resolution, smaller TADs may further delimit cell-type-specific long-range contacts, thus contributing to cell-type-specific gene expression programs [6, 9, 16–22].

Hi-C contact matrices show that frequencies of long-range interactions (LRIs) largely depend on distances, as explained by polymer physics models [14, 23]. LRIs highlight how TADs are insulated from neighboring domains due to either self-assembly properties of TADs or to their delimitation by “insulators/boundaries” [12, 20, 24–27]. TAD insulators may restrict LRIs with sites localized in adjacent TADs. Accordingly, removal of a TAD border results in gene deregulation. This may involve long-range contacts between a gene and the regulatory sequences localized in the adjacent TAD [9, 28, 29]. The existence of TADs and domains may however not solely rely on borders but also on intrinsic self-assembling/propagation properties, e.g., shown for polycomb repressive complexes (PRC1 or 2) in inactive TADs [1, 15, 18, 30–33], or of transcription factors in context of active TADs [34] along with epigenetic mechanisms involving non-coding RNAs, DNA methylation, or post-translational modifications (PTMs) of histones [7, 14, 35–38]. 3D clustering of factors may then lead to “phase separation” involving multimerized DNA-protein complexes or liquid droplet [31, 34, 39], possibly accounting for the maintenance of gene expression programs [30, 40]. High-resolution mapping of TADs in *Drosophila* highlights their good correspondence with repressed domains including at levels of highly resolved loop-based (sub-)TADs [16, 20, 21]. Strikingly, repressive TADs are dynamic structures defining nano-compartments visible in single cells [1, 15]. How much epigenetically marked repressed TADs maintain their identity depending on self-maintenance or on TAD borders remains unknown.

Of interest, TAD borders fall into sites recognized by a family of factors called insulator proteins that notably include CCCTC-binding factor (CTCF) [41, 42]. Additional insulator proteins are being identified defining a growing family of factors from *Drosophila* to human [43]. Major factors include CTCF, GAGA-binding factor (GAF) [44], M1BP [20, 45], and Boundary Element-Associated Factor of 32 KDa (Beaf32) [46]. Insulator proteins “insulate” a (*trans*-)gene from its environment [47] and from activation by promiscuous enhancers [48]. Insulating activity relies on interactions with co-factors including cohesin or CP190 to stabilize loops through evolutionary conserved principles [19, 22, 25, 26, 49–51]. Remarkably, inversion of CTCF sites impairs genome topology and enhancer-promoter long-range contacts [28, 52]. Insulator-based regulation of long-range contacts may contribute to link transcriptional programs to 3D folding, as shown upon stem cell differentiation [6, 9, 19, 24, 37].

Insulators further act as chromatin barrier insulators (CBI) [53] that participate in Hox-based para-segment identity in flies [54, 55]. CTCF (and other insulator protein) sites are specifically enriched at heterochromatin domain borders [42, 46, 56–58]. In *Drosophila*, borders lacking dCTCF harbor other types of insulator proteins whose binding is required to block the spreading of repressive histone marks, including

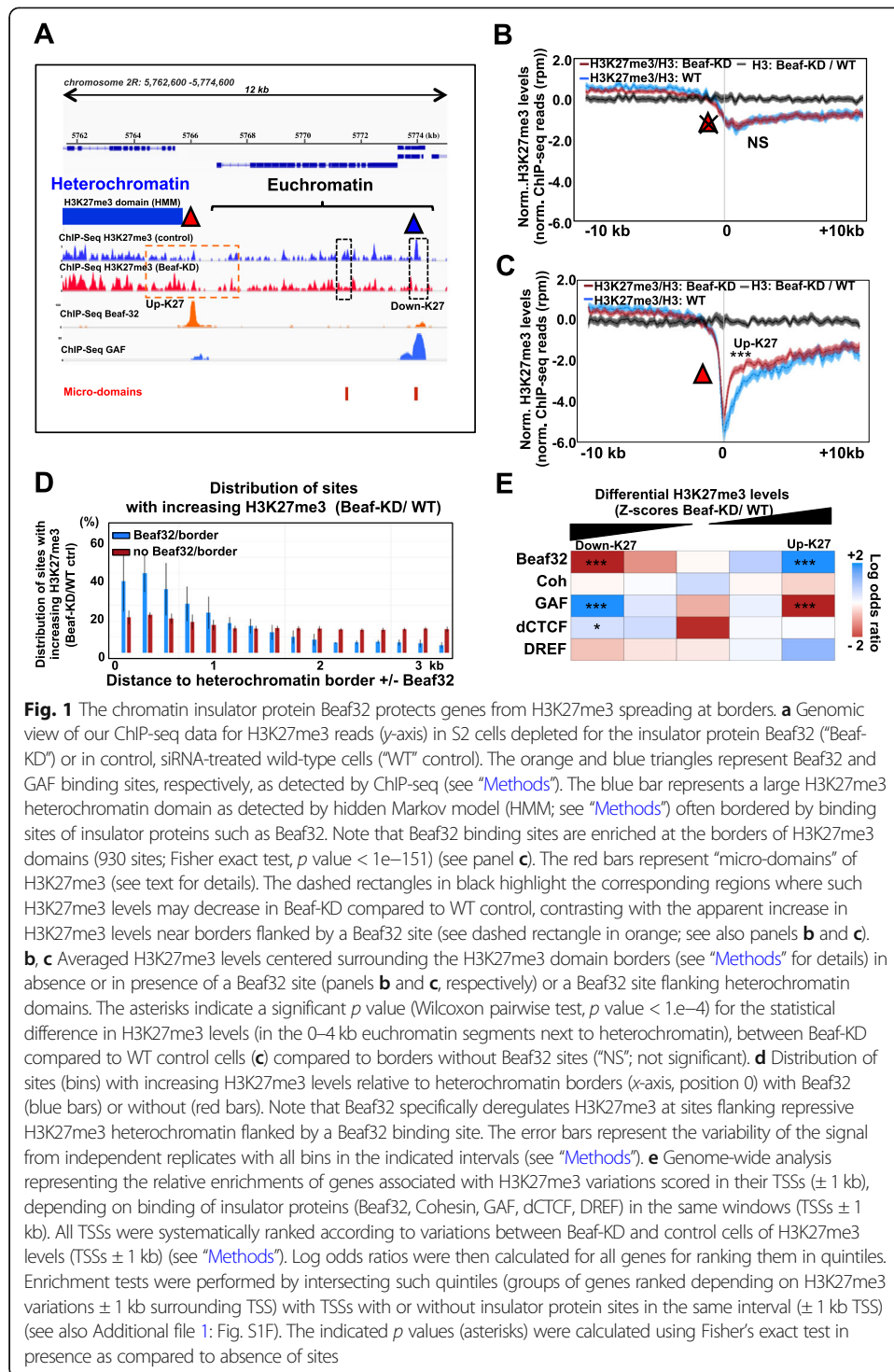
histone H3 trimethylated on -lysine 9 (H3K9me3) [46] and on -lysine 27 (H3K27me3) [59, 60]. Removal of insulators may not systematically lead to spreading defects for all borders. Moreover, the role of insulators is unclear with respect to the highly dynamic nature of TAD compartments in single cells [1, 15], suggesting possible interactions between H3K27me3 domains and the flanking euchromatin.

Here, we analyzed the spreading of heterochromatin H3K27me3 marks depending on insulator proteins and long-range interactions (LRIs) by analyzing Hi-C data [16, 20, 21] aggregated onto TAD borders. Removal of insulator proteins Beaf32 leads to H3K27me3 spreading locally, across borders. In addition, Beaf32 promotes spreading onto distant euchromatin sites named “micro-domains.” Systematic measurements of LRIs suggest that H3K27me3 micro-domains do not form due to the weakness of TAD borders. Rather, micro-domains were visible at sites showing high levels of LRIs, including distant dCTCF and GAF insulator sites bound by the looping co-factor CP190. Also, micro-domain formation appears to depend on such specific insulator-mediated LRIs utilized to spread H3K27me3 to distant sites through looping. Supporting these results, specific synthetic mutants that impair LRIs compromise distant spreading over micro-domains. Distant spreading at micro-domains is further associated with insulator-based control of genes and it influences H3K27me3 throughout developmental stages of *Drosophila*. Our data highlight how specific LRIs encoded by insulator-mediated loops contribute to the regulation of H3K27me3 spreading over the distance. We propose that micro-domains reflect how insulators participate to chromatin folding dynamics in 3D, aside additional factors required to separate heterochromatin nano-compartments from nearby euchromatin domains.

## Results

### H3K27me3 micro-domains are associated with dCTCF and GAF insulator binding sites

Insulator proteins often bind to sites flanking heterochromatin domains from *Drosophila* to human [46, 56], as illustrated for Beaf32 (Fig. 1a). In such contexts, removal of Beaf32 was accompanied with the downregulation of adjacent genes due to heterochromatin spreading [46, 59]. Increasing levels of H3K27me3 levels could be detected near Beaf32 sites flanking heterochromatin (Fig. 1a–c). Systematic measurements showed that Beaf32 depletion (“Beaf-KD”) led to a relatively modest yet significant increase in H3K27me3 levels as compared to siRNA-treated control cells (Fig. 1b, c;  $p$  value of  $1e^{-4}$ ) (Additional file 1: Fig. S1A). Such increase was specific of heterochromatin domains with a Beaf32 site as compared to control domains without a Beaf32 site. Of note, the increase was not detected for total histone H3 reads (Fig. 1b, c), indicating that it is specific of the H3K27me3 mark. Furthermore, such increase in H3K27me3 levels was preferentially associated with genes being downregulated upon Beaf-KD, unlike control or upregulated genes (Additional file 1: Fig. S1B) [59]. Furthermore, genes encoding the subunits of Polycomb repressor complex showed no variation in expression upon such depletion of insulator proteins (Additional file 1: Fig. S1E), arguing against indirect defects in regulating H3K27me3 at least due to PRC2 deregulation. Actually, the distribution of H3K27me3 spreading upon Beaf32 depletion was detected specifically for heterochromatin borders flanked by Beaf32 sites as shown (Fig. 1d), thus confirming a specific defect in spreading.



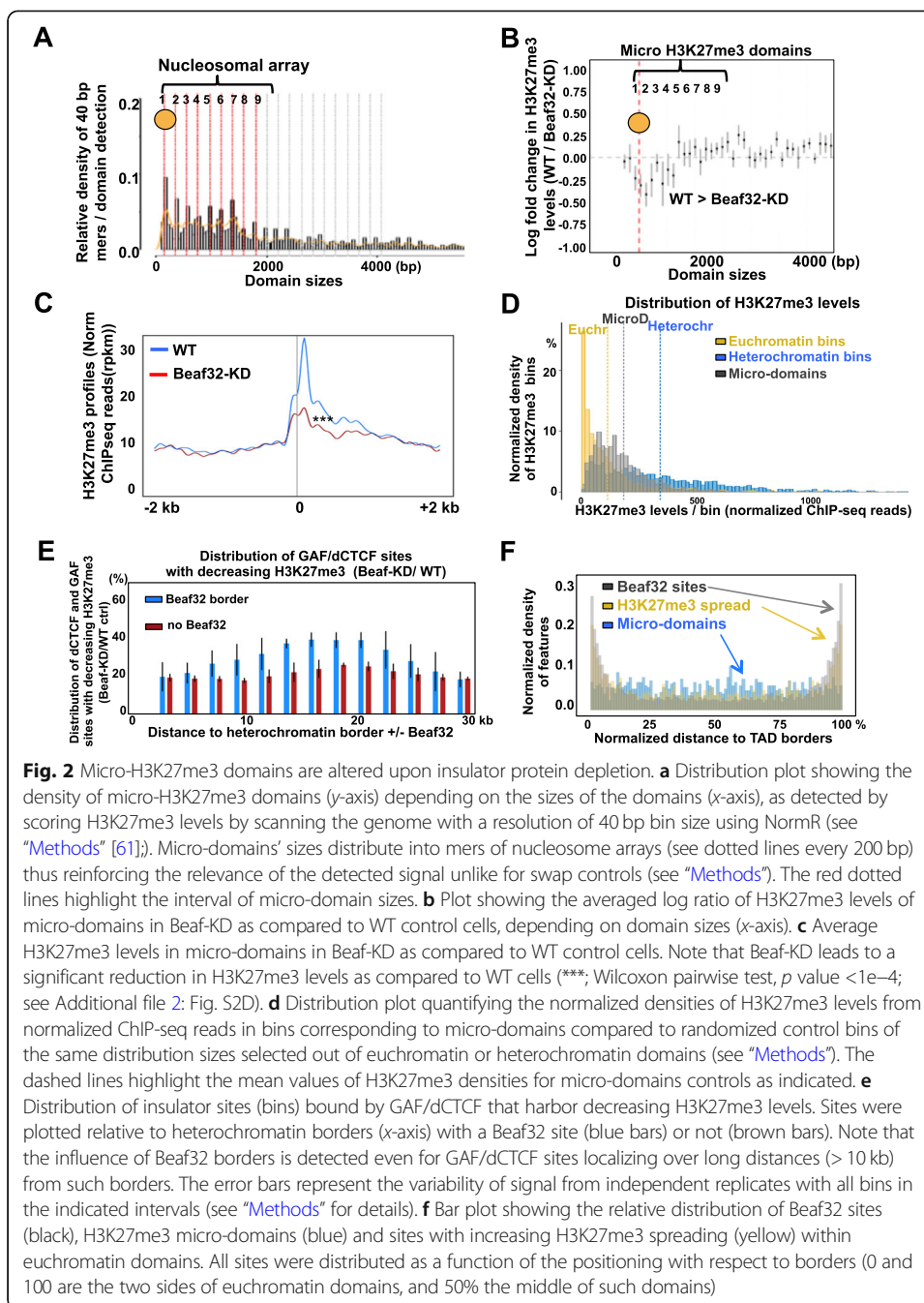
**Fig. 1** The chromatin insulator protein Beaf32 protects genes from H3K27me3 spreading at borders. **a** Genomic view of our ChIP-seq data for H3K27me3 reads (y-axis) in S2 cells depleted for the insulator protein Beaf32 (“Beaf-KD”) or in control, siRNA-treated wild-type cells (“WT” control). The orange and blue triangles represent Beaf32 and GAF binding sites, respectively, as detected by ChIP-seq (see “Methods”). The blue bar represents a large H3K27me3 heterochromatin domain as detected by hidden Markov model (HMM; see “Methods”) often bordered by binding sites of insulator proteins such as Beaf32. Note that Beaf32 binding sites are enriched at the borders of H3K27me3 domains (930 sites; Fisher exact test,  $p$  value  $< 1e-151$ ) (see panel **c**). The red bars represent “micro-domains” of H3K27me3 (see text for details). The dashed rectangles in black highlight the corresponding regions where such H3K27me3 levels may decrease in Beaf-KD compared to WT control, contrasting with the apparent increase in H3K27me3 levels near borders flanked by a Beaf32 site (see dashed rectangle in orange; see also panels **b** and **c**). **b, c** Averaged H3K27me3 levels centered surrounding the H3K27me3 domain borders (see “Methods” for details) in absence or in presence of a Beaf32 site (panels **b** and **c**, respectively) or a Beaf32 site flanking heterochromatin domains. The asterisks indicate a significant  $p$  value (Wilcoxon pairwise test,  $p$  value  $< 1.e-4$ ) for the statistical difference in H3K27me3 levels (in the 0–4 kb euchromatin segments next to heterochromatin), between Beaf-KD compared to WT control cells (**c**) compared to borders without Beaf32 sites (“NS”; not significant). **d** Distribution of sites (bins) with increasing H3K27me3 levels relative to heterochromatin borders (x-axis, position 0) with Beaf32 (blue bars) or without (red bars). Note that Beaf32 specifically deregulates H3K27me3 at sites flanking repressive H3K27me3 heterochromatin flanked by a Beaf32 binding site. The error bars represent the variability of the signal from independent replicates with all bins in the indicated intervals (see “Methods”). **e** Genome-wide analysis representing the relative enrichments of genes associated with H3K27me3 variations scored in their TSSs ( $\pm 1$  kb), depending on binding of insulator proteins (Beaf32, Cohesin, GAF, dCTCF, DREF) in the same windows (TSSs  $\pm 1$  kb). All TSSs were systematically ranked according to variations between Beaf-KD and control cells of H3K27me3 levels (TSSs  $\pm 1$  kb) (see “Methods”). Log odds ratios were then calculated for all genes for ranking them in quintiles. Enrichment tests were performed by intersecting such quintiles (groups of genes ranked depending on H3K27me3 variations  $\pm 1$  kb surrounding TSS) with TSSs with or without insulator protein sites in the same interval ( $\pm 1$  kb TSS) (see also Additional file 1: Fig. S1F). The indicated  $p$  values (asterisks) were calculated using Fisher’s exact test in presence as compared to absence of sites

Our results showed that the influence of Beaf32 was not drastic, raising the possibility that additional factors may be required to block heterochromatin. Since two distant insulators can interact to form a loop, we sought to test if two insulators could better block H3K27me3 spreading. However no difference in spreading was detected depending on the presence of one or two insulators bracketing the domain (Additional file 1: Fig. S1D; see below). Alternatively, the moderate spreading of H3K27me3 could suggest

a requirement for additional factors that participate in blocking heterochromatin. Since > 91% of Beaf32 sites co-localize ( $\pm 1$  kb) with TSSs, we sought to better evaluate the influence of Beaf32 or of other factors by taking assessing H3K27me3 in an otherwise similar genomic context,  $\pm 1$  kb of TSSs (see “Methods”). A systematic scoring of H3K27me3 variations between Beaf32-depleted cells compared to wild-type control confirmed that in this context, Beaf32 sites was the insulator proteins that was specifically associated with increasing levels of H3K27me3 (Fig. 1e). Of interest, the opposite effect—i.e., the decrease in H3K27me3 levels upon depletion of Beaf32 compared to control cells—was detected at certain insulator factor sites including GAF and to a lesser extent dCTCF sites (Fig. 1a, e). These results were also confirmed when scoring variations in H3K27me3 surrounding all insulator sites, independently of TSSs (Additional file 1: Fig. S1F).

Our above results prompted us to systematically detect regions where decreasing H3K27me3 levels might be detected upon Beaf32 depletion, genome-wide and without a priori. We thus re-analyzed our chromatin immunoprecipitation experiments (ChIP-seq) for H3K27me3 in control or Beaf32-depleted cells and scanned the genome with NormR [61]. Briefly, we scored normalized reads in sliding windows (bins) of 40 bp compared to input and then compared to depleted conditions (see “Methods”). As a result, novel “micro-domains” of H3K27me3 were identified (Fig. 2a). Of note, micro-domains could not be previously detected by classic, e.g., hidden Markov model (HMM) methods, in part because of their relatively small sizes and low H3K27me3 levels (see below). Plotting the density of 40 bp-bins showed a non-random distribution of their lengths, corresponding to nucleosome mers (Fig. 2a, b). Micro-domains corresponded to 2–8 nucleosomes with more than 65% of them of length < 2 kb. Most micro-domains further showed a significant reduction in the log ratio of H3K27me3 levels upon Beaf32 depletion compared to control cells (Fig. 2b). Such a decrease was more significant for micro-domains harboring 2 up to 4 nucleosome mers, as also confirmed by inspecting averaged profiles markedly impaired by the depletion compared to control cells (Fig. 2c; Additional file 1: Fig. S2A;  $p$  value of  $1e-6$ ). The decrease was confirmed by re-measuring H3K27me3 in micro-domains by qPCR (Additional file 1: Fig. S2B-C). The reduction was most significant for 2–4 nucleosome mers with no difference in spreading over 2 kb distances (Additional file 1: Fig. S2D-E). From these results, we defined a list of 1311 H3K27me3 micro-domains of sizes < 2 kb (Additional file 2: Table S1) for all subsequent genomic analyses, of which 722 flanked (< 1 kb) from a TSS (Additional file 3: Table S2). Micro-domains are distinct from known conventional heterochromatin domains as evident by differences in their sizes and intensities, as shown by genome-wide analyses of H3K27me3 levels in micro-domains, heterochromatin or euchromatin for bins of identical sizes (Fig. 2d). This illustrates how euchromatic micro-domains (730 bp average size), i.e., the equivalent of 3–4 nucleosomes, may be distinct from larger/denser and epigenetically stable heterochromatin domains.

Micro-domains could reflect the observed decrease upon Beaf32 depletion of H3K27me3 levels at GAF/dCTCF sites (Fig. 1e). Analyzing the distribution of such insulator sites with decreasing H3K27me3 levels showed their relative enrichment from 5 to 30 kb distances from a Beaf32 site as shown (Fig. 2e). Unlike H3K27me3 spreading, H3K27me3 micro-domains localized in euchromatin, away from Beaf32 borders

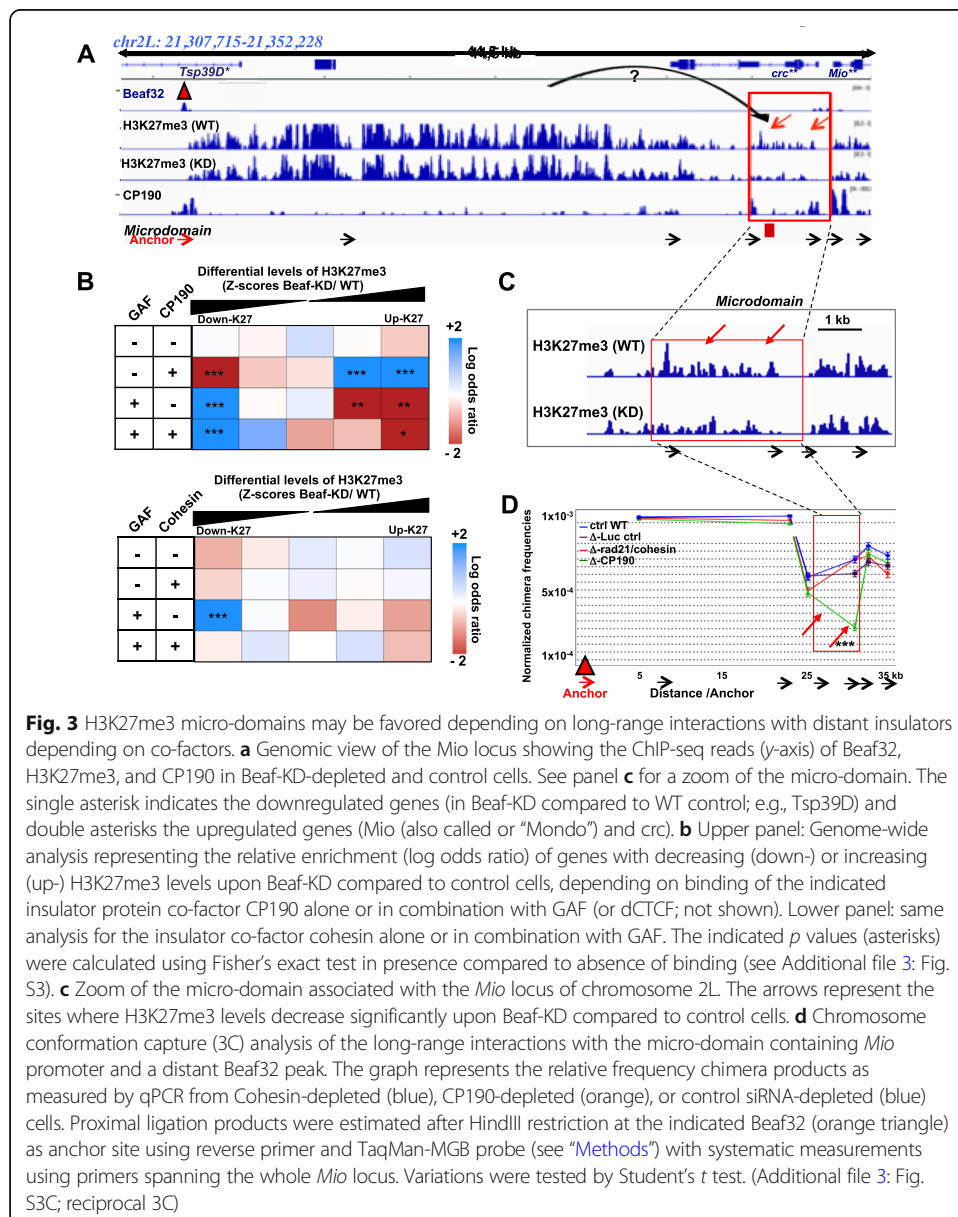


(Fig. 2f). Taken altogether, our data suggest that while the spreading of H3K27me3 levels occurs locally over Beaf32 borders, the concomitant decrease of H3K27me3 at distant micro-domains may involve long-range interactions with additional, distant GAF/dCTCF insulators.

### Micro-domain may form upon insulator-based long-range interactions

Insulator proteins like dCTCF or Beaf32 contribute to the folding of chromosomes into TADs [26]. We hypothesized that weak TADs unable to restrain H3K27me3 within

such topological unit might lead to micro-domain formation. An alternative possibility may involve the ability of insulator proteins to define specific LRIs with distant dCTCF or GAF [20, 50], independently of any contribution of insulators in assembling TADs. As an illustration, a H3K27me3 micro-domain was encountered at the distant GAF insulator sites flanking *Mio* locus, where Beaf32 establishes specific LRIs (Fig. 3a, c, red arrows) [50]. This micro-domain associated with *Mio* and *crc* genes was impaired upon Beaf32 depletion (Fig. 3a, c, red arrows). Beaf32 LRIs with GAF/dCTCF was shown to depend on co-factors including CP190 that is shared among all dCTCF, GAF, and Beaf32 types of insulators [50]. Accordingly, genome-wide analysis showed an enrichment of sites with most significant decreases in H3K27me3 levels in Beaf-KD cells when co-localizing with CP190, dCTCF, or GAF (Fig. 3b, upper matrix; *p* value of 1–4), in stark contrast to what was detected when it co-localizes with Beaf32



(Additional file 1: Fig. S3A-B). The involvement of CP190 was specific, contrasting with the additional co-factor cohesin that was not required for the decrease in H3K27me3 levels at GAF sites (Fig. 3b, lower matrix). Chromosome conformation capture (3C) further suggested that in contrast to CP190 depletion, depletion of cohesin did not affect long-range contacts at *Mio* as compared to control cells (Fig. 3d; Additional file 1: Fig. S3). Taken altogether, our data thus raised the possibility that H3K27me3 micro-domains form depending on presence of long-range interactions between insulator sites.

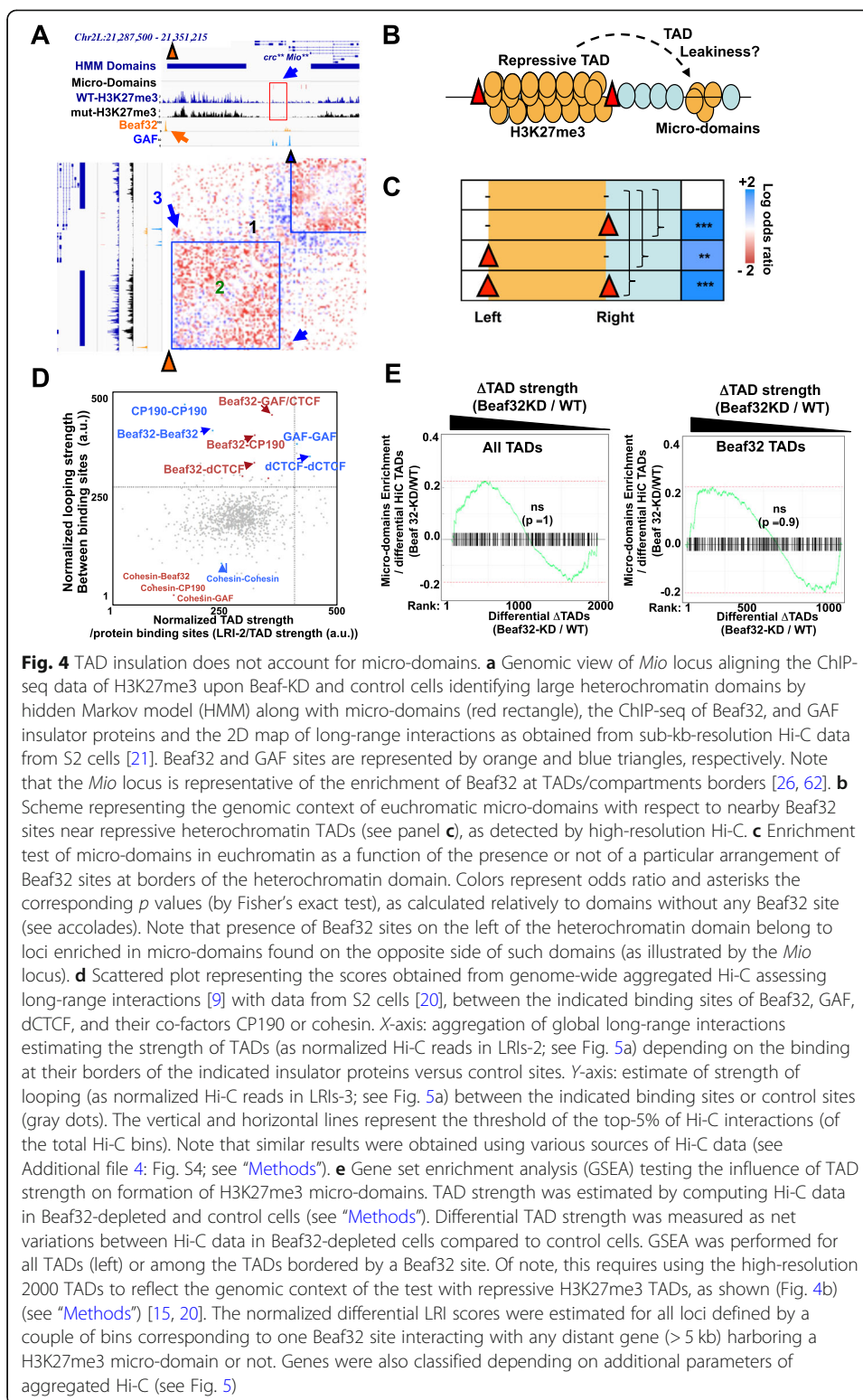
#### Specific long-range contacts rather than TAD leakiness may account for micro-domains

Our observations supported a model where Beaf32 regulates H3K27me3 micro-domains involving long-range interactions (Fig. 4a, b). In the case of *Mio*, the micro-domain was detected in the euchromatin domain localized on the opposite side of the Beaf32 site that flanks heterochromatin (Fig. 4a). Such an arrangement was found to be among the significant genomic contexts that favor micro-domains, providing the presence of a Beaf32 site on either side of heterochromatin (Fig. 4c; 336/722 micro-domains) or on both sides (4th row: 361/722 micro-domains).

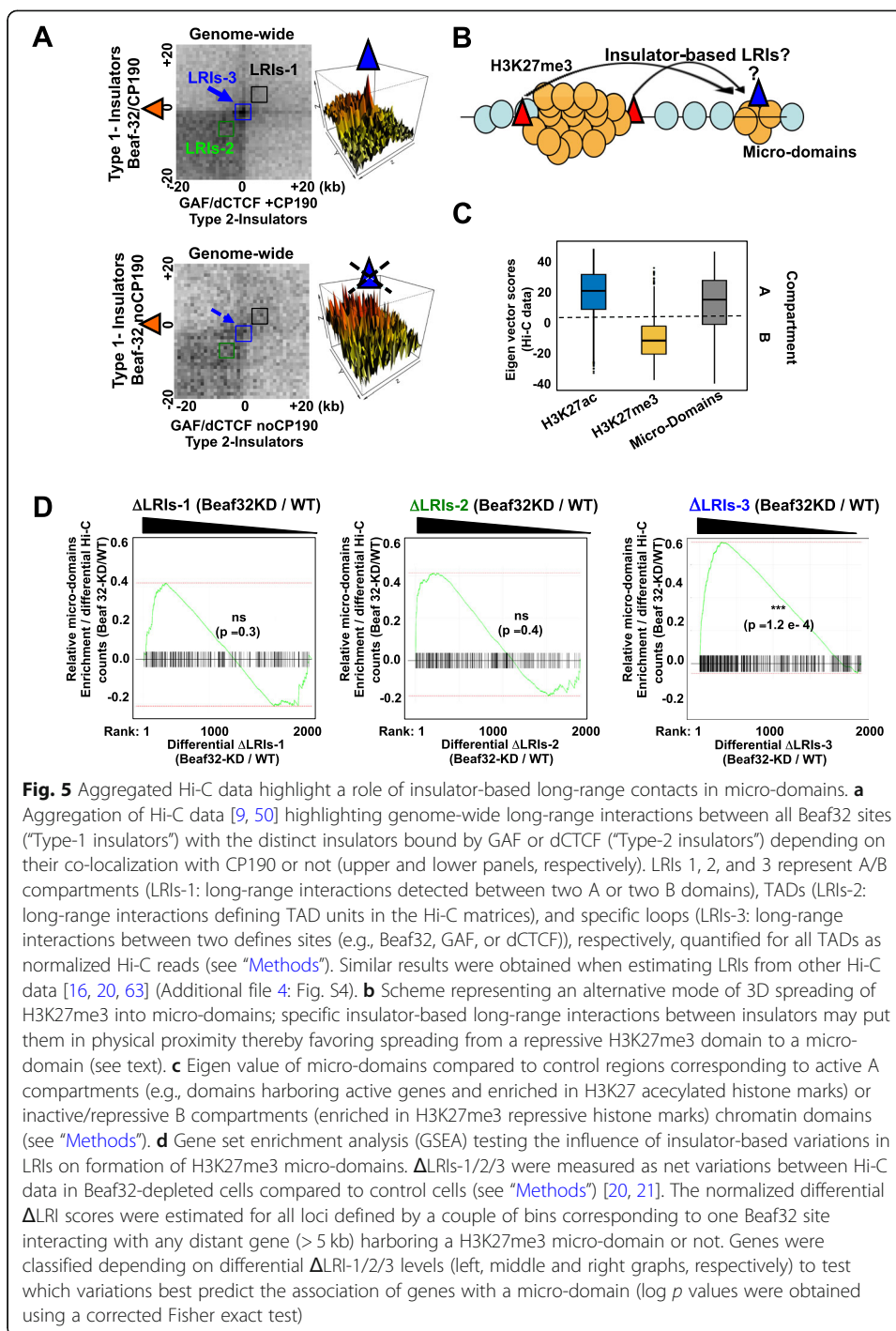
Given the contribution of dCTCF or Beaf32 in TADs [26], our above observations raised the possibility that micro-domains form when TAD strength is low, i.e., when H3K27me3 sites in a repressive TAD may randomly spread onto the flanking euchromatin. In this instance, spreading into micro-domains might reflect TAD “leakiness” or weakness. In contrast, robust TADs might contribute to insulate euchromatin from flanking heterochromatin. We thus evaluated TAD strength using genome-wide aggregation analyses, as developed previously [9, 50] (see “Methods”) depending on protein binding. This analysis shows that Beaf32 binds to the borders of the most robust TADs genome-wide (Fig. 4d), which also involves GAF, dCTCF, and CP190 proteins. We then assessed the influence of Beaf32 depletion on all TADs genome-wide, testing if the probability to detect a micro-domain in the flanking euchromatin domain could be explained by the reduction in TAD strength, as tested using gene set enrichment analysis (GSEA) (Fig. 4e; see “Methods”). Ranking according to the changes in Hi-C counts representing TAD robustness ( $\Delta$ LRI-2; see Fig. 5a) showed no significant correlation with the presence of micro-domains ( $p$  value =1 in both instances). As such, our results suggest that deregulation of TAD robustness by depletion of insulator proteins may not account for the presence of micro-domains.

Insulator binding sites not only bracket TADs, they also define sites with high levels of LRIs in the genome, as evidenced by aggregating Hi-C data onto their binding sites (Fig. 4d; Fig. 5a; see middle region (LRI-3) of the matrix). Such ability to form LRIs with distant sites is notably detected in presence of insulator proteins and cohesin or CP190 co-factors, reflecting how insulators are capable of forming long-range interactions (Fig. 5a, LRI-3). Of note, these are unique features specifically detected with insulator protein sites, and not found for control sites as shown by global assessment of LRIs as a function of protein binding (Fig. 4d, see  $y$ -axis). We thus reasoned that such loops between Beaf32 localized at the borders of repressive domains with distant sites (including GAF sites) inside euchromatin, may represent an alternative possibility accounting for H3K27me3 micro-domains (Fig. 5b). Inspection of the characteristic





Eigen's value reflecting euchromatin/heterochromatin into distinct A/B compartments (see "Methods") showed that micro-domains may not be totally separated from B compartments (Fig. 5c). Thus, an alternative rationale for micro-domain formation may also



**Fig. 5** Aggregated Hi-C data highlight a role of insulator-based long-range contacts in micro-domains. **a** Aggregation of Hi-C data [9, 50] highlighting genome-wide long-range interactions between all Beaf32 sites (“Type-1 insulators”) with the distinct insulators bound by GAF or dCTCF (“Type-2 insulators”) depending on their co-localization with CP190 or not (upper and lower panels, respectively). LRIs 1, 2, and 3 represent A/B compartments (LRIs-1: long-range interactions detected between two A or two B domains), TADs (LRIs-2: long-range interactions defining TAD units in the Hi-C matrices), and specific loops (LRIs-3: long-range interactions between two defines sites (e.g., Beaf32, GAF, or dCTCF)), respectively, quantified for all TADs as normalized Hi-C reads (see “Methods”). Similar results were obtained when estimating LRIs from other Hi-C data [16, 20, 63] (Additional file 4: Fig. S4). **b** Scheme representing an alternative mode of 3D spreading of H3K27me3 into micro-domains; specific insulator-based long-range interactions between insulators may put them in physical proximity thereby favoring spreading from a repressive H3K27me3 domain to a micro-domain (see text). **c** Eigen value of micro-domains compared to control regions corresponding to active A compartments (e.g., domains harboring active genes and enriched in H3K27 acetylated histone marks) or inactive/repressive B compartments (enriched in H3K27me3 repressive histone marks) chromatin domains (see “Methods”). **d** Gene set enrichment analysis (GSEA) testing the influence of insulator-based variations in LRIs on formation of H3K27me3 micro-domains.  $\Delta$ LRIs-1/2/3 were measured as net variations between Hi-C data in Beaf32-depleted cells compared to control cells (see “Methods”) [20, 21]. The normalized differential  $\Delta$ LRI scores were estimated for all loci defined by a couple of bins corresponding to one Beaf32 site interacting with any distant gene (> 5 kb) harboring a H3K27me3 micro-domain or not. Genes were classified depending on differential  $\Delta$ LRI-1/2/3 levels (left, middle and right graphs, respectively) to test which variations best predict the association of genes with a micro-domain (log  $p$  values were obtained using a corrected Fisher exact test)

be due to imperfect 3D compartmentalization of such euchromatic sites from heterochromatin.

To test these hypotheses in details, we first estimated the changes in long-range interactions upon Beaf32-depleted compared to control cells, reflecting either reduction in compartmentalization/phase separation (left:  $\Delta$ LRIs-1) or alternatively in reducing specific loops (right:  $\Delta$ LRIs-3) between insulator sites. We also compared such measures with possible changes in TAD robustness, as previously (middle:  $\Delta$ LRIs-2). All TADs

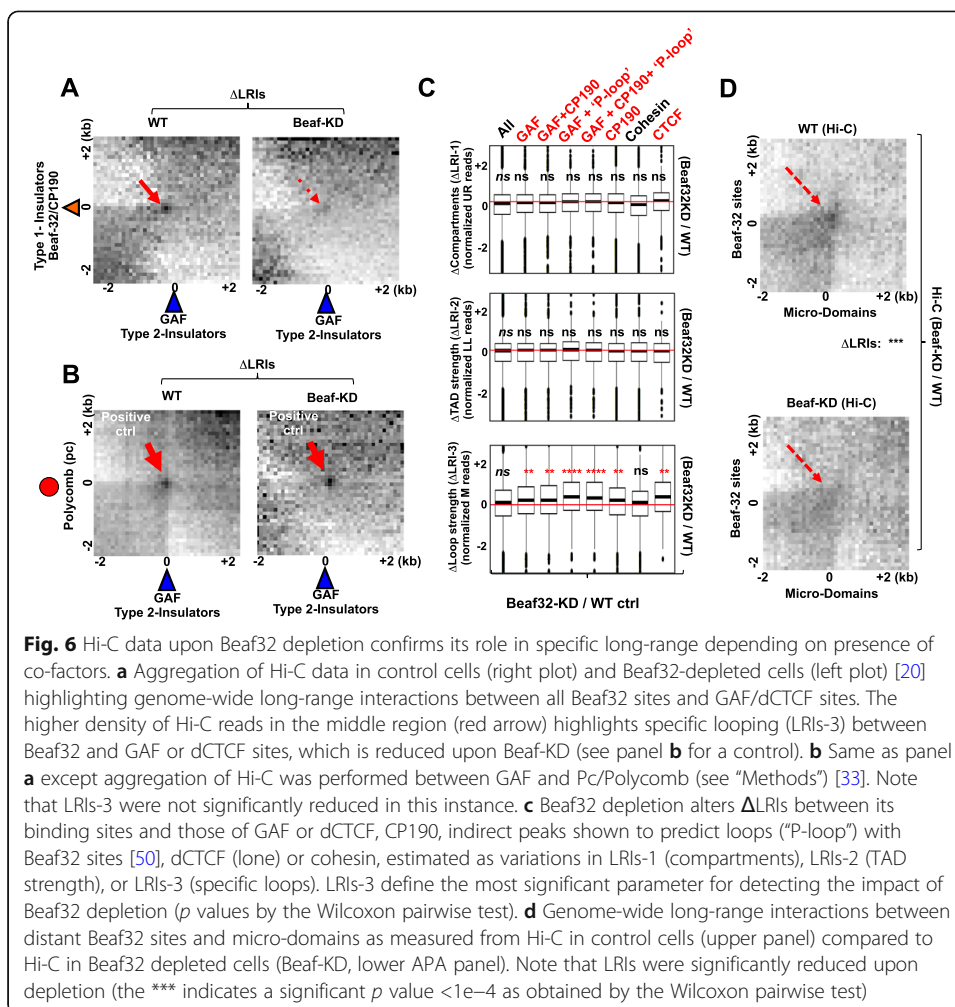
were then ranked according to the variations of each metric (Fig. 5d;  $\Delta$ LRI) [20, 21], providing with three different genome-wide rankings of TADs. The influence of  $\Delta$ LRI parameters was then tested using gene set enrichment analysis (GSEA) to assess which one best predicts the formation of micro-domains (Fig. 5d; see “Methods”). Ranking according to  $\Delta$ LRIs between A compartments (LRI-1) or TAD strength (LRI-2) show no significant prediction of micro-domains. In stark contrast, ranking according to specific LRIs between Beaf32 and distant insulator (GAF/dCTCF) sites ( $\Delta$ LRIs-3) show that  $\Delta$ LRIs-3 significantly predicted micro-domain formation ( $p$  value =  $1.2e-4$ ). Accordingly, distant sites with LRIs not influenced by Beaf32 depletion showed lower chances to harbor micro-domains (Fig. 5d; compare left and right part of the curve). Therefore, specific long-range contacts (LRIs-3) define the best parameter accounting for micro-domain formation, as confirmed using various sources of Hi-C data (Additional file 1: Fig. S4)(see “Methods”). We conclude that the influence of insulator proteins on micro-domains more likely reflect their ability to establish specific long-range interactions rather than a global contribution to insulate domains or to assemble TADs.

#### **Beaf-KD impairs LRIs depending on CP190 at genome-wide levels**

Additional aggregation of Hi-C data highlighted loops/LRIs between Beaf32 and distant GAF/dCTCF/CP190 insulator sites in control cells, which were actually impaired upon Beaf32 depletion (Fig. 6a). In contrast, the loops formed between GAF sites and Polycomb/Pc were retained in depleted cells (Fig. 6b), confirming a specific influence. Most significant reductions in  $\Delta$ LRI-3 were observed in presence of GAF, dCTCF, and CP190 binding indeed (Fig. 6c; Additional file 1: Fig. S5), whereas a systematic influence on LRIs assessing compartments or TAD strength could not be detected (Fig. 6c;  $\Delta$ LRI-1 and  $\Delta$ LRI-2, respectively). Beaf32 indirect peaks that predict loops [50] were enriched among the sites influenced for LRIs with distant Beaf32 sites upon Beaf32 depletion (Fig. 6c; Predicted “P-loop”). Importantly, micro-domains themselves formed significant LRIs with the distant Beaf32 sites, which were impaired by Beaf-KD (Fig. 6d). Therefore, our analyses show that Beaf32 is required for specific LRIs with distant insulators, which may account for the presence of H3K27me3 micro-domains.

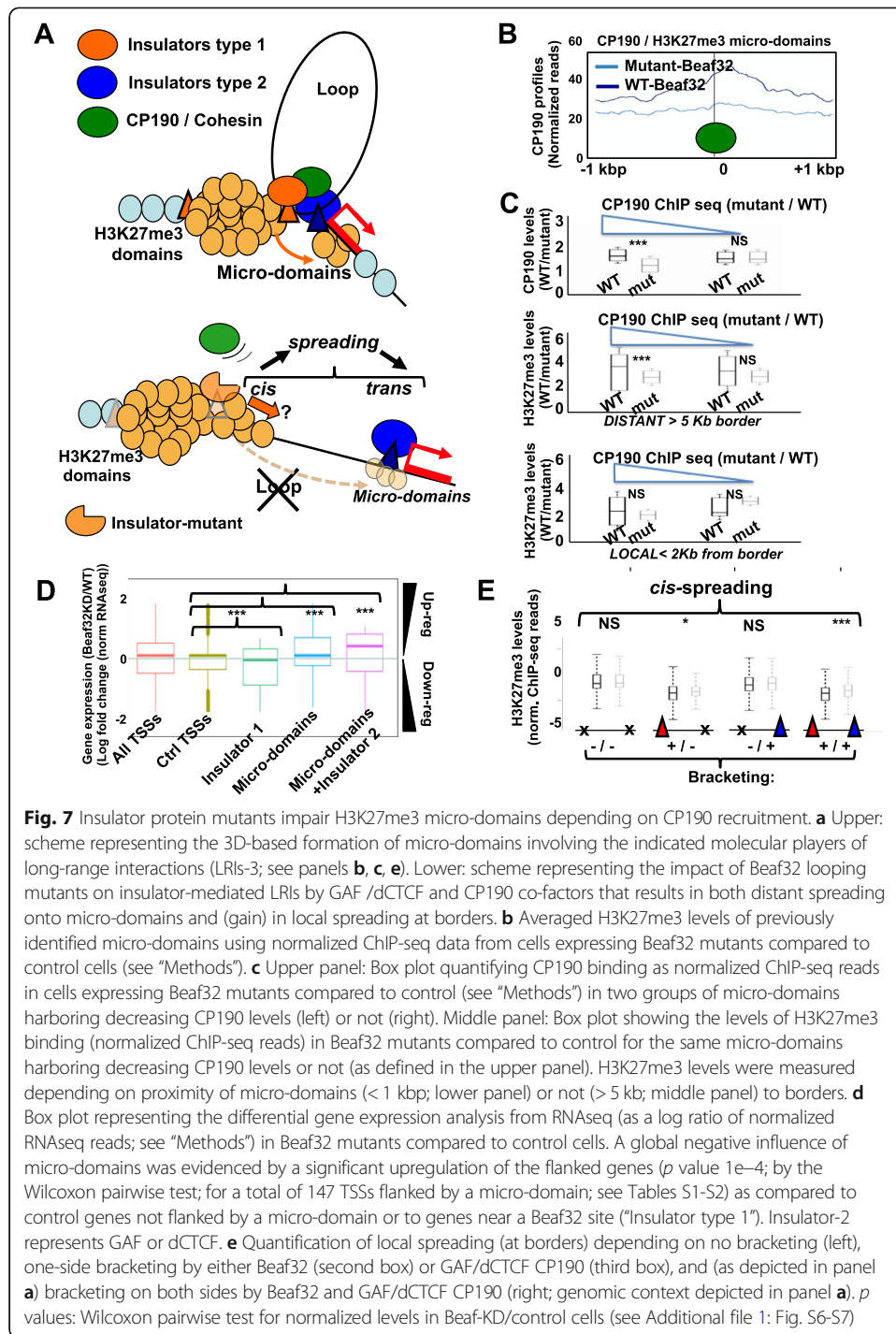
#### **Synthetic insulator proteins impair both CP190 loading and H3K27me3 micro-domains**

We previously designed specific Beaf32 mutants that impaired looping due to their impaired ability to recruit CP190 onto insulator sites (Fig. 7a) [50], in complete agreement with the major role of CP190 in LRIs. We thus asked whether Beaf32 mutants could impair micro-domains due to failure to promote CP190-dependent looping with distant GAF/dCTCF insulators. Beaf32 mutants were expressed as previously [50], followed by ChIP-seq to score H3K27me3 variations systematically compared to control cells (see “Methods”). Enrichment tests showed that of the micro-domains identified in wild-type and that were lost in Beaf32-depleted cells, 55.5% (500/901) were also impaired by looping mutants (Additional file 1: Fig. S6A-E,  $p$  value of  $1e-75$ ), as confirmed by the reproducible decrease in H3K27me3 levels at micro-domains (Additional file 1: Fig. S6F). These results strongly supported the view that looping is a key feature required for micro-domain formation. GAF/dCTCF and CP190 binding sites were enriched in micro-domains harboring the most significant decreases in H3K27me3 levels in presence of mutants (Additional file 1: Fig. S6E, rows 1–2), supporting a



central role of CP190 in micro-domain formation at distant GAF/dCTCF sites. Averaged CP190 profiles were decreased by the mutants (Fig. 7b) concomitantly with the decrease in H3K27me3 levels, for sites where CP190 was also decreased (Fig. 7c; upper and middle box plot, respectively). Of interest, the decreases in CP190 and H3K27me3 were most specific of micro-domains localized away (> 5 kb) from Beaf32 borders (Fig. 7c; middle box plot). In stark contrast, micro-domains flanking Beaf32 heterochromatin borders showed no decrease (Fig. 7c; Additional file 1: Fig. S6E, lower box plot), as such borders are subjected to H3K27me3 spreading locally (Fig. 1c), as confirmed by enrichment tests (Additional file 1: Fig. S6E).

Our work identifies micro-domains of H3K27me3 where heterochromatin components may “use” 3D loops to spread over distant sites (Fig. 7a). Such phenomenon was detected at hundreds of sites depending on specific long-range contacts with the insulator proteins GAF and dCTCF and their shared co-factors CP190 (Additional file 1: Fig. S6-S7), which is specifically impaired by expressing looping mutants as shown (Fig. 7a; lower scheme). Of interest, micro-domains contributed to control the expression of nearby genes that become upregulated upon depletion of Beaf32 (Fig. 7d; *p* value of 1e−4; see also Additional file 1: Fig. S6G). Such genes pertain to specific gene ontologies associated with distant spreading, such as the immune response, cellular homeostasis,



and signal transduction (Additional file 1: Fig. S7), which are distinct from genes being regulated locally at Beaf32/dCTCF insulators [46, 64]. In the latter case, pairing of Beaf32 with GAF conditions the presence of micro-domains and it favors spreading locally (Fig. 7e;  $p$  value  $< 1e-6$ ). Hence, combinations of distinct insulators may be required to detect spreading across the heterochromatin borders. Thus, insulator bracketing may contribute to spreading in 3D, for micro-domain formation, and also for the demarcation of euchromatin from heterochromatin.

Taken altogether, our data support a functional implication of specific LRIs into gene expression programs. We propose that such LRIs contribute to regulate the spreading of H3K27me3 to distant sites, giving rise to micro-domains that participate to insulator-mediated homeostasis of gene expression throughout development (see “Discussion”).

## Discussion

Chromosome compartmentalization in 3D reinforces the demarcation of euchromatin from heterochromatin to control gene expression globally. The identification of micro-domains highlights that heterochromatin can further influence genes through specific long-range contacts in euchromatin. Micro-domain formation requires insulator-based LRIs between heterochromatin TAD borders and micro-domains, which does not contradict compartmentalization principles. The 3D organization of heterochromatin may therefore also influence expression through specific LRIs participating in H3K27me3 deposition locally, in micro-domains, thereby regulating distant euchromatic genes.

Compartmentalization principles may reinforce the global demarcation of TADs [57, 65]. Remarkably, recent high-resolution approaches in single cells have unraveled small “nano-compartments” that define TADs [1, 15]. Nano-compartments thus reflect how higher-order chromatin organization promotes interactions among domains sharing the same epigenetic state (A-A or B-B compartments) and self-interactions within the same folding TADs. Although H3K27me3 nano-compartments are self-maintainable, it remains unclear whether insulator factors, or transcription, participate to the demarcation of these domains from neighboring euchromatin. Our data highlight specific long-range contacts between the borders of nano-compartments with distant sites in euchromatin, through specific insulator-mediated loops. The resulting H3K27me3 micro-domains do not imply that TADs are not strong or that nano-compartments are ill-defined. Actually, LRIs between nano-compartments and nearby euchromatin are poor predictors of micro-domains. Rather, LRIs involved in micro-domain formation specifically involve TAD borders and they depend on insulator proteins. Therefore, micro-domains challenge classic models of insulator-based demarcation of H3K27me3. Rather, insulators do not solely “protect” nearby genes from spreading, as insulator-mediated looping also favors H3K27me3 spreading to distant sites in 3D.

Insulator proteins and additional factors participate to DNA looping between TAD borders [12, 20, 50], thereby contributing to the demarcation of epigenetic domains [30, 40, 65, 66]. Yet a structural role of insulators would predict that their removal alter heterochromatin-euchromatin barriers more systematically than what has been observed [59, 60]. Our work supports the view that the barrier activity of insulators further relies on combinations of insulator factors (Beaf32, GAF, and dCTCF or additional insulator proteins) (Fig. 7), in complete agreement with recent high-resolution Hi-C data [20, 55]. H3K27me3 spreading can actually occur through loops between two distant insulators [67], which may depend on insulator combinations and orientations [55]. Multiple insulators appear thus required for efficient H3K27me3 blocking at borders, while allowing spreading through 3D looping, depending on genomic contexts.

Pioneer work showed that CTCF participated in gene expression homeostasis [68], possibly due to CTCF/cohesion facilitating enhancer-promoter contacts inside TADs.

Our data raise the possibility that a complementary contribution of insulators in expression homeostasis could involve loop-based H3K27me3 deposition. Actually, systematic detection of H3K27me3 throughout developmental stages of *Drosophila* embryos highlights high correlation coefficients in H3K27me3 levels among micro-domains compared to control euchromatin sites (Additional file 1: Fig. S7C-D). As LRIs are transient, persistence of H3K27me3 micro-domains through development may rely on Polycomb-encoded memory and histone-based positive feedback in 1D and in 3D [30, 69, 70]. Similar to Heterochromatin Protein 1-based liquid droplets [31] or to super-enhancers clustering [34], insulator-based micro-domains maintenance may depend on 3D clustering and phase separation principles [71]. Such clustering may serve to counteract high turnover dynamics by erasers/demethylases [40, 71–73]. A sub-fraction of micro-domains overlap with 9.7% of genomic enhancers (Additional file 1: Fig. S7E) that may also be regulated by Polycomb [74]. These observations suggest that co-regulation of H3K27me3 levels in micro-domains further involve shared transcriptional activators to subsets of enhancers.

Micro-domains are not unique in that previous observations identified dispersed, heterochromatin-like H3K9me2/3 islands, which may also depend on 3D organization [75]. Specific long-range interactions are involved in the nucleation of PRC2-mediated repression before allosteric spreading [76], which may involve CTCF-based assembly of TADs or looping [77]. Fly para-segment identity actually relies on specific LRIs at endogenous chromatin boundary insulators [54, 55]. Homeotic gene full repression requires Hox clustering through LRIs for full PRC2-dependent repression during development [78], even though repressive TADs may self-assemble [1]. Further studies should unravel how specific LRIs regulating H3K27me3 at distant genes, depending on dynamics of Pc clusters and co-factors binding at enhancers, TSSs, or insulators, could serve to progressively acquire gene expression homeostasis during development.

## Methods

### Cell culture, insulator mutants, RNAi, and gene expression analyses

Exponentially growing S2 cells were depleted by double-stranded RNAs (dsRNAs) against Beaf32, CP190, or cohesin (*rad21*) compared to mock-depletions (dsRNAs against luciferase) as previously described [50, 59], using the indicated oligos (see Additional file 4: Table S3). Gene expression analyses by RNAseq were performed as previously described [50] on cells depleted of Beaf32 or in cells expressing mutant or WT Beaf32 (GSE52887).

### Chromatin immunoprecipitation analyses and micro-domains detection

Chromatin immunoprecipitations were done as previously described [59] followed by high-throughput sequencing (ChIP-seq) with affinity-purified anti-CP190 antibodies [59] and anti-H3K27me3-specific antibodies (Upstate #07-449) performed in independent replicates in Beaf32-depleted cells and mock-depleted control cells, as well as in 2 × 2 cell replicates expressing mutant- or WT-Beaf32 (see “Methods” for details). For detection of micro-domains, we used all four ChIP-seq datasets analyzed as replicates of control cells compared Beaf32 depleted normalized to input, using normR package version 1.8.0, <https://github.com/your-highness/normR> developed by Helmuth and

Chung for automated normalization and difference calling in ChIP-seq data [61], with the *enrichR* function using 40-bp bin sizes. Robustness of domain detection was tested according to various bin sizes (20 to 200 bp) and selection of domain sizes was performed for domains < 2 kb, based on variations ( $FDR < 5e-2$ ) of the signal between depleted and control conditions (Additional Methods for details).

### 3C/Hi-C experimental and data analysis

All scripts used in this manuscript are available at: [https://github.com/CuvierLab/K27me3\\_mdom\\_spreading/tree/master/src](https://github.com/CuvierLab/K27me3_mdom_spreading/tree/master/src). Hi-C data in both S2 cells and KC cells were normalized using K-R norm function Knight-Ruiz [9]. Aggregation analysis was performed as previously in 1D/2D/3D plots [9, 50, 79] using various sources of high-resolution Hi-C data [16, 20, 21] aggregated onto the H3K27me3 borders of repressive sub-TADs (of median size of 16 kb) depending on presence or absence of the indicated insulator proteins Beaf32, dCTCF, and GAF together with CP190 or cohesin binding by integrating previous ChIP-Seq data [45, 80]. Long-range interactions (LRIs) were estimated as previously described [9, 50] by extracting normalized intensities of the indicated LRIs at specific binding sites in Beaf-KD and control cells [20, 21] (see “Methods”). 3C measurements of LRIs in CP190-depleted, rad21-depleted, or control depletion (dsRNA against luc) conditions as performed by qPCR using TaqMan MGB probes as previously described [50]. Frequency of chimera was estimated in triplicates relatively to products from random ligation estimated using BACs that span the same loci (see Additional information for details).

### Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13059-020-02106-z>.

**Additional file 1: Figure S1.** Genomic contexts of the influence of Beaf32 on H3K27me3 spreading. **Figure S2.** Validation of micro-domains by quantitative PCR analysis of ChIP. **Figure S3.** Regulation of H3K27me3 spreading by Beaf32 and CP190 and depending on genomic contexts. **Figure S4.** Regulation of long-range interactions by insulator proteins. **Figure S5.** Beaf32 depletion affects specific insulator-based LRIs by Beaf32, GAF, dCTCF and CP190. **Figure S6.** Beaf32 looping mutants alter H3K27me3 levels in micro-domains involving regulation of CP190 recruitment. **Figure S7.** Regulation of H3K27me3 trans-spreading may contribute to co-regulate specific gene functions through development.

**Additional file 2: Table S1.** List of genes associated with H3K27me3 micro-domains.

**Additional file 3: Table S2.** List of genes associated with the binding sites of insulator proteins.

**Additional file 4: Table S3.** List of oligos used in this study.

**Additional file 5.** Review history.

### Acknowledgements

We thank Sylvain Foissac, Pascal G. Martin, Frederic Bantignies, Thierry Forné, William Ritchie, Fabian Erdel, and “Gen2i” for suggestions regarding data analysis in 3D and Adélaïde Cucci for help with 3C; the Genotoul for providing access to high-performance computer cluster with INRA and BGI for high-throughput sequencing services.

### Review history

The review history is available as Additional file 5.

### Peer review information

Andrew Cosgrove was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

### Authors' contributions

A.H. did all metagenomic analyses and data integration (Figs. 1, 2, 3, 4, 5, and 6) with the help of D.D. for gene expression analyses (e.g., 3B, S1C, S3A-B, S7B) and St.S. for Hi-C (S4A, 5A-B). C.P. did all ChIP-seq, RNAseq, and ChIP-qPCR with the help of O.F. for ChIP including micro-domain validation (Figure S2B-C). The author(s) read and approved the final manuscript.



**Funding**

This work was supported by the “Fondation pour la Recherche Médicale” (FRM team grant number DEQ20160334940) including a fellowship to A.H. and by a thesis fellowship to D.D. and O.F. (Ministry of Research and Technology), the CNRS (C.P.), and INSERM (St.S. and O.C.).

**Availability of data and materials**

All source codes pertaining to this manuscript are released in compliant with the Open source initiative (OSI) under MIT license and are accessible in GitHub [https://github.com/CuvierLab/H3K27me3\\_micro-Dom\\_spreading](https://github.com/CuvierLab/H3K27me3_micro-Dom_spreading) [81] and the Zenodo doi: [https://zenodo.org/record/3889838#Xut4gpMza\\_u](https://zenodo.org/record/3889838#Xut4gpMza_u) [82]. All ChIP-seq data pertaining to this manuscript were deposited to GEO of NCBI (GSE130211) [83], RNAseq are accessible through GSE52887 [50]. The corresponding lists of H3K27me3 micro-domains are provided in Tables S1-S2, alone or in association with nearby genes, respectively (see Additional information).

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Chromatin Dynamics and Cell Proliferation, Center of Integrative Biology (CBI), Laboratoire de Biologie Moléculaire Eucaryote (LBME), CNRS, Université Fédérale Paul Sabatier de Toulouse, France. F-31000 Toulouse, France. <sup>2</sup>St. Patrick Research Group in Basic Oncology, Laval University Cancer Research Center, Centre Hospitalier Universitaire de Québec City, Québec, QC G1R 3S3, Canada.

Received: 24 July 2019 Accepted: 14 July 2020

Published online: 03 August 2020

**References**

1. Cattoni DI, Cardozo Gizzi AM, Georgieva M, Di Stefano M, Valeri A, Chamousset D, et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat Commun.* 2017;8(1):1753.
2. Rouquette J, Cremer C, Cremer T, Fakan S. Functional nuclear architecture studied by microscopy: present and future. *Int Rev Cell Mol Biol.* 2010;282:1–90.
3. Dekker J. The three ‘C’ s of chromosome conformation capture: controls, controls, controls. *Nat Methods.* 2006;3(1):17–21.
4. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326(5950):289–93.
5. Dekker J, Misteli T. Long-range chromatin interactions. *Cold Spring Harb Perspect Biol.* 2015;7(10):a019356.
6. Dixon JR, Jung J, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature.* 2015;518(7539):331–6.
7. Dostie J, Bickmore WA. Chromosome organization in the nucleus - charting new territory across the hi-Cs. *Curr Opin Genet Dev.* 2012;22(2):125–31.
8. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation Centre. *Nature.* 2012;485(7398):381–5.
9. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 2014;159(7):1665–80.
10. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature.* 2012;489(7414):109–13.
11. Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, et al. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res.* 2013;23(12):2066–77.
12. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell.* 2012;148(3):458–72.
13. Cremer M, Cremer T. Nuclear compartmentalization, dynamics, and function of regulatory DNA sequences. *Genes Chromosomes Cancer.* 2019;58(7):427–36. <https://doi.org/10.1002/gcc.22714>.
14. Marti-Renom MA, Almouzni G, Bickmore WA, Bystrycky K, Cavalli G, Fraser P, et al. Challenges and guidelines toward 4D nucleome data and model standards. *Nat Genet.* 2018;50(10):1352–8.
15. Szabo Q, Jost D, Chang JM, Cattoni DI, Papadopoulos GL, Bonev B, et al. TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci Adv.* 2018;4(2):eaar8082.
16. Eagen KP. Principles of chromosome architecture revealed by hi-C. *Trends Biochem Sci.* 2018;43(6):469–78.
17. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature.* 2013;503(7475):290–4.
18. Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong CT, et al. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell.* 2015;58(2):216–31.
19. Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell.* 2013;153(6):1281–95.
20. Ramirez F, Bhardwaj V, Arrigoni L, Lam KC, Gruning BA, Villaveces J, et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun.* 2018;9(1):189.

21. Wang Q, Sun Q, Czajkowsky DM, Shao Z. Sub-kb Hi-C in *D. melanogaster* reveals conserved characteristics of TADs between insect and mammalian cells. *Nat Commun.* 2018;9(1):188.
22. Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci U S A.* 2014;111(3):996–1001.
23. Yaffe E, Tanay A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat Genet.* 2011;43(11):1059–65.
24. Downen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell.* 2014;159(2):374–87.
25. Mizuguchi T, Fudenberg G, Mehta S, Belton JM, Taneja N, Folco HD, et al. Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature.* 2014;516(7531):432–5.
26. Mourad R, Li L, Cuvier O. Uncovering direct and indirect molecular determinants of chromatin loops using a computational integrative approach. *PLoS Comput Biol.* 2017;13(5):e1005538.
27. Phillips-Cremins JE, Corces VG. Chromatin insulators: linking genome organization to cellular function. *Mol Cell.* 2013;50(4):461–74.
28. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell.* 2015;162(4):900–10.
29. Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, Li CH, et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science.* 2016;351(6280):1454–8.
30. Erdel F. How Communication between nucleosomes enables spreading and epigenetic memory of histone modifications. *Bioessays.* 2017;39(12). <https://doi.org/10.1002/bies.201700053>.
31. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature.* 2017;547(7662):236–40.
32. Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ 3rd, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature.* 2009;461(7265):762–7.
33. Schuettengruber B, Oded Elkayam N, Sexton T, Entrevan M, Stern S, Thomas A, et al. Cooperativity, specificity, and evolutionary stability of polycomb targeting in *Drosophila*. *Cell Rep.* 2014;9(1):219–33.
34. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A phase separation model for transcriptional control. *Cell.* 2017;169(1):13–23.
35. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet.* 2016;17(8):487–500.
36. Cavalli G. Chromosomes: now in 3D! *Nat Rev Mol Cell Biol.* 2014;15(1):6.
37. Fraser J, Ferrai C, Chiariello AM, Schueler M, Rito T, Laudanno G, et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol Syst Biol.* 2015;11(12):852.
38. Heard E, Martienssen RA. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell.* 2014;157(1):95–109.
39. Erdel F, Rippe K. Formation of chromatin subcompartments by phase separation. *Biophys J.* 2018;114(10):2262–70.
40. Cuvier O, Fierz B. Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells. *Nat Rev Genet.* 2017;18(8):457–72.
41. Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell.* 2009;137(7):1194–211.
42. Van Bortle K, Ramos E, Takenaka N, Yang J, Wahi JE, Corces VG. *Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res.* 2012;22(11):2176–87.
43. Mourad R, Cuvier O. TAD-free analysis of architectural proteins and insulators. *Nucleic Acids Res.* 2018;46(5):e27.
44. Negre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, et al. A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.* 2010;6(1):e1000814.
45. Li J, Gilmour DS. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J.* 2013;32(13):1829–41.
46. Emberly E, Blattes R, Schuettengruber B, Hennion M, Jiang N, Hart CM, et al. BEAF regulates cell-cycle genes through the controlled deposition of H3K9 methylation marks into its conserved dual-core binding sites. *PLoS Biol.* 2008;6(12):2896–910.
47. Kellum R, Schedl P. A position-effect assay for boundaries of higher order chromosomal domains. *Cell.* 1991;64(5):941–50.
48. Cai H, Levine M. Modulation of enhancer-promoter interactions by insulators in the *Drosophila* embryo. *Nature.* 1995;376(6540):533–6.
49. Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell.* 2012;48(3):471–84.
50. Liang J, Lacroix L, Gamot A, Cuddapah S, Queille S, Lhoumaud P, et al. Chromatin immunoprecipitation indirect peaks highlight long-range interactions of insulator proteins and pol II pausing. *Mol Cell.* 2014;53(4):672–81.
51. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell.* 2008;132(3):422–33.
52. Merkschlager M, Nora EP. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu Rev Genomics Hum Genet.* 2016;17:17–43.
53. Vogelmann J, Valeri A, Guillou E, Cuvier O, Nollmann M. Roles of chromatin insulator proteins in higher-order chromatin organization and transcription regulation. *Nucleus.* 2011;2(5):358–69.
54. Fedotova A, Aoki T, Rossier M, Mishra RK, Clendinen C, Kyrchanova O, et al. The BEN domain protein insensitive binds to the Fab-7 chromatin boundary to establish proper segmental identity in *Drosophila*. *Genetics.* 2018;210(2):573–85.
55. Kyrchanova O, Mogila V, Wolle D, Deshpande G, Parshikov A, Cleard F, et al. Functional dissection of the blocking and bypass activities of the Fab-8 boundary in the *Drosophila* Bithorax complex. *PLoS Genet.* 2016;12(7):e1006188.
56. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* 2009;19(1):24–32.
57. Ho JW, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, et al. Comparative analysis of metazoan chromatin organization. *Nature.* 2014;512(7515):449–52.
58. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, et al. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature.* 2011;471(7339):480–5.

59. Lhoumaud P, Hennion M, Gamot A, Cuddapah S, Queille S, Liang J, et al. Insulators recruit histone methyltransferase dMes4 to regulate chromatin of flanking genes. *EMBO J.* 2014;33:1599–613.
60. Schwartz YB, Linder-Basso D, Kharchenko PV, Tolstorukov MY, Kim M, Li HB, et al. Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Res.* 2012;22(11):2188–98.
61. Kinkley S, Helmuth J, Polansky JK, Dunkel I, Gasparoni G, Frohler S, et al. reChIP-seq reveals widespread bivalency of H3K4me3 and H3K27me3 in CD4(+) memory T cells. *Nat Commun.* 2016;7:12514.
62. Rowley MJ, Corces VG. The three-dimensional genome: principles and roles of long-distance interactions. *Curr Opin Cell Biol.* 2016;40:8–14.
63. Rowley MJ, Corces VG. Minute-made data analysis: tools for rapid interrogation of hi-C contacts. *Mol Cell.* 2016;64(1):9–11.
64. Bushey AM, Ramos E, Corces VG. Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev.* 2009;23(11):1338–50.
65. Jost D, Carrivain P, Cavalli G, Vaillant C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res.* 2014;42(15):9553–61.
66. Sexton T, Yaffe E. Chromosome folding: driver or passenger of epigenetic state? *Cold Spring Harb Perspect Biol.* 2015;7(2):a018721.
67. Comet I, Schuettengruber B, Sexton T, Cavalli G. A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc Natl Acad Sci U S A.* 2011;108(6):2294–9.
68. Ren G, Jin W, Cui K, Rodriguez J, Hu G, Zhang Z, et al. CTCF-mediated enhancer-promoter interaction is a critical regulator of cell-to-cell variation of gene expression. *Mol Cell.* 2017;67(6):1049–58. e6.
69. Bantignies F, Cavalli G. Polycomb group proteins: repression in 3D. *Trends Genet.* 2011;27(11):454–64.
70. Cheutin T, Cavalli G. Polycomb silencing: from linear chromatin domains to 3D chromosome folding. *Curr Opin Genet Dev.* 2014;25:30–7.
71. Erdel F, Greene EC. Generalized nucleation and looping model for epigenetic memory of histone modifications. *Proc Natl Acad Sci U S A.* 2016;113(29):E4180–9.
72. Audergon PN, Catania S, Kagansky A, Tong P, Shukla M, Pidoux AL, et al. Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. *Science.* 2015;348(6230):132–5.
73. Raganathan K, Jih G, Moazed D. Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science.* 2015;348(6230):1258699.
74. Erceg J, Pakozdi T, Marco-Ferreres R, Ghavi-Helm Y, Girardot C, Bracken AP, et al. Dual functionality of cis-regulatory elements as developmental enhancers and Polycomb response elements. *Genes Dev.* 2017;31(6):590–602.
75. Li Q, Tjong H, Li X, Gong K, Zhou XJ, Chiolo I, et al. The three-dimensional genome organization of *Drosophila melanogaster* through data integration. *Genome Biol.* 2017;18(1):145.
76. Oksuz O, Narendra V, Lee CH, Descostes N, LeRoy G, Raviram R, et al. Capturing the onset of PRC2-mediated repressive domain formation. *Mol Cell.* 2018;70(6):1149–62. e5.
77. Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science.* 2015;347(6225):1017–21.
78. Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, et al. Polycomb-dependent regulatory contacts between distant Hox loci in *Drosophila*. *Cell.* 2011;144(2):214–26.
79. Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, et al. Evolutionarily conserved principles predict 3D chromatin organization. *Mol Cell.* 2017;67(5):837–52. e7.
80. Zouaz A, Auradkar A, Delfini MC, Macchi M, Barthez M, Ela Akoa S, et al. The Hox proteins Ubx and AbdA collaborate with the transcription pausing factor M1BP to regulate gene transcription. *EMBO J.* 2017;36(19):2887–906.
81. Heurteau A, Perrois C, Depierre D, Fosseprez O, Humbert J, Schaak S, et al. Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes. Github. 2020; [https://github.com/CuvierLab/H3K27me3\\_micro-Dom\\_spreading](https://github.com/CuvierLab/H3K27me3_micro-Dom_spreading). Accessed 11 June 2020.
82. Heurteau A, Perrois C, Depierre D, Fosseprez O, Humbert J, Schaak S, et al. Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes. Zenodo. 2020; <https://zenodo.org/record/3889838> - XuvFkZMza\_s. Accessed 11 June 2020.
83. Heurteau A, Perrois C, Depierre D, Fosseprez O, Humbert J, Schaak S, et al. Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes. Datasets. Gene Expression Omnibus. 2019; <https://www.ncbi.nlm.nih.gov/insb.bib.cnrs.fr/geo/query/acc.cgi?acc=GSE130211>. Accessed 28 Apr 2020.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

