# The *Drosophila* speciation factor HMR localizes to genomic insulator sites

**Thomas Andreas Gerland**[1,2], **Bo Sun**[1], **Pawel Smialowski**[1,3], **Andrea Lukacs**[1], **Andreas Walter Thomae**[1,4], **Axel Imhof**[1,2]*

**1** Biomedical Center, Histone Modifications Group, Department of Molecular Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany, **2** Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität München, Munich, Germany, **3** Biomedical Center, Core Facility Computational Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany, **4** Biomedical Center, Core Facility Bioimaging, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

* imhof@lmu.de

## Abstract

Hybrid incompatibility between *Drosophila melanogaster* and *D. simulans* is caused by a lethal interaction of the proteins encoded by the *Hmr* and *Lhr* genes. In *D. melanogaster* the loss of HMR results in mitotic defects, an increase in transcription of transposable elements and a deregulation of heterochromatic genes. To better understand the molecular mechanisms that mediate HMR's function, we measured genome-wide localization of HMR in *D. melanogaster* tissue culture cells by chromatin immunoprecipitation. Interestingly, we find HMR localizing to genomic insulator sites that can be classified into two groups. One group belongs to *gypsy* insulators and another one borders HP1a bound regions at active genes. The transcription of the latter group genes is strongly affected in larvae and ovaries of *Hmr* mutant flies. Our data suggest a novel link between HMR and insulator proteins, a finding that implicates a potential role for genome organization in the formation of species.

## Introduction

Biodiversity is the result of the emergence and the extinction of species. New species form by pre- and post-zygotic isolation mediated by genetic incompatibility [1]. One of the best characterized examples of hybrid incompatibility is the gene pair *Hybrid male rescue* (*Hmr*) and *Lethal hybrid rescue* (*Lhr*). *Hmr* and *Lhr* cause hybrid incompatibility between the closely related fly species *Drosophila melanogaster* and *D. simulans*. *Hmr* diverged in both *Drosophila* sibling species under positive selection [2]. HMR and LHR from both species interact physically and localize predominantly to centromeric regions [3]. A reduction of HMR expression results in a misregulation of transposable elements, satellite DNAs and heterochromatic genes [3–5]. The major difference between HMR and LHR in *D. melanogaster* and *D. simulans* is their substantial difference in protein amounts [3,6], which has been proposed to result in a lethal gain of function in male hybrids [3]. High levels of HMR and

LHR in hybrids and overexpression of these proteins in pure species lead to an increased number of binding sites of the complex [3]. Such spreading phenomena based on protein amount have been observed for several chromatin-associated complexes such as the dosage compensation complex [7,8], the polycomb complex [9] or components of pericentromeric heterochromatin [10,11]. In most cases, the precise mechanisms for targeting and spreading are not fully understood. Interestingly, several of the components involved in these processes show signs of adaptive evolution and differ substantially even in very closely related organisms [12–14]. This observation has spurred a model of a dynamic genome that drives the adaptive evolution of chromatin-associated factors [15].

Eukaryotic genomes of closely related species differ mostly in the amount and sequence of repetitive DNA [16–18]. This DNA is often derived from transposable elements, which are highly mutagenic and are therefore under tight transcriptional control by the cellular machinery. During evolution transposons or transposon-derived sequences occasionally adopted structural or novel *cis*-regulatory functions, thereby contributing to the evolution of new, species-specific, phenotypic traits [19–21]. Genomic insulators are a particular class of such novel, fast evolving, *cis*-regulatory elements that show signs of transposon ancestry [22,23]. A strong expansion of these elements is observed in arthropods, which also experienced a successive gain in the number of insulator binding proteins during evolution [24]. In fact, the *Drosophila* genome harbours a large variety of insulator proteins such as CTCF, BEAF-32, Su(Hw), Mod(mdg4) and CP190, which all affect nuclear architecture [25]. Different *Drosophila* species underwent multiple genomic rearrangements and transposon invasions [26,27], which presumably resulted in an adaptive response of regulatory DNA binding factors to maintain spatial and temporal gene expression. For example, binding sites for the insulator proteins BEAF-32 and CTCF show a high degree of variability when compared among very closely related species [26,27]. The gain of new insulator sites is associated with chromosome rearrangements, new born genes and species-specific transcription regulation [19,23]. Similar to insulator proteins, which tend to cluster in specific nuclear regions [28], the speciation factor HMR clusters at centromeres or pericentromeric regions in diploid cells [3,6] but is also detected at distinct euchromatic regions along the chromosome arms in polytene chromosomes [3]. A unifying feature for many of these sites is their close proximity to binding sits of the Heterochromatin Protein 1 (HP1a), a HMR interactor and a well-characterized heterochromatic mark.

Various studies describe HMR's localization to heterochromatin, but the molecular details on HMR's binding sites and its recruitment to these sites are not well understood. To get new insights into HMR's association to chromatin, we measured HMR's genome-wide localization by chromatin immunoprecipitation (ChIP) in the *D. melanogaster* embryonic S2 cell line. We demonstrate an extensive colocalization of HMR with a subset of insulator sites across the genome. HMR's binding to genomic *gypsy* insulators, which constitute the major group of its binding sites, is dependent on the residing insulator protein complex. In a second group, HMR borders heterochromatin together with the insulator protein BEAF-32. In agreement with previous low-resolution techniques in cell lines and fly tissue [3], these binding sites are enriched at pericentromeric regions, the cytological region 31 on the 2nd chromosome and the entire 4th chromosome. At most of these sites, HMR associates to the promoters of actively transcribed genes. Interestingly, these genes code for transcripts that have been reported to be downregulated in *Hmr* mutant larvae and ovaries. Altogether, our data provide evidence for a functional link between HMR and insulator proteins, which potentially results in hybrid incompatibilities due to the adaptive evolution of these genome-organizing complexes.

## Materials and methods

### Cell culture and RNAi

*D. melanogaster* S2-DRSC cells were obtained from the DGRC and grown at 26˚C in Schneider's *Drosophila* medium (Invitrogen) supplemented with 10% fetal calf serum and antibiotics (100 units/mL penicillin and 100 µg/mL streptomycin).

For RNAi experiments cells were incubated in serum-free medium containing 10 mg/mL dsRNA. After 1 hr of incubation, the serum-containing medium was supplied. Samples were taken after 7 days. The dsRNA was prepared using the MEGAScript T7 Transcription Kit (Thermo Fisher Scientific) following the manufacturers instructions with primers listed in S1 Table.

### Chromatin immunoprecipitation, Real-Time PCR and sequencing

For chromatin immunoprecipitation (ChIP) cells were crosslinked with 1% formaldehyde for 5 min at room temperature. Upon cell lysis, protease inhibitors and proteasome inhibitor MG-132 (Enzo Life Sciences) were applied. The chromatin was isolated and sheared with adaptive focused acoustics (Covaris) to an average size of 200 base pair (bp). For each ChIP reaction, chromatin isolated from $1–2 \times 10^6$ cells was incubated with following antibodies precoupled to Protein A/G Sepharose: rat anti-HMR 2C10 (RRID: AB2569849) [3] with rabbit IgG anti-rat IgG (RRID: AB2339804), mouse anti-HP1a C1A9 (RRID: AB528276) [29], rabbit anti-H3 (RRID: AB302613), rabbit anti-H3K9me3 (RRID:AB2532132) and mouse anti-FLAG (RRID: AB262044). Real-Time PCR was performed with Fast SYBR Green master mix (Applied Biosystems) using a LightCycler 480 II (Roche). For deep sequencing, all libraries were prepared using MicroPlex (Diagenode) or NEBNext (NEB) Library Preparation kit and single-end, 50 bp sequenced with the Illumina HiSeq2000. An overview of all ChIP-Seq samples used and the number of uniquely aligned sequence reads is provided as S2 Table. A list of HMR peaks used for further analyses is provided as S3 Table. All sequencing data are publicly available as described below.

### Data analysis

The raw reads were aligned to the *D. melanogaster* genome assembly (dm3) using Bowtie 2.2.6 with unique mapping criteria and exclusion of chromosome Uextra [30]. The raw read quality was accessed using FASTQC 11.5 [31] and read filtering was performed using FastX 0.0.13 [32]. Sequencing tracks were visualized using IGB [33] and IGV [34] genome viewers. Peak calling, motif search and peak annotation were performed using HOMER 4.8 with peak size of 200 bp [35] and ChIPseeks implementation of HOMER [36]. For downstream analysis, peaks identified in two out of three biological replicates were taken. Downstream analysis steps were performed using Python and R and parts of data preprocessing was done using ChipPeakAnno [37]. For repeat analysis, reads from ChIP-Seq experiments were mapped to RepBase version 19.10 [38] using Bowtie [30]. Only unique reads were kept for analysis. For each repetitive element, the log2 fold change was calculated. Following genome-wide binding data sets derived from S2 cells (unless stated otherwise) were used: CP190, Su(Hw), CTCF and mod(mdg4) from GEO GSE41354 [39], BEAF-32 from GEO GSE32815 [40]. RNA expression data for untreated S2 cells was taken from GEO GSE46020. For *D. melanogaster* larvae and ovaries, RNA-Seq data were taken from NCBI BioProject PRJNA236022 [4] and analyses were performed with cuffdiff 2 [41]. An extended description of the bioinformatics tools and methods used is provided in S1 Methods.

## Western blot analysis

Samples were boiled in loading buffer, separated on SDS-PAGE gels (Serva), processed for western blot using standard protocols and detected using rat anti-HMR 2C10 (1:20) (RRID: AB2569849), rabbit anti-CP190 (RRID: AB2615894) [42], rabbit anti-H3K9me3 (1:2000) (RRID:AB2532132) and mouse anti-Tubulin (1:800) (RRID: AB2241150) antibodies. Secondary antibodies included sheep anti-mouse (1:5000) (RRID: AB772210), goat anti-rat (1:5000) (RRID: AB772207), donkey anti-rabbit (1:5000) (RRID: AB772206) coupled to horseradish peroxidase.

## Data access

ChIP-Seq data from this study are publicly available at NCBI GEO (GSE86106).

## Results

### Genome-wide binding map of HMR in *D. melanogaster*

Immunohistological studies revealed a binding of HMR to centromeric or pericentromeric regions in diploid cells and to several euchromatic and telomeric regions in polytene chromosomes [3,6]. However, detailed information on HMR's binding to chromatin was so far lacking. To better understand the molecular mechanisms that govern HMR's binding within the genome, we mapped the genomic binding sites of HMR in cultured *D. melanogaster* S2 cells. We used a highly specific monoclonal antibody against HMR [3] to purify associated chromatin followed by next generation sequencing (ChIP-Seq) and derived a set of 794 HMR binding sites, which were present in at least two out of three biological replicates (Fig 1A). A composite plot of all HMR binding sites found in the genome revealed a sharp peak of HMR binding with a width of approximately 200 nucleotides, which is reminiscent of sequence specific transcription factors (Fig 1B). To validate the identified HMR binding sites, we applied multiple strategies. First, we measured enrichment of the HMR binding sites in ChIP experiment using an anti-FLAG antibody, an epitope that is not expressed in wild type cells (Fig 1B and S1A Fig). Second, we performed RNAi knock-down experiments to reduce HMR protein level and compared the enrichment of HMR between HMR RNAi treated cells and Control (Ctrl) RNAi treated cells. Although we observe an overall reduction of HMR binding at most HMR peaks (S1B Fig), we rarely see a complete loss of binding despite the high efficiency of the HMR knock-down. This apparent discrepancy suggests that chromatin-bound HMR is rather resistant towards a RNAi-mediated removal. The existence of such RNAi-resistant binding sites in ChIP experiments has been observed before and was attributed to high-affinity binding sites [43] or an incomplete removal of the chromatin-bound factors. Third, we used the CRISPR/ *cas9* system to edit the HMR locus in S2 cells such that the cell line exclusively expresses an HMR allele, which carries a double FLAG-tag at the C-terminus. ChIP-qPCR using HMR and FLAG antibody in wild type and HMR-Flag$_2$ expressing cells showed specific and reproducible enrichment of HMR at selected HMR binding sites (S1C Fig).

We find HMR binding sites on all chromosomes and distributed along the whole chromosome arms with a marked increase in peak density at pericentromeric regions and at the 4th chromosome where we also observe a higher density of binding sites for HP1a, a known interaction partner of HMR [3,4] (Fig 1C and S1D Fig). Unfortunately, the centromeric regions are not present in the current *Drosophila* genome assembly, preventing read mapping and analysis in this area of the genome. However, the increased number of peaks at pericentromeric regions (S1D Fig) is consistent with the strong centromeric HMR signal we previously observed when staining S2 cells with an anti-HMR antibody [3]. Besides the pericentromeric region, we also
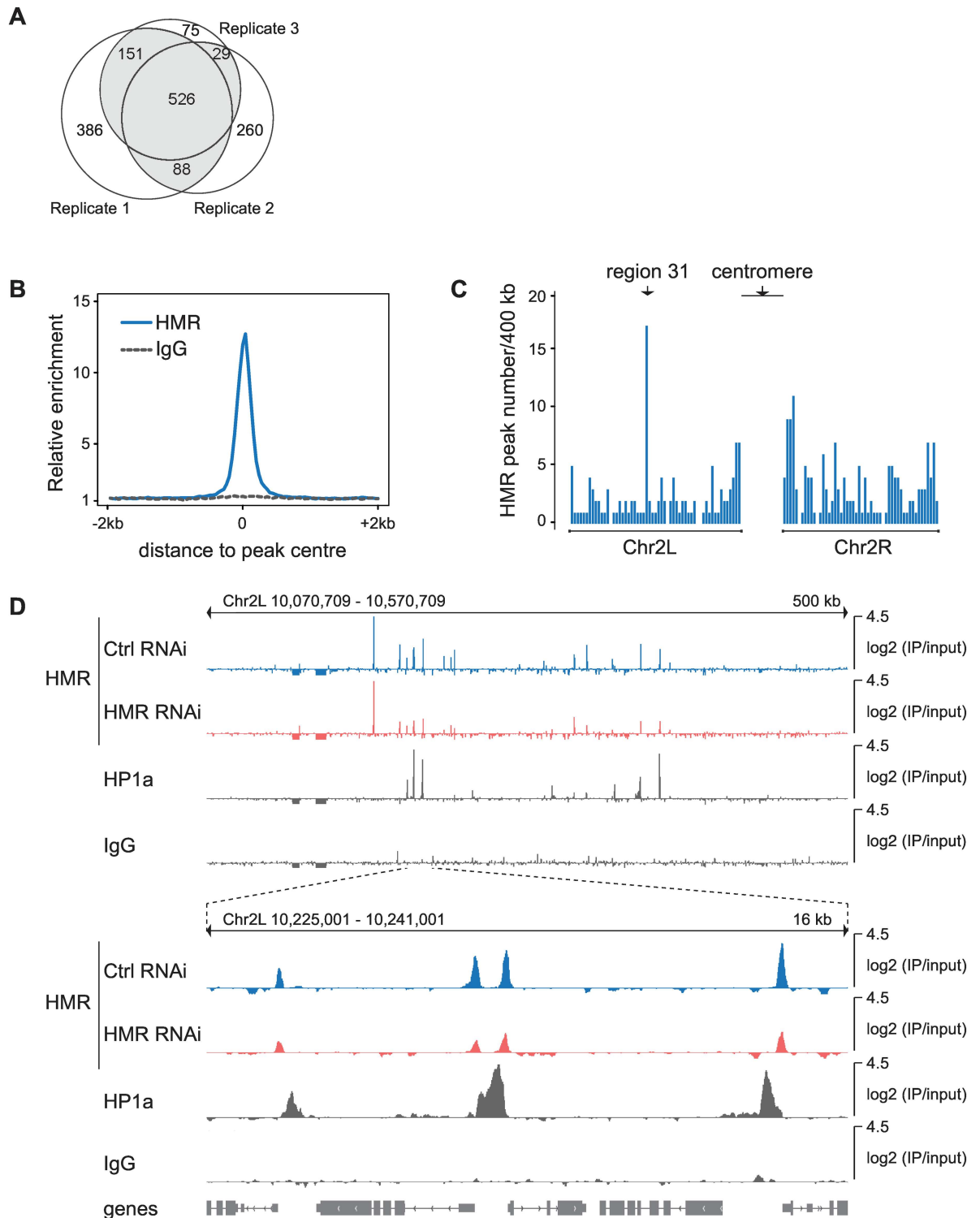
**Fig 1. Identification of HMR binding sites in *D. melanogaster* S2 cells. (A)** Venn diagram of HMR peaks showing the number of peaks identified in three independent biological replicates. Peaks identified in at least two out of three replicates were used for further analysis and are highlighted in *grey*. **(B)** Composite analysis of HMR and control IgG (anti-FLAG) ChIP signals at genomic HMR peak positions. **(C)** Histogram of HMR peak density across the left arm (2L) and right arm (2R) of the 2nd chromosome. The cytological region 31 and centromere-proximal regions are indicated **(D)** Genome browser view of HMR, HP1a and control IgG (anti-FLAG) ChIP signals at region 31. HMR ChIP signals obtained upon knock-down using control RNAi and HMR RNAi are shown with the same amplitude.

observe a strong clustering of HMR peaks at the cytological region 31 on the left arm of the second chromosome where we also see HMR binding in polytene chromosomes [3]. Interestingly, HMR binding sites within these regions do not completely overlap with HP1a bound regions but rather localize at their edges (Fig 1D).

## HMR binding sites largely overlap with genomic insulator sites

We next asked whether HMR binding sites are enriched for specific DNA sequence motifs. A motif analysis of HMR-bound regions revealed three DNA sequence motifs that were significantly enriched and present in up to 26% of all HMR peaks (Fig 2A). These motifs are highly related to the recognition motifs of the insulator DNA binding proteins Su(Hw) and BEAF-32 (S2A Fig), both containing a zinc-finger DNA-binding domain [44,45], suggesting that HMR binds to insulator regions. Indeed, we observe a substantial overlap of our HMR binding profiles with the published ones of known insulator proteins such as CP190, Mod(mdg4), Su(Hw), CTCF and BEAF32 [39,43] (Fig 2B and S2B Fig). Insulator binding sites can be subclassified depending on their composition of known insulator proteins [43]. One of the best characterized family of insulators are derived from the *gypsy* retrotransposon and are strongly bound by Su(Hw), Mod(mdg4) and CP190 [46–48]. Consistent with the strong enrichment of Su(Hw)-recognition motifs in the binding sites of HMR, we find about half of all HMR sites belonging to this *gypsy*-like family of insulators (Fig 2B). However, only 7% of all Su(Hw) binding sites and 11% of *gypsy*-like elements classified as bound by Su(Hw), Mod(mdg4) and CP190 are also bound by HMR.

Given the extensive colocalization of HMR with non-repetitive *gypsy*-like insulators (Fig 2C) and the effect of a *Hmr* mutations on the expression of retrotransposons [3,4], we wondered whether HMR is also enriched at repetitive DNA. We therefore mapped sequences obtained from our ChIP-Seq experiments using anti-HMR and anti-HP1a antibodies as well as published binding profiles for Su(Hw), Mod(mdg4)2.2 and CP190 [39] against the RepBase repeat database [38]. In agreement with previous studies we observe a strong enrichment of HP1a at the centromeric heterochromatin-associated Dodeca satellite (DMSAT6) [49] and the transposable elements Rt1a and Rt1b (DMRT1A, DMRT1B) [50] (S2C Fig). In contrast to HP1a, the only repetitive elements that show a substantial enrichment for HMR are the retrotransposons *gypsy*, and *gtwin* (Fig 3A). At these elements HMR binds together with Su(Hw), Mod(mdg4) and CP190 to the 5' insulator region (Fig 3B and S2D Fig).

A key element for the formation of insulator complexes at *gypsy*-like elements is the presence of the CP190 adaptor protein. A reduction of CP190 levels has been shown to strongly affect binding of insulator proteins to these elements but not to others [43]. To test whether CP190 also impacts the binding of HMR to *gypsy*-like binding sites, we performed RNAi knock-down experiments to reduce CP190 protein level (Fig 4A) and measured HMR binding. Strikingly, we observe a substantial reduction of HMR binding only for the *gypsy*-like group of binding sites (Fig 4B and 4C), suggesting that HMR's binding to the *gypsy*-like insulator class is indeed dependent on CP190. A HMR RNAi knock-down in contrast affects HMR binding equally in both classes (Fig 4D and S3 Fig). As insulator sites are known to contain less nucleosomes [51], nucleosome occupancy can serve as a proxy for insulator complex integrity at these sites [43]. We therefore performed a Histone H3 ChIP upon CP190 RNAi knock-down to monitor changes in insulator complex integrity [52]. Consistent with the importance of CP190 for maintaining the *gypsy* insulator, nucleosome occupancy only increases in the *gypsy*-like HMR binding sites (Fig 4C). Taken together, these results demonstrate an extensive colocalization of HMR with genomic insulator proteins, which play an important role in mediating its binding to chromatin.
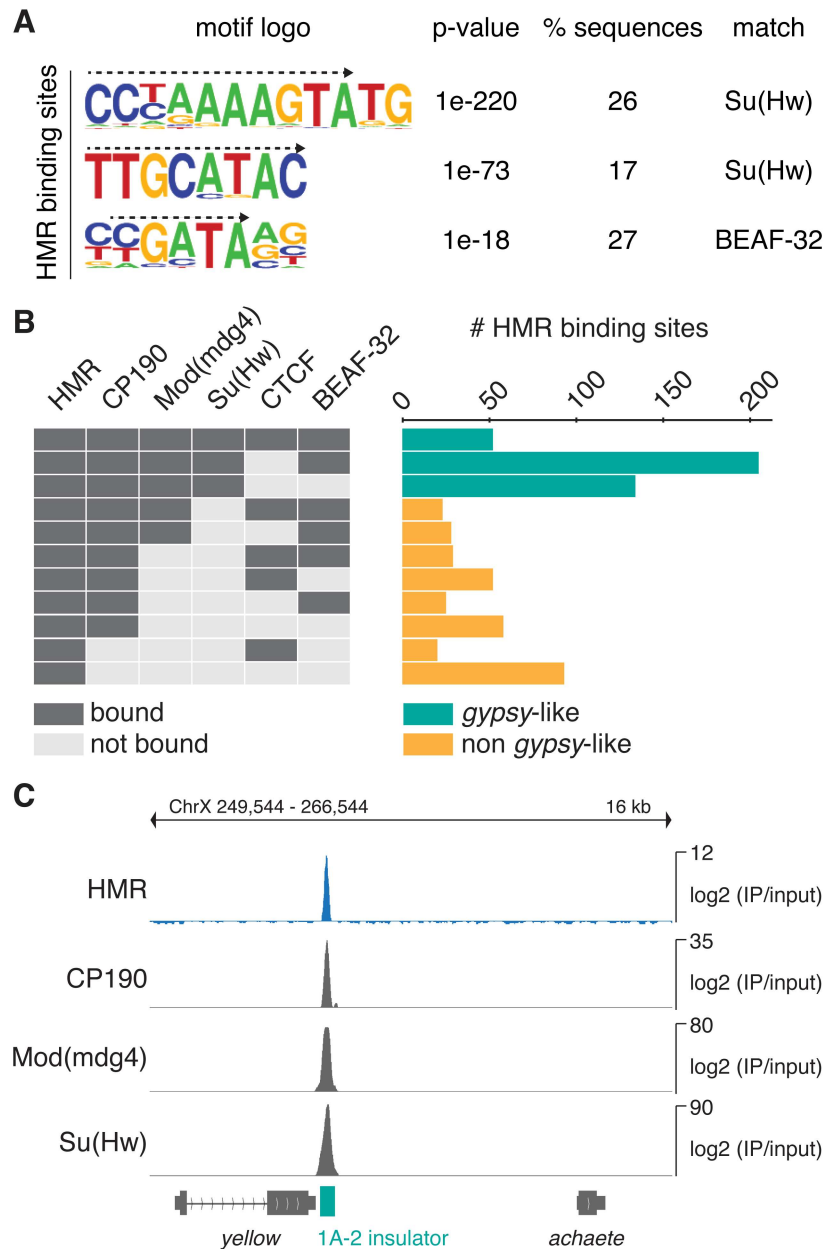
**Fig 2. HMR localizes to genomic insulators and the *gypsy* transposon. (A)** Sequence motifs identified within HMR peak regions. The corresponding motif logo, p-value of enrichment, percentage of regions with this motif and putative binding factors are indicated. Dashed arrows mark the sequence that matches the published binding sites of Su(Hw) and BEAF-32 (see also S2A Fig) **(B)** Peak overlap of HMR with peaks of the insulator proteins CP190, Mod(mdg4), Su(Hw), CTCF [39] and BEAF-32 [40]. The number of HMR peaks is indicated depending on their colocalization with known boundary factors. Groups with less than 11 members are not displayed. Su(Hw)-containing *gypsy*-like groups are depicted in *green*, non *gypsy*-like groups in *orange*. Combinations that contain less than ten HMR peaks are not shown. **(C)** Genome browser view of the Su(Hw) binding region 1A-2. ChIP signals of HMR and known *gypsy* binding factors are shown. The 1A-2 insulator is highlighted in *green*.
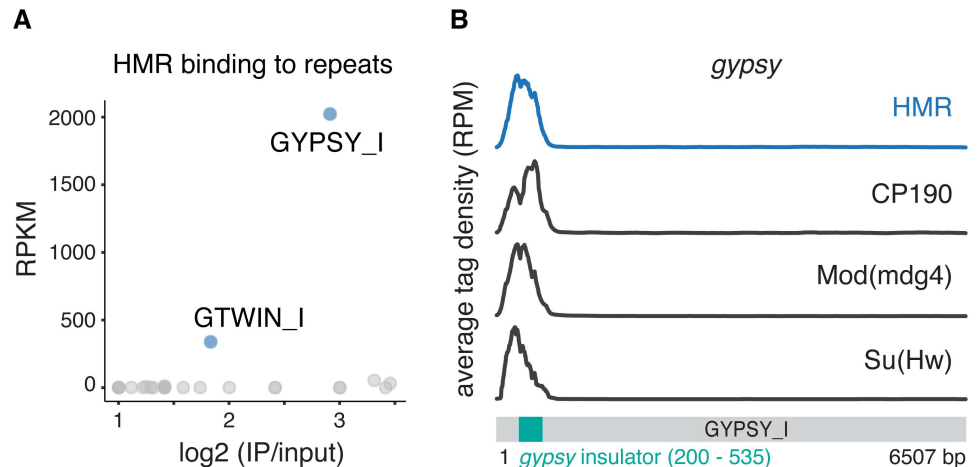
doi:10.1371/journal.pone.0171798.g002

**Fig 3. HMR localisation to repetitive elements. (A)** HMR ChIP tag enrichment at repetitive DNA elements. To identify enriched sequences the enrichment (log2-fold) over input is plotted against the RPKM of an individual repeat sequence from RepBase. Repeats with less than 2-fold enrichment are not displayed. **(B)** ChIP tag density of HMR and the *gypsy*-insulator proteins CP190, Mod(mdg4), Su(Hw) [39] across the repetitive *gypsy* retrotransposon sequence. The *gypsy* insulator sequence at the 5' end is highlighted in *green*.

## HMR borders HP1a domains at active promoters

Although a large portion of HMR binding sites is associated with *gyspy and gypsy*-like insulators, there is a considerable number of HMR-bound peaks that do not localize with Su(Hw), Mod(mdg4) and CP190 (Fig 4B). We noticed that many of these non *gypsy*-like sites are in close proximity to HP1a bound regions (Fig 1D). Indeed, when we sorted all HMR peaks according to the presence of HP1a in their proximity, we observed an almost complete lack of *gypsy* insulator binding proteins at these sites (Fig 5A). Consistent with the lack of Su(Hw) binding to this class of HMR peaks, a motif search revealed no enrichment of the Su(Hw) recognition site among those peaks but rather an enrichment for BEAF-32 binding sites (Fig 5B). To better understand a possible role of HMR at these sites, which we termed class 1 binding sites, we analyzed them with regards to their annotation. Interestingly, almost all HP1a-associated HMR binding sites (90%) are in close proximity to transcriptional start sites (TSS), whereas the other HMR binding sites show a somewhat broader distribution among various functional elements (Fig 5C). Strikingly, HMR binds very closely to the TSS at the boundary between HP1a containing domains and the gene body (S4A Fig). The genes in proximity of these HMR binding sites are classified as transcriptionally active suggesting that HMR might prevent the repressive influence of HP1a on neighbouring genes (Fig 5D). To investigate whether HMR loss has an impact on HP1a or H3K9me3 domains at these genomic regions, we performed HP1a ChIP and H3K9me3 ChIP upon HMR knockdown. However, we could not confirm extensive spreading of the heterochromatin marks HP1a or H3K9me3 after HMR loss (S4C and S4D Fig). Nevertheless, the genes associated with this class of HMR binding sites are transcriptionally down-regulated in *Hmr* mutant larvae and ovaries (Fig 5E, S4E Fig and [5]). This seems to be particularly important within regions that are rich in heterochromatin such as the 4th chromosome or the pericentromeric regions where we find the class of HP1a-associated binding sites highly enriched (Fig 5F). In summary, we can classify HMR's genomic binding sites into two groups: One being associated with *gypsy* insulators, and another one associated with active promoters in pericentromeric heterochromatin where HMR borders
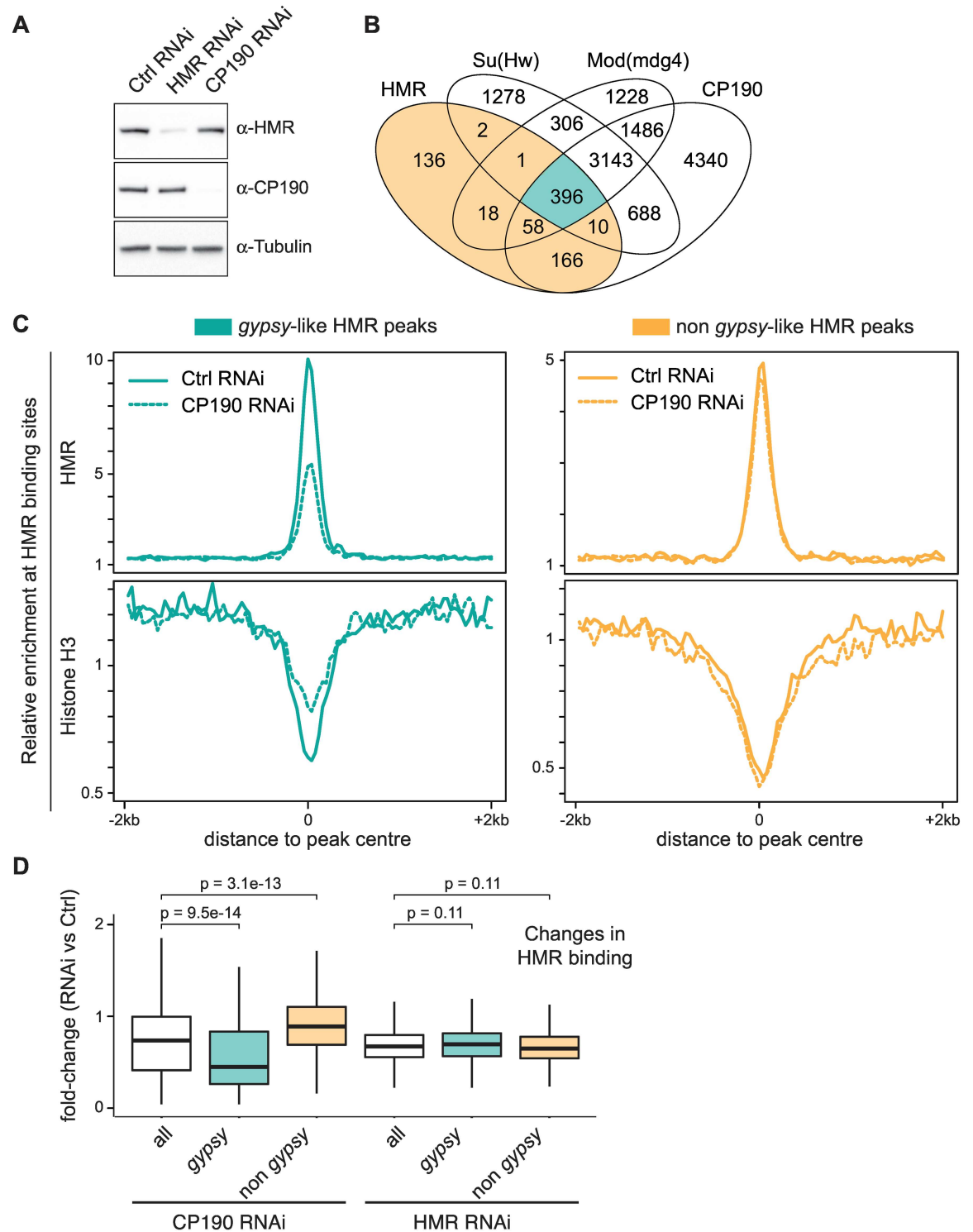
**Fig 4. HMR genomic localization to *gypsy*-like insulator sites is dependent on CP190. (A)** Western Blot of cell lysates after treatment with specific and control dsRNA shows an efficient knock-down of HMR and CP190. **(B)** Venn diagram of the overlap between HMR, CP190, Mod(mdg4) and Su(Hw) peaks [39] classifying HMR peaks as *gypsy*-like (highlighted in *green*) and non *gypsy*-like (highlighted in *orange*). **(C)** Composite analysis of HMR ChIP signals and Histone H3 ChIP signals at genomic HMR peak positions according to the groups defined in **(B)**. **(D)** Quantification of the fold-change of HMR ChIP enrichment upon CP190 RNAi and HMR RNAi. Box plots represent the fold-change of normalized HMR ChIP tag number aligned to 200 bp wide HMR peak regions. Peak regions with less than 50 aligned tags were excluded from the analysis. Significance of difference was estimated with p-values calculated with Wilcoxon rank sum test [72].
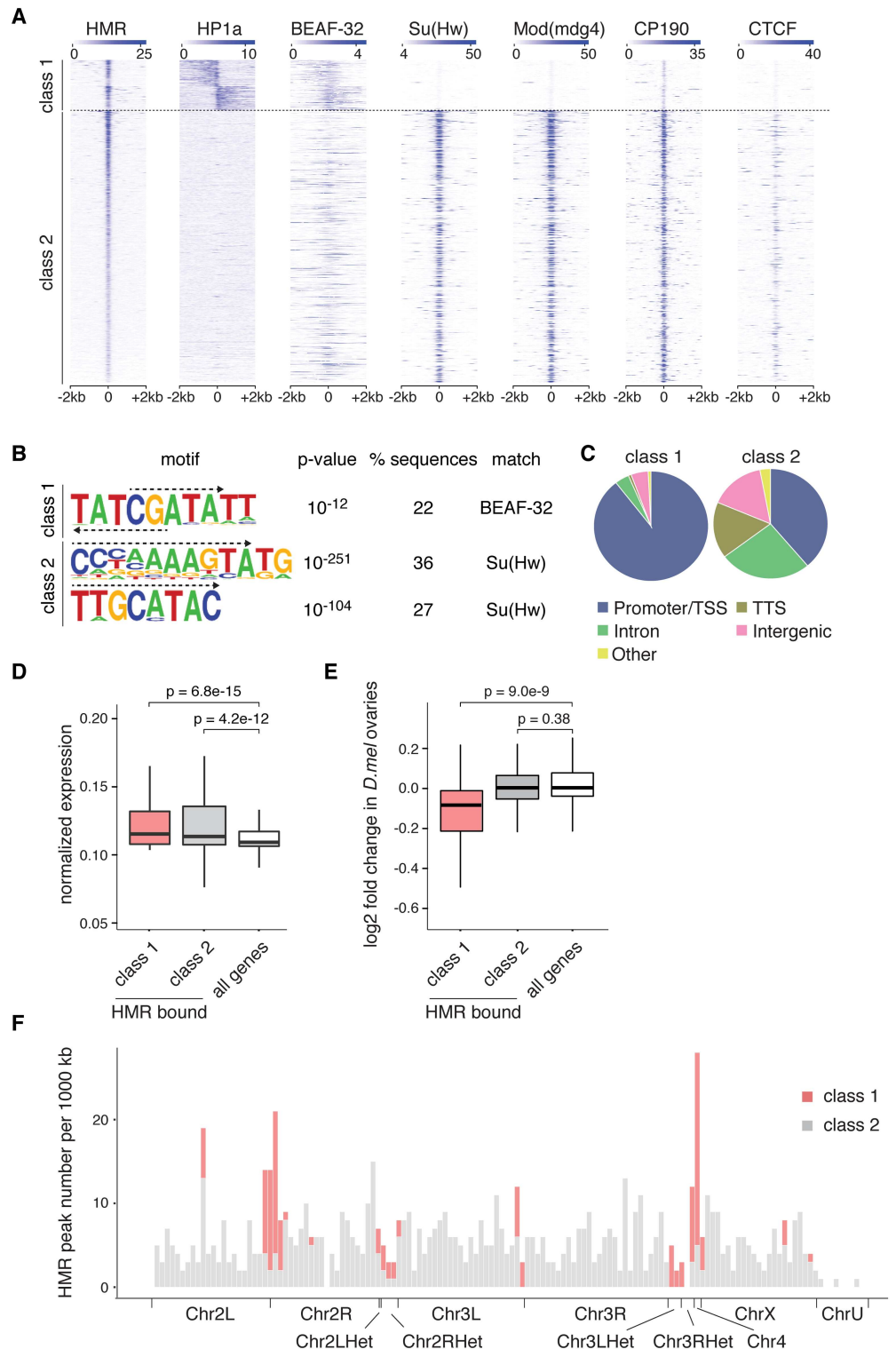
doi:10.1371/journal.pone.0171798.g004

**Fig 5. HMR borders HP1a together with BEAF-32 at the TSS of actively transcribed genes and enhances their transcription. (A)** Heatmaps of HMR, HP1a, BEAF-32 [40], Su(Hw), Mod(mdg4), CP190 and CTCF [39] ChIP signals. All signals are centered around the HMR binding sites, clustered according to adjacent HP1a signals and sorted by HMR intensity. **(B)** Sequence motifs identified within HMR peak regions from class 1 and class 2 based on HOMER motif analysis. The corresponding motif logo, p-value of enrichment, percentage of regions with this motif, and putative binding factors are indicated. Dashed arrows

mark the sequence that matches the published binding sites of Su(Hw) and BEAF-32 (see also S2A Fig). **(C)** Distribution of class 1 and class 2 HMR peaks among various genomic landmarks. **(D)** Box plot showing the normalized RNA expression of all genes and HMR-bound genes (promoter/TSS annotated) in class 1 and in class 2. S2 cells RNA expression levels were used according to [73]. Significance of difference was estimated with p-values calculated with Wilcoxon signed rank test [72]. **(E)** Box plot showing the log2 fold change of protein coding gene transcripts of all analyzed genes and HMR-bound genes (promoter/TSS annotated) in class 1 and in class 2 comparing *Hmr* mutant against wild type flies. The RNA-Seq data comes from experiments done in *D. melanogaster* ovaries [4]. Significance of difference was estimated with p-values calculated with Wilcoxon rank sum test [72]. For both box plots the box represents the interval that contains the central 50% of the data with the line indicating the median. The length of the whiskers is 1.5 times the interquartile distance (IQD). **(F)** Histogram showing HMR peak density across the annotated *D. melanogaster* genome. Class 1 HMR binding sites are enriched at region 31, centromere-proximal regions and the 4th chromosome.

doi:10.1371/journal.pone.0171798.g005

HP1a-containing chromatin regions together with BEAF-32 and potentially promotes gene transcription.

## Discussion

HMR localizes to centromeric and pericentromeric regions in *D. melanogaster* cell lines as well as in mitotically dividing embryonic cells where it has been suggested to act as a repressor of transposable elements [3–5]. Mutations in *Hmr* lead to overexpression of satellite DNA and transposable elements in ovaries and larvae [4]. Such a derepression is also observed in hybrid flies [53], where HMR and LHR levels are higher than the ones in pure species and result in a widespread distribution of the HMR/LHR complex [3]. To better understand the targeting principles that mediate HMR binding within the *D. melanogaster* genome, we wondered whether we could identify HMR binding sites by applying ChIP-Seq in the *D. melanogaster* S2 cell line. Combining this approach with RNAi mediated knockdown experiments we uncover a strong colocalization of HMR with *gypsy* insulator binding sites and demonstrate that HMR binding to these sites depends on the presence of the residing insulator protein complex. Notably, HMR associates only with a subset of all Su(Hw) binding sites, but almost all those sites can be classified as *gypsy*-like sites bound by CP190 and mod(mdg4) in addition to Su(Hw).

Besides dispersed binding of HMR at genomic *gypsy* insulator sites along the chromosome arms, we observe dense clusters of HMR binding sites around the centromere and on the 4th chromosome where it potentially serves to separate HP1a binding domains from highly active genes. This dense clustering of binding sites around the centromere correlates well with the strong colocalization of HMR signals with the centromeric H3 variant CID in immunolocalization experiments [3]. Due to its biochemical interaction and partial colocalization with the heterochromatin protein HP1a in *Drosophila* embryos, HMR has been suggested to be a *bona-fide* heterochromatin component [3,4,6,54]. However, in contrast to HP1a, we detect very distinct HMR binding sites within the genome. When we find HMR close to an HP1a binding domain, it rather borders it than covering the whole domain. The sharp HMR binding signals and the fact that almost all euchromatic HMR binding sites contain putative insulator elements, suggest a role of HMR in separating chromatin domains. A distinct boundary that separates constitutive heterochromatin from the core centromere has also been postulated by Olszak and colleagues who suggest that transition zones between heterochromatin and euchromatin are hotspots for sites of CID misincorporation [55]. Unfortunately, centromeres are notoriously difficult to study by next generation sequencing due to their highly repetitive nature [56,57]. In addition, the microscopic resolution is not sufficiently high to allow a distinction between a binding to the core centromere chromatin and the chromatin immediately adjacent to it. Therefore, we cannot rule out the possibility that HMR binds large domains at the central region of the *Drosophila* centromere. However, the fact that the purification of

chromatin containing the centromeric H3 variant CID did not identify HMR [58], suggests that it may very well also form a boundary between pericentromeric heterochromatin and the core centromere. To which extent and by which mechanism HMR fulfils a functional role at these genomic sites remains to be elucidated.

The genomic sites, where we find HMR bound next to an HP1a domain, are highly enriched for recognition sites of the insulator protein BEAF-32. Interestingly, a depletion of BEAF-32 in S2 cells results in an increased rate of mitotic defects [45], which is very reminiscent of the phenotype detected when HMR is depleted [3]. Similarly to flies carrying a mutation in the *Hmr* gene, flies in which BEAF-32 is only contributed maternally have defects in female fertility [59,60]. BEAF-32's role in maintaining associated promoter regions in an environment that facilitates high transcription levels [61] has been suggested to be functionally relevant for this phenotype [45]. Strikingly, we find most HMR/BEAF-32 binding sites located between HP1a containing heterochromatin and the transcription start site of a highly active gene. HP1a chromatin might fulfil a repressive function at these genomic regions and HMR might block this repressive impact on the neighbouring gene body. However, we do not see extensive spreading of HP1a or H3K9me3 upon HMR knockdown suggesting that the repressive effect is not directly mediated by HP1a binding or the HMR knock down not efficient enough. As there is evidence that HP1a can also promote gene transcription [62], HMR may also function as a co-activator for HP1a. Currently, we therefore consider HMR binding next to HP1a containing chromatin as a unifying feature of transcriptionally affected genes but can only speculate about potential mechanism by which HMR exerts its function.

Although HMR depletion has a substantial effect on the transcription of multiple transposons, we find HMR only enriched at the 5' insulator region of the *gypsy* or *gtwin* retrotransposons and to similar sites within the genome that are presumably derived from these elements. These sites are occupied by insulator proteins Su(Hw), CP190 and Mod(mdg4) and often display enhancer blocking activity in transgenic assays [43,63–65]. Artificial targeting of HMR to DNA placed between an enhancer and a promoter of a reporter gene can block the transcription activity [3], suggesting that HMR may indeed play a role in setting up endogenous boundary elements. Similar to what is known for Su(Hw), HMR binding to this class of binding sites is dependent on the presence of the structural protein CP190, which has a key function in the stabilization of insulator protein complexes [22]. However, as we do not observe a strong physical interaction between CP190 and HMR, the loss of HMR binding upon a reduction of CP190 levels may also be the result of increased nucleosome occupancy. Such increase in Histone H3 binding cannot be observed upon HMR removal suggesting that HMR acts downstream of CP190. Interestingly, CP190 loss impairs HMR binding to *gypsy*-like insulator sites but has weak effect on HMR binding to sites containing BEAF-32 recognition motifs. Notably, in contrast to BEAF-32, CP190 is not required for oogenesis [66], suggesting that the lack of HMR binding to the class 1 sites may be responsible for the female sterility phenotype observed in *Hmr* mutant flies.

How can we integrate our findings with the lethal phenotype of increased HMR/LHR levels in male hybrids? It is tempting to speculate that multiple additional binding sites that are observed in hybrids and on polytene chromosomes of fly strains over-expressing HMR [3] constitute boundary regions. An increased binding to such boundaries, which have been shown to cluster and form aggregates *in vivo* [48,67,68], may trigger a massive change in nuclear architecture. In turn, this could indirectly activate multiple transposable elements similar to what is observed when centromere clustering is disturbed [69]. Such a disturbed nuclear architecture may then trigger the activation of a cell cycle checkpoint which has been previously suggested to be a major cause of hybrid lethality [70,71].

Altogether, our data provide a novel link between HMR and *cis*-regulatory elements bound by insulator proteins. We speculate that divergent evolution of such genomic elements and their corresponding binding factors in sibling species is triggering hybrid incompatibilities.

## Supporting information

**S1 Fig. Control experiments of HMR ChIP-Seq studies. (A)** Venn diagram showing the lack of overlap between HMR peaks (peaks identified in at least two out of three independent biological replicates, highlighted in *grey*) and control IgG (anti-FLAG) ChIP peaks (pool of peaks from two independent biological replicates). **(B)** Changes in HMR ChIP enrichment upon HMR RNAi versus a control RNAi (GST) in two biological replicates. Each data point represents a mapped HMR peak. The scatter plot on the left displays fold changes of normalized HMR ChIP tag number mapped to a 200 bp HMR peak region in two biological replicates. Peak regions with less than 50 aligned tags were excluded from the analysis. The histogram on the right shows the frequency of peaks displaying a reduction of HMR binding upon knockdown. Shown are average values of replicate 1 and replicate 2. **(C)** ChIP-qPCR showing specific HMR enrichment at HMR binding sites. HMR ChIP is enriched for HMR binding sites in both wild type and HMR-Flag$_2$ expressing cells. FLAG ChIP is enriched for HMR binding sites only in HMR-Flag$_2$ expressing cells but not in wild type cells lacking the Flag$_2$ epitope. Data are represented as mean ± SD of three technical replicates. **(D)** Genome browser view of HMR ChIP, HP1a ChIP and control IgG ChIP signal at a large centromere-proximal region at the right arm of the 2nd chromosome.
(TIF)

**S2 Fig. Overlap of HMR binding sites with known insulator regions and repetitive DNA.** **(A)** Sequence motifs identified within Su(Hw) [39] and BEAF-32 [40] peak regions. The corresponding motif logo, p-value of enrichment and percentage of regions with this motif are indicated. Dashed arrows mark the sequence that matches the published binding sites of Su(Hw) and BEAF-32. (B) Genome browser view of ChIP signals showing combinatoric binding pattern for HMR and the insulator proteins CP190, Mod(mdg4), Su(Hw), CTCF [39] and BEAF-32 [40]. (C) Su(Hw) and HP1a ChIP tag enrichment at repetitive DNA elements. Each point in the scatter plot represents the enrichment (log2 fold) over input and the RPKM of an individual repeat from Repbase. Repeats with less than 2-fold enrichment are not displayed. (D) ChIP tag density of HMR and the *gypsy*-insulator proteins CP190, Mod(mdg4), Su(Hw) [39] across the *gypsy-twin* repeat.
(TIF)

**S3 Fig. Selective effect of CP190 RNAi on HMR binding to *gypsy*-like elements.** Composite analysis of HMR ChIP signal and Histone H3 ChIP signal at genomic HMR peak positions according to the groups defined in Fig 4B. The ChIP signals were obtained upon control RNAi and HMR RNAi. The HMR ChIP signals are similarly affected in both groups, whereas Histone H3 ChIP signals are retained.
(TIF)

**S4 Fig. Additional information for HP1a-associated HMR binding sites. (A)** Composite analysis of HMR, HP1a and BEAF-32 ChIP signals at class 1 genomic sites relative to the transcriptional start site (TSS) and the gene body. Shown are normalised and scaled read density plots **(B)** Peak overlap of HMR with peaks of the insulator proteins CP190, Mod(mdg4), Su (Hw), CTCF [39] and BEAF-32 [40] for class 1 and for class 2 HMR binding sites. **(C)** Composite analysis of HP1a and H3K9me3 ChIP signals at class 1 HMR binding sites after HMR knockdown. Class 1 is defined in Fig 5A but oriented according to HP1a ChIP signal. **(D)**

Western Blot analysis on cell lysates to assay protein levels after HMR knockdown. Tubulin protein detection served as control. **(E)** Same as described in Fig 5E, but the RNA-Seq data comes from experiments done in *D. melanogaster* male larvae [4].
(TIF)

**S1 Table. List of primers used for CRISPR/*cas9* genome editing, RNAi experiments and ChIP Real-Time PCR.** List of primers used in this study. Primers used in ChIP Real-Time PCR were designed with help of Primer3.
(DOCX)

**S2 Table. ChIP-Seq sample overview and number of uniquely aligned sequence reads.** ChIP-Seq sample overview and number of uniquely aligned sequence reads. The percentage of uniquely mapped reads in ChIP-Seq experiments can largely vary and depends on the nature of the ChIPed protein. Proteins that bind repetitive regions (such as HMR or HP1a) give substantially lower percentages of uniquely mapped reads.
(DOCX)

**S3 Table. HMR peaks used for downstream analysis.** HMR peak list derived from HOMER peak calling on three biological replicates (Fig 1A). First three columns provide information on the peak position within the genome (chromosome, peak start and end using dm3), followed by peak annotation obtained from ChIPseeks implementation of HOMER (Fig 5C) and classification according to adjacent HP1a signals (Fig 5A).
(XLSX)

**S1 Methods. Supporting information on methods.** *Hmr* gene editing using CRISPR/*cas9*, extended ChIP Real-Time PCR methods, extended ChIP-seq data analysis methods.
(DOCX)

## Acknowledgments

## Author Contributions

**Conceptualization:** AI AWT TAG.

**Data curation:** TAG BS.

**Formal analysis:** BS PS TAG.

**Funding acquisition:** AI.

**Investigation:** TAG.

**Methodology:** AI TAG.

**Project administration:** AI TAG.

**Resources:** AL TAG.

**Software:** BS.

**Supervision:** AI TAG AWT.

**Validation:** TAG.

**Visualization:** AI TAG.

**Writing – original draft:** AI TAG.

**Writing – review & editing:** AI TAG AL.

## References

1. Maheshwari S, Barbash DA. The Genetics of Hybrid Incompatibilities. Annu Rev Genet. 2011; 45: 331–355. doi: 10.1146/annurev-genet-110410-132514 PMID: 21910629

2. Barbash DA, Awadalla P, Tarone AM. Functional divergence caused by ancient positive selection of a Drosophila hybrid incompatibility locus. PLoS Biol. 2004; 2: e142. doi: 10.1371/journal.pbio.0020142 PMID: 15208709

3. Thomae AW, Schade GOM, Padeken J, Borath M, Vetter I, Kremmer E, et al. A Pair of Centromeric Proteins Mediates Reproductive Isolation in Drosophila Species. Dev Cell. 2013; 27: 412–424. doi: 10.1016/j.devcel.2013.10.001 PMID: 24239514

4. Satyaki PRV, Cuykendall TN, Wei KH-C, Brideau NJ, Kwak H, Aruna S, et al. The hmr and lhr hybrid incompatibility genes suppress a broad range of heterochromatic repeats. Malik HS, editor. PLoS Genet. Public Library of Science; 2014; 10: e1004240.

5. Wei KHC, Clark AG, Barbash DA. Limited Gene Misregulation Is Exacerbated by Allele-Specific Upregulation in Lethal Hybrids between Drosophila melanogaster and Drosophila simulans. Mol Biol Evol. 2014; 31: 1767–1778. doi: 10.1093/molbev/msu127 PMID: 24723419

6. Maheshwari S, Barbash DA. Cis-by-Trans Regulatory Divergence Causes the Asymmetric Lethal Effects of an Ancestral Hybrid Incompatibility Gene. Begun DJ, editor. PLoS Genet. 2012; 8: e1002597. doi: 10.1371/journal.pgen.1002597 PMID: 22457639

7. Straub T, Grimaud C, Gilfillan GD, Mitterweger A, Becker PB. The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS Genet. 2008; 4: e1000302. doi: 10.1371/journal.pgen.1000302 PMID: 19079572

8. Gorchakov AA, Alekseyenko AA, Kharchenko P, Park PJ, Kuroda MI. Long-range spreading of dosage compensation in Drosophila captures transcribed autosomal genes inserted on X. Genes Dev. 2009; 23: 2266–2271. doi: 10.1101/gad.1840409 PMID: 19797766

9. 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. Elsevier; 2011; 144: 214–226.

10. Tartof KD, Hobbs C, Jones M. A structural basis for variegating position effects. Cell. 1984; 37: 869–878. PMID: 6086148

11. Talbert PB, Henikoff S. Spreading of silent chromatin: inaction at a distance. Nat Rev Genet. 2006; 7: 793–803. doi: 10.1038/nrg1920 PMID: 16983375

12. Rodriguez MA, Vermaak D, Bayes JJ, Malik HS. Species-specific positive selection of the male-specific lethal complex that participates in dosage compensation in Drosophila. Proc Natl Acad Sci USA. 2007; 104: 15412–15417. doi: 10.1073/pnas.0707445104 PMID: 17878295

13. Ross BD, Rosin L, Thomae AW, Hiatt MA, Vermaak D, la Cruz de AFA, et al. Stepwise Evolution of Essential Centromere Function in a Drosophila Neogene. Science. 2013; 340: 1211–1214. doi: 10.1126/science.1234393 PMID: 23744945

14. Bayes JJ, Malik HS. Altered Heterochromatin Binding by a Hybrid Sterility Protein in Drosophila Sibling Species. Science. 2009; 326: 1538–1541. doi: 10.1126/science.1181756 PMID: 19933102

15. Sawamura K. Chromatin Evolution and Molecular Drive in Speciation. International Journal of Evolutionary Biology. 2012; 2012: 1–9.

16. Lerat E, Burlet N, Biémont C, Vieira C. Comparative analysis of transposable elements in the melanogaster subgroup sequenced genomes. Gene. Elsevier B.V; 2011; 473: 100–109.

17. Vieira C, Biémont C. Transposable element dynamics in two sibling species: Drosophila melanogaster and Drosophila simulans. Genetica. 2004; 120: 115–123. PMID: 15088652

18. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, et al. Evolution of genes and genomes on the Drosophila phylogeny. Nature. 2007; 450: 203–218. doi: 10.1038/nature06341 PMID: 17994087

19. Ni X, Zhang YE, Nègre N, Chen S, Long M, White KP. Adaptive evolution and the birth of CTCF binding sites in the Drosophila genome. PLoS Biol. 2012; 10: e1001420. doi: 10.1371/journal.pbio.1001420 PMID: 23139640

20. Indjeian VB, Kingman GA, Jones FC, Guenther CA, Grimwood J, Schmutz J, et al. Evolving New Skeletal Traits by cis-Regulatory Changes in Bone Morphogenetic Proteins. Cell. 2016; 164: 45–56. doi: 10.1016/j.cell.2015.12.007 PMID: 26774823

21. Dai H, Chen Y, Chen S, Mao Q, Kennedy D, Landback P, et al. The evolution of courtship behaviors through the origination of a new gene in Drosophila. National Acad Sciences; 2008; 105: 7478–7483.

22. Bushey AM, Dorman ER, Corces VG. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. Mol Cell. 2008; 32: 1–9. doi: 10.1016/j.molcel.2008.08.017 PMID: 18851828

23. Yang J, Ramos E, Corces VG. The BEAF-32 insulator coordinates genome organization and function during the evolution of Drosophila species. Genome Res. 2012; 22: 2199–2207. doi: 10.1101/gr.142125.112 PMID: 22895281

24. Heger P, George R, Wiehe T. Successive gain of insulator proteins in arthropod evolution. Evolution. 2013; 67: 2945–2956. doi: 10.1111/evo.12155 PMID: 24094345

25. 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. Elsevier Inc; 2012; 148: 458–472.

26. Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, et al. Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. Nature. 2007; 450: 219–232. doi: 10.1038/nature06340 PMID: 17994088

27. Bosco G, Campbell P, Leiva-Neto JT, Markow TA. Analysis of Drosophila species genome size and satellite DNA content reveals significant differences among strains as well as between species. Genetics. 2007; 177: 1277–1290. doi: 10.1534/genetics.107.075069 PMID: 18039867

28. Labrador M, Corces VG. Setting the boundaries of chromatin domains and nuclear organization. Cell. 2002; 111: 151–154. PMID: 12408858

29. James TC, Elgin SC. Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. Mol Cell Biol. 1986; 6: 3862–3872. PMID: 3099166

30. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10: R25. doi: 10.1186/gb-2009-10-3-r25 PMID: 19261174

31. Andrews S. FATSQC, a quality control tool for high throughput sequence data. In: http://www.bioinformatics.babraham.ac.ukprojectsfastqc. 2010.

32. Hannon GJ. FASTX Toolkits FASTA/Q short reads preprocessing kit [Internet].

33. Nicol JW, Helt GA, Blanchard SG, Raja A, Loraine AE. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics. 2009; 25: 2730–2731. doi: 10.1093/bioinformatics/btp472 PMID: 19654113

34. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinformatics. 2013; 14: 178–192. doi: 10.1093/bib/bbs017 PMID: 22517427

35. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010; 38: 576–589. doi: 10.1016/j.molcel.2010.05.004 PMID: 20513432

36. Chen T-W, Li H-P, Lee C-C, Gan R-C, Huang P-J, Wu TH, et al. ChIPseek, a web-based analysis tool for ChIP data. BMC Genomics. BioMed Central; 2014; 15: 539.

37. Zhu LJ, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics. 2010; 11: 237. doi: 10.1186/1471-2105-11-237 PMID: 20459804

38. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mobile DNA. 2015; 6: 11. doi: 10.1186/s13100-015-0041-9 PMID: 26045719

39. Ong C-T, Van Bortle K, Ramos E, Corces VG. Poly(ADP-ribosyl)ation Regulates Insulator Function and IntrachromosomalInteractions in Drosophila. Cell. Elsevier Inc; 2013; 155: 148–159.

40. Riddle NC, Minoda A, Kharchenko PV, Alekseyenko AA, Schwartz YB, Tolstorukov MY, et al. Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. Genome Res. 2011; 21: 147–163. doi: 10.1101/gr.110098.110 PMID: 21177972

41. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013; 31: 46–53. doi: 10.1038/nbt.2450 PMID: 23222703

42. Frasch M, Glover DM, Saumweber H. Nuclear antigens follow different pathways into daughter nuclei during mitosis in early Drosophila embryos. J Cell Sci. 1986; 82: 155–172. PMID: 3098744

43. Schwartz YB, Linder-Basso D, Kharchenko PV, Tolstorukov MY, Kim M, Li H-B, et al. Nature and function of insulator protein binding sites in the Drosophila genome. Genome Res. Cold Spring Harbor Lab; 2012; 22: 2188–2198.

44. Spana C, Harrison DA, Corces VG. The Drosophila melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. Genes Dev. 1988; 2: 1414–1423. PMID: 2850261

45. 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. Misteli T, editor. PLoS Biol. 2008; 6: e327–15.

46. Ghosh D, Gerasimova TI, Corces VG. Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. EMBO J. 2001; 20: 2518–2527. doi: 10.1093/emboj/20.10.2518 PMID: 11350941

47. Gause M, Morcillo P, Dorsett D. Insulation of enhancer-promoter communication by a gypsy transposon insert in the Drosophila cut gene: cooperation between suppressor of hairy-wing and modifier of mdg4 proteins. Mol Cell Biol. 2001; 21: 4807–4817. doi: 10.1128/MCB.21.14.4807-4817.2001 PMID: 11416154

48. Pai C-Y, Lei EP, Ghosh D, Corces VG. The centrosomal protein CP190 is a component of the gypsy chromatin insulator. Mol Cell. Elsevier; 2004; 16: 737–748.

49. Abad JP, Carmena M, Baars S, Saunders RD, Glover DM, Ludeña P, et al. Dodeca satellite: a conserved G+C-rich satellite from the centromeric heterochromatin of Drosophila melanogaster. Proc Natl Acad Sci USA. 1992; 89: 4663–4667. PMID: 1584802

50. Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, et al. The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. Genome Biol. 2002; 3: RESEARCH0084. doi: 10.1186/gb-2002-3-12-research0084 PMID: 12537573

51. Bartkuhn M, Straub T, Herold M, Herrmann M, Rathke C, Saumweber H, et al. Active promoters and insulators are marked by the centrosomal protein 190. EMBO J. EMBO Press; 2009; 28: 877–888.

52. Bohla D, Herold M, Panzer I, Buxa MK, Ali T, Demmers J, et al. A functional insulator screen identifies NURF and dREAM components to be required for enhancer-blocking. Dean A, editor. PLoS ONE. 2014; 9: e107765. doi: 10.1371/journal.pone.0107765 PMID: 25247414

53. Kelleher ES, Edelman NB, Barbash DA. Drosophila Interspecific Hybrids Phenocopy piRNA-Pathway Mutants. Noor MAF, editor. PLoS Biol. 2012; 10: e1001428. doi: 10.1371/journal.pbio.1001428 PMID: 23189033

54. Brideau NJ, Barbash DA. Functional conservation of the Drosophila hybrid incompatibility gene Lhr. BMC Evol Biol. BioMed Central Ltd; 2011; 11: 57.

55. Olszak AM, van Essen D, Pereira AJ, Diehl S, Manke T, Maiato H, et al. Heterochromatin boundaries are hotspots for de novo kinetochore formation. Nat Cell Biol. 2011; 13: 799–808. doi: 10.1038/ncb2272 PMID: 21685892

56. Sun X, Wahlstrom J, Karpen G. Molecular structure of a functional Drosophila centromere. Cell. 1997; 91: 1007–1019. PMID: 9428523

57. Sun X, Le HD, Wahlstrom JM, Karpen GH. Sequence Analysis of a Functional Drosophila Centromere. Genome Res. 2003; 13: 182–194. doi: 10.1101/gr.681703 PMID: 12566396

58. Barth TK, Schade GOM, Schmidt A, Vetter I, Wirth M, Heun P, et al. Identification of novel Drosophila centromere-associated proteins. Proteomics. 2014; 14: 2167–2178. doi: 10.1002/pmic.201400052 PMID: 24841622

59. Aruna S, Flores HA, Barbash DA. Reduced fertility of Drosophila melanogaster hybrid male rescue (Hmr) mutant females is partially complemented by Hmr orthologs from sibling species. Genetics. 2009; 181: 1437–1450. doi: 10.1534/genetics.108.100057 PMID: 19153254

60. Roy S, Gilbert MK, Hart CM. Characterization of BEAF mutations isolated by homologous recombination in Drosophila. Genetics. Genetics; 2007; 176: 801–813. doi: 10.1534/genetics.106.068056 PMID: 17435231

61. Jiang N, Emberly E, Cuvier O, Hart CM. Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in Drosophila melanogaster links BEAF to transcription. Mol Cell Biol. American Society for Microbiology; 2009; 29: 3556–3568.

62. Yasuhara JC, Wakimoto BT. Oxymoron no more: the expanding world of heterochromatic genes. Trends Genet. Elsevier; 2006; 22: 330–338.

63. Parnell TJ, Viering MM, Skjesol A, Helou C, Kuhn EJ, Geyer PK. An endogenous suppressor of hairy-wing insulator separates regulatory domains in Drosophila. Proc Natl Acad Sci USA. 2003; 100: 13436–13441. doi: 10.1073/pnas.2333111100 PMID: 14597701

64. Golovnin A, Biryukova I, Birukova I, Romanova O, Silicheva M, Parshikov A, et al. An endogenous Su (Hw) insulator separates the yellow gene from the Achaete-scute gene complex in Drosophila. Development. 2003; 130: 3249–3258. PMID: 12783795

65. Nègre N, Brown CD, Ma L, Bristow CA, Miller SW, Wagner U, et al. A cis-regulatory map of the Drosophila genome. Nature. 2011; 471: 527–531. doi: 10.1038/nature09990 PMID: 21430782

66. Baxley RM, Soshnev AA, Koryakov DE, Zhimulev IF, Geyer PK. The role of the Suppressor of Hairy-wing insulator protein in Drosophila oogenesis. Dev Biol. 2011; 356: 398–410. doi: 10.1016/j.ydbio.2011.05.666 PMID: 21651900

67. Gerasimova TI, Byrd K, Corces VG. A chromatin insulator determines the nuclear localization of DNA. Mol Cell. 2000; 6: 1025–1035. PMID: 11106742

68. Gerasimova TI, Corces VG. Polycomb and trithorax group proteins mediate the function of a chromatin insulator. Cell. 1998; 92: 511–521. PMID: 9491892

69. Padeken J, Mendiburo MJ, Chlamydas S, Schwarz H-J, Kremmer E, Heun P. The Nucleoplasmin Homolog NLP Mediates Centromere Clustering and Anchoring to the Nucleolus. Mol Cell. Elsevier Inc; 2013;: 1–14.

70. Bolkan BJ, Booker R, Goldberg ML, Barbash DA. Developmental and Cell Cycle Progression Defects in Drosophila Hybrid Males. Genetics. 2007; 177: 2233–2241. doi: 10.1534/genetics.107.079939 PMID: 17947412

71. Phadnis N, Baker EP, Cooper JC, Frizzell KA, Hsieh E, la Cruz de AFA, et al. An essential cell cycle regulation gene causes hybrid inviability in Drosophila. Science. 2015; 350: 1552–1555. doi: 10.1126/science.aac7504 PMID: 26680200

72. Wilcoxon F. Individual comparisons by ranking methods. Biometrics bulletin. 1945; 1s: 80–83.

73. Rus F, Flatt T, Tong M, Aggarwal K, Okuda K, Kleino A, et al. Ecdysone triggered PGRP-LC expression controls Drosophila innate immunity. EMBO J. 2013; 32: 1626–1638. doi: 10.1038/emboj.2013.100 PMID: 23652443