Context Differences Reveal Insulator and Activator Functions of a Su(Hw) Binding Region

Alexey A. Soshnev1,9, Xingguo Li2,3, Misty D. Wehling1, Pamela K. Geyer1,2*  
1 Program in Molecular and Cellular Biology, University of Iowa, Iowa City, Iowa, United States of America, 2 Department of Biochemistry, University of Iowa, Iowa City, Iowa, United States of America

Abstract

Insulators are DNA elements that divide chromosomes into independent transcriptional domains. The Drosophila genome contains hundreds of binding sites for the Suppressor of Hairy-wing (Su(Hw)) insulator protein, corresponding to locations of the retroviral gypsy and non-gypsy binding regions (BRs). The first non-gypsy BR identified, 1A-2, resides in cytological region 1A. Using a quantitative transgene system, we show that 1A-2 is a composite insulator containing enhancer blocking and facilitator elements. We discovered that 1A-2 separates the yellow (y) gene from a previously unannotated, non-coding RNA gene, named yar for y-achaete (ac) intergenic RNA. The role of 1A-2 was elucidated using homologous recombination to excise these sequences from the natural location, representing the first deletion of any Su(Hw) BR in the genome. Loss of 1A-2 reduced yar RNA accumulation, without affecting mRNA levels from the neighboring y and ac genes. These data indicate that within the 1A region, 1A-2 acts an activator of yar transcription. Taken together, these studies reveal that the properties of 1A-2 are context-dependent, as this element has both insulator and enhancer activities. These findings imply that the function of non-gypsy Su(Hw) BRs depends on the genomic environment, predicting that Su(Hw) BRs represent a diverse collection of genomic regulatory elements.

Introduction

In eukaryotic genomes, neighboring genes often display distinct spatial and temporal patterns of transcription, even though intergenic distances are within the range of enhancer and silencer action. These observations suggest that constraints exist that limit promiscuous interactions between long distance regulatory elements and non-target promoters. Chromatin insulators represent one class of genomic elements that restrict enhancer and silencer action [1–5].

Insulators have been identified based on two functional properties. First, insulators prevent enhancer and silencer modulation of a promoter in a position-dependent manner, such that an enhancer or silencer is blocked only when the insulator is located between these elements and a promoter. Second, insulators protect gene expression from positive and negative chromosomal position effects associated with ectopic placement of genes within genomes, an activity referred to as barrier function. Sequences with one or both of these properties have been identified in most eukaryotic genomes and have been implicated in the regulation of diverse cellular processes, ranging from centromere function in yeast to imprinting in mammals [6,7]. These observations imply that insulators are fundamental components of eukaryotic genomes.

One of the best-characterized insulators resides in the 5′ untranslated region of the Drosophila gypsy retrovirus. This versatile gypsy insulator blocks over twenty enhancers active in different tissues and developmental stages, prevents repressive effects caused by Polycomb group complexes and heterochromatin and protects an origin of DNA replication from chromosomal position effects [2,5]. The gypsy insulator consists of a cluster of twelve repeats that are bound by the zinc finger Suppressor of Hairy-wing (Su(Hw)) protein [8]. At least three additional proteins are associated with the gypsy insulator, including Modifier of (mdg4) 67.2 (Mod67.2), Centrosomal Protein of 190 kD (CP190) and Enhancer of y2 [E(y)2]. In general, Mod67.2 and CP190 are required for enhancer and silencer blocking by the gypsy insulator, while E(y)2 has been shown to be required only for barrier function [9–13].

The Su(Hw) protein associates with hundreds of non-gypsy regions in the Drosophila genome that have a largely unknown function. The extensive co-localization of the four gypsy insulator proteins at non-gypsy regions has led to the proposal that these represent chromatin insulators. Yet, non-gypsy Su(Hw) binding regions are different in sequence and organization from the gypsy insulator, with the majority of BRs containing single Su(Hw) binding sites (BSs) [14–18]. This observation is striking, as at least four tightly spaced Su(Hw) sites from the gypsy insulator were required for robust enhancer blocking [19–21]. Direct tests of the non-gypsy BRs in transgene assays show that most, but not all, interfere with enhancer-activated transcription [15–18]. These findings imply that non-gypsy regions contain additional elements that assist the insulator function of Su(Hw).
The first non-gypsy Su(Hw) BR identified, named 1A-2, is a cluster of two Su(Hw) BSs located in cytological region 1A (Figure 1). Here we investigated the properties of 1A-2, using two strategies. First, we employed a quantitative transgene system to define the 1A-2 sequences required for enhancer blocking. Second, we performed homologous recombination to establish lines carrying a deletion of 1A-2 at the natural genomic location, representing the first deletion of non-gypsy Su(Hw) BR in the Drosophila genome. Effects of the loss of these sequences on gene expression in the 1A region were determined, leading to the discovery that 1A-2 contributes to transcriptional activation of a novel, non-coding RNA gene. Taken together, our studies demonstrate that 1A-2 has both activator and insulators properties, depending on the context tested. These findings imply that properties of non-gypsy Su(Hw) BRs are influenced by the genomic environment, predicting that Su(Hw) BRs represent a diverse collection of elements with distinct regulatory functions.

Results

1A-2 Is a Composite Insulator

The Su(Hw) BR 1A-2 is a 520 bp element that contains two Su(Hw) BSs [18] and a CP190 BS [9]. Previous studies using qualitative transgene assays demonstrated that 1A-2 blocked enhancer-activated transcription in a position-dependent manner, a key feature of insulator activity [17,18]. We employed the quantitative Fat Body Enhancer (FBE1)-yolk protein (yp2)-LacZ transgene to define DNA sequences required for 1A-2 enhancer blocking (Figure 1), a system previously used to characterize properties of the gypsy insulator [20,22]. A reporter transgene was constructed wherein full length 1A-2(520) was inserted between FBE1 and the yp2 promoter. Multiple P[F-IA-2(520)-yp2] transgenic lines with single insertions were established. Quantitative β-galactosidase activity assays were completed to define the level of yp2 promoter activity. Protein extracts were isolated from adult females and organization from the gypsy insulator. To gain insights into the role of Su(Hw) in genome organization, we defined the properties of the first non-gypsy Su(Hw) binding region identified, 1A-2. Our studies reveal differences in 1A-2 activity, depending on the context tested. We show that 1A-2 is an insulator in enhancer blocking studies but functions as a transcriptional activator within the natural genomic location. Our findings are reminiscent of properties of binding regions that associate with the vertebrate CTCF protein, which have defined insulator, activator, and repressor functions. Finally, our studies indicate that a noncoding RNA gene may contribute to independent transcriptional regulation in the genome.

The minimal sequences required for 1A-2 insulator function were determined by generation of transgenic lines carrying transposons with insertion of subregions of 1A-2 between FBE1 and yp2-Lac Z ([Figure 2]). P[F-IA-2(153)-yp2] females showed a strong enhancer block (aau 0.62). As this subregion lacks the CP190 BS [9], these findings indicate that direct CP190 binding is not required for insulator function. 1A-2(157) was further divided into two parts, one containing the two Su(Hw) BSs, 1A-2(79), and one containing the remaining sequences, 1A-2(78). Transgenic P[F-IA-2(79)-yp2] females showed a two-fold weaker enhancer block than 1A-2(157) (aau 1.29, P = 0.02), whereas P[F-IA-2(78)-yp2] females showed high yp2 activity levels, close to those obtained for the control P[F-yp2] females (aau 5.9 versus 5.97). These data suggest that 1A-2(78) contributes to the blocking effectiveness of the 1A-2 Su(Hw) BSs, but cannot itself block enhancer-promoter interactions.

We considered two possibilities to account for the contributions made by 1A-2(78). First, these sequences might contain a binding site(s) for a second insulator protein that cooperates with the Su(Hw) BSs for insulator function. Second, 1A-2(78) might improve the activity of the Su(Hw) BSs, perhaps by increasing in vivo association. We reasoned that if 1A-2(78) contained a binding site for a novel insulator protein, then insulator effects might require a reiterated element, as observed previously when individual binding sites for other insulator proteins were tested [23,24]. To this end, we generated P[F-IA-2 (78×4)-yp2] that carried four copies of 1A-2(78) inserted between FBE1 and the yp2 promoter. Surprisingly, these transgenic females had higher yp2 activity than the control P[F-yp2] females (aau 18.78 versus 5.97 aau, P = 6.3×10^{-9}). Transgenic P[F-IA-2(78×4)-yp2] males showed no yp2 activity (data not shown). Based on the retained transcriptional specificity of the P[F-IA-2 (78×4)-yp2] transgene, we conclude that 1A-2(78) is not a general transcriptional enhancer but improves the activity of FBE1. These data imply that 1A-2(78) may possess a general activity that facilitates factor association. To test this postulate, we determined whether 1A-2(78) restored enhancer blocking to a synthetic Su(Hw) BR containing three reiterated gypsy BSs (3R:3) that was previously shown to be inactive in this transgene system [20]. Supporting a facilitator function of 1A-2(78) we found that transgenic P[F- 3R:3-IA-2(78×4)-yp2] females had low yp2 activity (aau 0.22). These studies show that in the presence of 1A-2(78), 3R:3 became a strong insulator. As previous findings suggest that the effectiveness of enhancer blocking by the Su(Hw) protein is limited by the in vivo accessibility of Su(Hw), we conclude 1A-2(78) is a facilitator that may improve transcription factor binding to chromosomes.

The y-ac Intergenic Region Contains a Novel, Non-Coding RNA Gene

As a first step in defining the role of 1A-2 within the y-ac region, we evaluated whether the existing annotation reflected the transcriptional potential of this region. These analyses were motivated by the recent studies showing widespread transcription in intergenic regions of the Drosophila genome [25]. A search of the NCBI databases uncovered a small, novel, processed EST of ~400 nt that was transcribed from the y-ac intergenic sequences. Sequences corresponding to this EST are located ~1.4 kb downstream of the y termination signal and transcribed in the same direction as the y and ac genes. Northern analyses of embryonic polyA+ RNA using a radiolabeled probe representing the intergenic EST identified a family of RNAs, with the most
abundant species sized at ~1.6 kb (Figure 3). Accumulation of these RNAs began ~7 hours after the start of embryogenesis, in agreement with the expression profile obtained using tiling array studies of embryonic RNAs [25]. These data suggest that the y-ac intergenic region contains a previously uncharacterized gene, which we call yar, for y-ac intergenic RNA. Activation of genes in the 1A locus is temporally in an order following chromosomal position, such that ac, then yar and then y is transcribed.

The structure of the yar RNAs was defined using rapid amplification of cDNA ends (RACE, Figure 4). Sequence analysis of the 5' RACE products revealed three discrete transcription start sites within an ~200 bp region, with the most distal RNA starting ~1.2 kb downstream of the y gene. Each putative start site showed weak homology to Drosophila transcriptional control elements [26], with two having a partial match to the TATA consensus sequence located 17 to 35 bp upstream of the start site. Sequence
analysis of the 3′ RACE products identified multiple splice variants, each ending in a common exon that contained an unconventional polyadenylation signal sequence AAATACA, previously estimated to be present in 3–3% of Drosophila genes [27], that was located 12 bp upstream of the string of As in the RACE products. Predicted translation of the yar RNAs indicated that no transcript would encode a protein of more than 75 amino acids, implying that yar is a non-coding RNA gene.

Loss of 1A-2 Does Not Alter Adult Phenotypes Generated by y and ac Expression

Ends out gene targeting was used to delete 1A-2 from the y-ac region (Figures 4, 5). Gene targeting is a two step processes that requires establishment of transgenic flies that carry a transposon with the replacement gene, followed by the introduction of endonucleases to stimulate homologous recombination between the replacement gene and its endogenous homologue. To delete 1A-2, we constructed P[y^{A1A2} target]. This transposon carried a modified y gene, wherein 1A-2 was replaced by the hypomorphic u^{ab} gene that was flanked by loxP sites (Figure 5). Transgenic lines were established in a y^{1} u^{116} background, where the endogenous y gene carried a mutation of the translation start codon, and the endogenous u gene carried a deletion of the promoter. P[y^{A1A2} target] flies had orange eyes and dark pigmentation of all cuticle structures except the wing, as the y gene lacked the wing enhancer. To stimulate recombination, transgenic y^{1} u^{116}, P[y^{A1A2} target] males were crossed to females carrying the heat shock (hs)-FLP recombinase and the hs-I-SceI endonuclease transgenes and progeny of this cross were heat shocked to produce the endonucleases. Over 100 resulting females were crossed to y^{1} u^{116} males and homologous recombinants were identified among the offspring of this cross in two ways. First, flies were screened for dark wings, as recombination at the endogenous y^{1} gene would reconstitute a wild type y transcription unit with all enhancers, whereas progeny with ectopic insertions of the replacement y gene would produce flies with lightly colored wings due to the absent
wing enhancer. Second, we conducted genetic analyses to determine whether the $w^+$ phenotype was linked to the $X$ chromosome. Five putative homologous recombination lines were established based on dark wing pigmentation. Further genetic analyses showed that in one line, XGL339-23-38, the $w$ marker mapped to the $X$ chromosome, suggesting a correct targeting event. Southern analyses confirmed the structure of the $y$ gene in these flies (Figure S1). This targeted allele was named, $y_{D1A-2w}$.

We reasoned that if 1A-2 was an insulator in the $y$-$ac$ locus, then deletion of 1A-2 would release constraints on the $y$ and $ac$ enhancers, causing changes in gene expression that would alter cuticle pigmentation and bristle number in $y_{D1A-2w}$ relative to wild type flies [28,29]. However, we found that adult phenotypes of $y_{D1A-2w}$ flies were indistinguishable from wild type flies. In $y_{D1A-2w}$, the $whs$ gene replaced 1A-2. To rule out the possibility that this gene served as a surrogate insulator by carrying a promoter that captured the $y$ and $ac$ enhancers, $y_{D1A-2w}$ flies were crossed to flies carrying a source of Cre recombinase to remove the $whs$ gene. Southern and PCR analyses confirmed the structure of the $y$ gene in $y_{D1A-2}$ flies (Figure S1). Again, the cuticle and bristle phenotypes of the $y_{D1A-2}$ flies were indistinguishable from wild type. Taken together, these data imply that 1A-2 is not an insulator at the endogenous genomic location.

Within the $y$-$ac$ intergenic region, we identified a second cluster of Su(Hw) binding sites, which we called 1A-2. These sites differ from the Su(Hw) consensus sequence at multiple highly conserved positions (Figure 4). Electrophoretic mobility shift assays demonstrated that 1A-2 had 3-fold lower affinity for Su(Hw) than 1A-2. Even so, we considered it possible that weaker 1A-2 Su(Hw) BR might provide a redundant function with 1A-2 to define regulatory interactions in the $y$-$ac$ region. For this reason,
Transcriptional Effects of the 1A-2 Insulator

Figure 5. Ends-out targeting strategy to generate deletions of 1A-2 at the endogenous 1A locus. A. Transgenic flies were generated that carried the mutant y' allele (structure shown where C indicates the mutation of the translation ATG start) at the endogenous X chromosome location (left) and the Py[*1A-2 target] transposon on a different chromosome (right) that carries a y gene lacking the wing enhancer, but encodes a wild type RNA (A indicates the presence of the correct translation ATG start). In this transposon, the y gene, flanked by FRT sites (white arrows) and I-Sce sites, is within a P transposon (inverted black triangles). Other symbols representing the y, yar and ac genes are as described in Figure 1. Transgenic flies y' w[118] Py[*1A-2 target] had a light wing color and orange eyes. B. FLP and I-Sce enzymes catalyzed replacement of the y' allele at the endogenous locus, with y[1A-2w], in which the 1A-2 insulator is substituted by the w' gene (raised triangle) inserted between loxP sites (black arrowheads on raised triangle). The recombinant y'[1A-2w] flies had dark wings and red eyes. C. Cre recombinase deleted the w' gene, leaving behind a single loxP site to form y[1A-2]. In the case of the y[1A-2], the remaining loxP site was mutated, forming a new EcoRV site (RV), whereas in the similarly derived y'[1A-2] flies a wild type loxP site remained. The bar under the y gene indicates the probe used in the Southern analyses (see Figure S1).

doi:10.1371/journal.pgen.1000159.g005

We generated a second targeting vector, Py[*1A-2/1A-2' target], wherein the w' gene replaced an ~1.0 kb deletion that encompassed both 1A-2 and 1A-2'. Following the procedure described above, six putative homologous recombinant lines were identified based on dark wing pigmentation. Further genetic analyses showed that one of these lines, XGL426-41-4, had marker linkage to the X chromosome. This allele was named y[1A-2/1A-2']w. Flies from this line were used to obtain a derivative line y[1A-2]w. In the case of the y[1A-2'w] allele at the endogenous locus, with y[1A-2w], in which the 1A-2 insulator is substituted by the w' gene (raised triangle) inserted between loxP sites (black arrowheads on raised triangle). The recombinant y'[1A-2w] flies had dark wings and red eyes. C. Cre recombinase deleted the w' gene, leaving behind a single loxP site to form y[1A-2]. In the case of the y[1A-2], the remaining loxP site was mutated, forming a new EcoRV site (RV), whereas in the similarly derived y'[1A-2'] flies a wild type loxP site remained. The bar under the y gene indicates the probe used in the Southern analyses (see Figure S1).

doi:10.1371/journal.pgen.1000159.g005

**yar Expression Is Lowered in the Absence of 1A-2.** We postulated that changes in y and ac gene expression might occur, but that these differences may not be readily observed in analyses of adult phenotypes. For this reason, levels of RNA accumulation were quantified using reverse transcriptase PCR (Figure 6). Total RNA was isolated from staged collections of Canton S (wild type), y[1A-2'] and y[1A-2/1A-2'] embryos and pupae, representing the developmental periods where the y and ac genes are maximally expressed [30]. Following conversion to cDNA, templates were amplified using primers against y, ac, yar, and Ras64B, a constitutively expressed RNA [16,31]. In a first set of experiments, PCR products obtained from cDNA amplification in the linear range were run on an agarose gel and visualized by ethidium bromide staining (Figure 6A semi-quantitative, Figure S2). These studies revealed that the loss of 1A-2 and 1A-2' did not change the timing or level of y and ac RNA accumulation, consistent with the lack of phenotypic changes. In contrast, amplification of y[1A-2] and y[1A-2/1A-2'] cDNA showed reduced yar levels relative to Canton S, without a temporal change. These findings indicate that 1A-2 and 1A-2' are required for yar expression.

Quantitative real time PCR analyses (Q-PCR) were undertaken to test the semi-quantitative results (Figure 6B, Figure S3). In these studies, we included analysis of scale (ac) RNA accumulation, the gene downstream of ac. A cycle threshold (C_T) for each primer set was determined and a corresponding ΔC_T was calculated, using the Ras64B C_T for standardization. These analyses identified a significant increase in ΔC_T for yar within y[1A-2'] and y[1A-2/1A-2']
Figure 6. Loss of 1A-2 and 1A2’ reduces yar RNA accumulation during embryogenesis. A. Ethidium bromide stained PCR products obtained from semi-quantitative PCR to evaluate ac, y, yar and Ras 64B RNA levels in wild type (Canton S), y^1A-2 (Δ1A-2) and y^1A-2/1A2’ (Δ1A-2/Δ1A-2’). Ras64B is constitutively expressed and serves as a control. The minus (–) RT lanes control for genomic DNA contamination. Different stages of embryonic and mixed pupal RNA were analyzed. B. Quantitative real time PCR (Q-PCR) was used to determine levels of yar mRNA accumulation from RNAs isolated during development from wild type and mutant lines. Individual transcript levels defined by Q-PCR were normalized to Ras64B for amount of input cDNA (ΔCt). A larger ΔCt indicates a reduction in RNA. Error bars indicate standard deviation of values obtained from analyses of three independently isolated RNAs. Significant changes in RNA accumulation relative to wild type are as indicated (*, P = <0.01; †, P<0.001, Student’s two-tailed t-test).

doi:10.1371/journal.pgen.1000159.g006
samples, relative to Canton S. These data correspond to a 7- and 25-fold decrease in y$^{1A-2}$ embryonic and pupal yar RNA respectively and a 32- and 41-fold decrease in y$^{1A-2/1A-2}$ embryonic and pupal yar RNA (Figure 6B, Figure S3). These data suggest that within the context of the y-ac genomic region, 1A-2 and 1A-2$'$ serve as an enhancer of the non-coding yar gene.

To determine whether the Su(Hw) contributes to expression of genes in the 1A region, we quantified $y$, $yar$, and $ac$ RNAs in a su(Hw)$^+/+$/Su(Hw)$^+/+$ mutant background, using Q-PCR. The su(Hw)$^+$ allele carries a promoter deletion and the su(Hw)$^+$ allele carries a point mutation that produces a full-length protein with an inactive finger 10 [32]. We found that only the level of pupal yar RNA was significantly changed in the su(Hw)$^{+/+}$/Su(Hw)$^{+/+}$ mutant background, associated with an $\sim$21-fold decrease (Figure 6B, Figure S3). These data indicate that Su(Hw) makes positive contribution to the normal low level of pupal yar transcription. The absence of expression changes in embryonic RNA may be confounded by an inability of Su(Hw)$^+$ to bind to 1A-2 in early embryos. Previous studies have shown that disruption of Su(Hw) zinc finger 10 limits chromosome accessibility, without altering DNA recognition [32]. It is possible that in early embryos, 1A-2 is in a more accessible chromatin structure, thereby allowing Su(Hw)$^+$ to bind 1A-2 and activate yar, but that this property is lost during development. We are unable to test yar expression in a su(Hw) null background, as complete loss of Su(Hw) blocks oogenesis. Regardless, our data imply that Su(Hw), along with contributions made by other proteins associated with 1A-2, function as an activator of yar transcription during development.

**Discussion**

Prevailing models of gypsy insulator function predict that the gypsy insulator establishes independent transcriptional domains through cooperation with genomic insulators defined by non-gypsy Su(Hw) BRs. Recent findings indicate that the sequence and organization of non-gypsy BSs differ from the Su(Hw) BR in the gypsy retrovirus [14–16]. These observations imply that properties of non-gypsy BSs may be distinct from those of the gypsy insulator. We defined the properties of 1A-2, to gain insights into mechanisms of Su(Hw) insulator action.

**Enhancer Blocking by 1A-2 Requires Su(Hw) BSs and a Facilitator**

We used the quantitative FBE1-yp2-LacZ reporter system to define the sequence requirements for enhancer blocking by 1A-2(520). Prior application of this system demonstrated that at least four gypsy Su(Hw) sites were needed for robust blocking [20]. Here, we show that 1A-2(157) provided as strong an enhancer blocker as the gypsy insulator (Figures 1, 2). A fragment containing only the Su(Hw) BSs [1A-2(79)] reconstituted a weaker enhancer blocker than 1A-2(157), but had a greater blocking capacity than the synthetic insulators made from reiterated copies of BS3 of the gypsy insulator [20]. While we do not know the reason for the more robust blocking, we note that these regions differ in sequence and distance of separation from Su(Hw) sites (Figure 4). Blocking effectiveness does not appear to be due to differences in DNA recognition, as the in vitro binding constants for Su(Hw) for the 1A-2 and gypsy BSs are similar [16]. Our experiments revealed that 1A-2 contains a second regulatory element located in 1A-2(78). When these sequences were positioned next to the inactive, synthetic Su(Hw) BR (3R:3), a functional insulator was reconstituted (Figure 2B). These data are consistent with previous findings that Su(Hw) chromosome association is limited [32]. Taken together, we propose that 1A-2 is a composite insulator that contains an enhancer blocking and a facilitator function that may improve Su(Hw) chromosome association. Further, we predict that in vivo effectiveness of enhancer blocking by the Su(Hw) protein is related to the accessibility of Su(Hw) BSs. If single or small clusters of Su(Hw) BSs are located in genomic regions of open chromatin, then these regions will demonstrate enhancer blocking, as defined in transgene assays. This proposal implies that genomic context greatly influences the properties of non-gypsy Su(Hw) BSs.

**A Novel Non-Coding RNA Gene Separates the y and ac Genes in the 1A Locus**

1A-2 is located between the independently regulated $y$ and $ac$ genes. Chromatin immunoprecipitation studies demonstrated that 1A-2 is associated with Su(Hw), Mod67.2 and $E(y)^{2}$ in vivo [12,16,18], suggesting that this element binds a complex competent for establishing a genomic insulator. Based on these properties, we postulated that 1A-2 was responsible for the regulatory independence of the $y$ and $ac$ genes in the 1A locus [16]. As a first step in testing this proposal, we investigated transcription in the $y$-ac region to evaluate the current accuracy of the genomic annotation of this region. These studies identified a previously unannotated gene, $yar$, located $\sim$1.2 kb downstream of the $y$ gene and $\sim$3.0 kb upstream of $ac$. Multiple, differentially spliced, polyA$^+$ RNAs are encoded by $yar$, with the largest translation product predicted to be 75 amino acids, indicating that this is a non-coding RNA gene. Emerging data suggest that non-coding RNAs are abundant in eukaryotes and have a wide repertoire of biological functions, ranging from structural components in protein complexes to regulatory molecules involved in transcription and translation [33–35]. It is unknown whether $yar$ has a function. As flies carrying a large genomic deletion that removes sequences upstream of $y$ and extends downstream of $ac$[$y^-ac^+$] are viable and fertile, $yar$ is a non-essential gene.

**1A-2 Is Required for Expression of a Non-Coding RNA Gene**

Having re-defined the transcriptional profile in the 1A locus, we tested the function of 1A-2 and a second, weaker Su(Hw) BR, 1A-2', on gene regulation, using gene targeting to delete these elements. Our studies represent the first deletional analysis of any non-gypsy Su(Hw) BR in the Drosophila genome. Two targeted deletion lines, $y^{1A-2}$ and $y^{1A-2/1A-2}$ were established (Figure 4). Levels of $y$, $ac$, $sc$ and $yar$ RNA accumulation during development were studied using quantitative PCR. We find that loss of 1A-2 and 1A-2' has no effect on the timing and level of $y$, $ac$ or $sc$ RNAs relative to the wild type control (Figure S3), but strongly reduced $yar$ RNA (Figure 6). These data suggest that the effects of loss of 1A-2 are limited to local changes of gene expression, implying that these sequences are not a chromatin insulator at the endogenous location. Instead, our data indicate that 1A-2 may be an activator of $yar$ expression, consistent with other studies that have suggested a role for Su(Hw) in gene activation [36–38]. These data, coupled with genetic studies on the effects of the loss of Su(Hw) on expression of genes adjacent to Su(Hw) BRs [16], demonstrate that Su(Hw) BRs have diverse functions in the genome.

The complexity of the transcriptional effects associated with Su(Hw) BRs is reminiscent of regions in mammalian genomes that bind the versatile regulatory protein CTCF. High throughput genomic analyses have identified hundreds of CTCF binding sites within the mouse and human genomes [7,39–41]. Although many of these sequences possess enhancer blocking activity [39,42,43], CTCF has been implicated in transcriptional activation [44–46], repression [47–50], and chromosome pairing [44,51,52]. These
observations suggest that, similar to the non-gypsy Su(Hw) BRs, genomic context will have an important influence on the properties of CTCF BRs within a given region.

The mechanism(s) used to maintain transcriptional autonomy in the IA locus are unclear. The discovery of yar provides an alternative explanation to the need for a chromatin insulator. Based on the developmental timing displayed by the IA genes, we postulate that activation of yar transcription may cause inactivation of ac through transcriptional interference. Similarly, activation of y may repress yar transcription. Although y\(^i\)A-2 and y\(^i\)A-2'/A-2 y flies show reduced yar expression, transcription is not abolished, suggesting that the remaining yar activity may be sufficient to turn off ac. Alternatively, other mechanisms can be considered that might influence enhancer preference, including selectivity of enhancers for certain classes of promoters [53,54], the presence of promoter targeting sequences that direct enhancer action [55,56], or promoter tethering elements that capture enhancers [57]. Further experiments to define the properties of DNA elements within the IA locus will resolve how transcriptional independence is achieved.

Materials and Methods
Fly Stocks and Crosses
Flies were raised at 25°C, 70% humidity on standard corn meal/agar medium. Description of the alleles used can be found at http://flybase.bio.indiana.edu.

Construction of FBE1-yp2 -LacZ Reporter Genes
The FBE1-yp2 -LacZ fusion gene [20] carried a BgllI site, positioned at +335 relative to the transcription start site (TSS) that was used for insertion of tested 1A-2 fragments. Resulting transgenes were inserted into a P element transformation vector, generating P[F-IA-2(520)-yp2] with the full length 1A-2, P[F-IA-2(157)-yp2] with a 157 bp region of 1A-2, P[F-IA-2(79)-yp2] with two 1A-2 Su(Hw) binding sites, P[F-IA-2(78)-yp2] with the 78 bp 3' region, P[F-IA-2(78×4)-yp2] with four tandem repeats of the 1A-2 78 bp element and P[F-3R(78)-yp2] with a hybrid insertion between a cluster of three tandem repeats of the gypsy Su(Hw) binding sites [nucleotides 732–759 [58]], as described in [20] and the 78 bp element. P transposons were injected into the host y\(^i\)w\(^{118}\) strain or y\(^i\)w\(^{118}\) mutant background. Gene targeting followed the procedure outlined in [59]. A combination of Southern and PCR analyses identified correctly targeted events. To remove the w\(^{118}\) gene, red-eyed males carrying a targeted deletion event were crossed to females carrying Cre recombinase, as described in [62]. The white-eyed flies were collected and used to establish homozygous stocks. Deletion events were confirmed by PCR amplification and sequence analyses.

Rapid Amplification of cDNA Ends (RACE)
The structures of the yar RNAs were determined using RACE of total RNA isolated from 6–12 hour CS embryos. In the 3'–RACE experiments, 5 μg of RNA were reverse transcribed using the adaptor oligo-dT primer (5'–RACE kit, Invitrogen), and cDNA was amplified using a yar specific primer (1 μM) and the abridged universal primer (80 nM, Invitrogen). Several products were identified by agarose gel electrophoresis, gel purified and cloned into the TOPO vector (Invitrogen). Sequencing and BLAST search identified three yar splice variants that shared a common distal exon and poly-A signal. In the 5'–RACE experiments, 5 μg of RNA were reverse transcribed with a yar specific primer (100 nM), purified over a S.N.A.P column (Invitrogen) to remove unincorporated nucleotides and primers, and C-tailed at 4°C for 2 hours, using terminal deoxynucleotidyl transferase. Tailed cDNAs were amplified with nested yar specific primers (400 nM) and an abridged anchor primer (400 nM, Invitrogen). PCR products were directly cloned into the TOPO vector. Forty-eight clones were analyzed by restriction digestion, revealing nine classes of insert. At least one representative of each class was sequenced. BLAST analyses of these data identified ten alternative splice variants and three alternative start sites. Both the 3'–RACE and 5'–RACE were performed on two independent RNA isolations. Gene-specific primer sequences are available upon request.

Northern and Real-Time PCR Analyses
RNA was isolated from staged embryos collected from cages of wild type (CS) flies, using the NaDodSO4/phenol technique [63]. Five μg of oligo-dT selected polyA+ RNA was used in northern analyses and hybridized with radiolabeled fragments corresponding to y (a ClaI-BglII fragment, representing +2466 to +4815 relative to the yTSS), yar (EST DN154052, 418 bp ) and ac (a PCR fragment representing +115 to +331 relative to the aTSS). Hybridization with sequences corresponding to the ribosomal gene, RpL32, served as a loading control. For real-time PCR experiments, RNA was isolated from embryos and pupae from three lines: CS, y\(^{Al}A-2\) line XGL339-23-8, y\(^{Al}A-2\) line XGL426-41-4. RNA isolation and real-time PCR analyses were performed as described in [16]. PCR primers amplified 100–200 bp fragments, y primers flanked the intron, yar primers were in the invariant fourth exon, to ensure quantification of all transcripts. Primer sequences are available upon request. Duplicate or triplicate reactions were performed and averaged, with the difference among the replicates no greater than 0.5 cycle threshold (C\(_T\)). At least three independent experiments were performed for
each primer set from two independent RNA samples. The expression level of each gene was determined using RsaI-4B as an internal control (ΔCT). The fold change in expression of each gene relative to the wild type (CS) value was determined with the ΔΔCT method.

Supporting Information

Figure S1 Southern analysis of y-ac locus in homologous recombinant lines. Genomic DNA was isolated from ten flies, digested with EcoRV (NEB) and run on a 1% agarose gel. Flies analyzed were the parental y\textsuperscript{u}\textsuperscript{11} flies, the P\textsuperscript{y\textsubscript{1A-2}}/y\textsuperscript{1A-2\textsubscript{x}} target or P\textsuperscript{y\textsubscript{1A-2}}/y\textsuperscript{1A-2\textsubscript{x}} transgenic (TG) lines, homologous recombinants carrying the y\textsuperscript{1A-2\textsubscript{e}} gene (P\textsuperscript{y\textsubscript{1A-2\textsubscript{e}}}) and homologous recombinants deleted for y\textsuperscript{1A-2\textsubscript{i}} gene (y\textsuperscript{1A-2\textsubscript{i}-y\textsuperscript{1A-2\textsubscript{i}}}). DNAs were transferred to Nytran and hybridized with a 32P-labeled probe made with Clal to EcolI fragment of y gene (black bar, Figure 5). The probe recognizes an endogenous band of 7.6 kb in y\textsuperscript{1A-2\textsubscript{e}} flies, and transgene band of 4.5 kb. Correct recombinant events removed the endogenous band. Excision of y\textsuperscript{1A-2\textsubscript{i}} gene with Cre recombinase lead to appearance of a new EcoRV site at the LoxP element in y\textsuperscript{1A-2\textsubscript{i}} line (3 kb band). A similar event did not occur in the y\textsuperscript{1A-2\textsubscript{i}-y\textsuperscript{1A-2\textsubscript{i}}} line, therefore a smaller band is seen due to the ~1.0 kb deletion of the Su(Hw) BS (6.7 kb band).

Found at: doi:10.1371/journal.pgen.1000159.s001 (5.15 MB TIF)

Figure S2 Definition of parameters for semi-quantitative PCR analyses. Indicated volumes of cDNA were used as a template for amplification by the ac, yar, y and RsaI-4B primers for the number of cycles shown at the right. Ethidium-stained PCR products from Cre recombinase lead to appearance of a new element in LoxP (Figure 5). The probe recognizes an endogenous band of 7.6 kb in y\textsuperscript{1A-2\textsubscript{e}} flies and transgene band of 4.5 kb. Correct recombinant events removed the endogenous band. Excision of y\textsuperscript{1A-2\textsubscript{i}} gene with Cre recombinase lead to appearance of a new EcoRV site at the LoxP element in y\textsuperscript{1A-2\textsubscript{i}} line (3 kb band). A similar event did not occur in the y\textsuperscript{1A-2\textsubscript{i}-y\textsuperscript{1A-2\textsubscript{i}}} line, therefore a smaller band is seen due to the ~1.0 kb deletion of the Su(Hw) BS (6.7 kb band).

Found at: doi:10.1371/journal.pgen.1000159.s002 (2.17 MB TIF)

References

1. Gaszner M, Felsenfeld G (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet 7: 703–715.
2. Kuhn EJ, Geyer PK (2003) Genomic insulators: connecting properties to mechanism. Curr Opin Cell Biol 15: 259–265.
3. Valenzuela L, Kamakaka RT (2006) Chromatin insulators. Annu Rev Genet 40: 107–130.
4. Maeda RK, Karch F (2007) Making connections: boundaries and insulators in Drosophila. Curr Opin Genet Dev 17: 394–399.
5. Dorman ER, Bushey AM, Corces VG (2007) The role of insulator elements in large-scale chromatin structure in interphase. Semin Cell Dev Biol 18: 682–690.
6. Engel N, Bartolomei MS (2003) Mechanisms of insulator function in gene expression. Int Rev Cytol 232: 89–127.
7. Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, et al. (2007) Large-scale chromatin structure in interphase. Semin Cell Dev Biol 18: 682–690.
8. Corces VG, Geyer PK (1991) Interaction of retrotransposons with the host genome: the case of the gypsy element of Drosophila. Trends in Genetics 7: 86–90.
9. Pai CY, Lei EP, Ghosh D, Corces VG (2004) The centromosomal protein CP190 is a component of the gypsy chromatin insulator. Mol Cell 16: 737–740.
10. Ghosh D, Gerasimova TI, Corces VG (2001) Interactions between the Su(Hw) and Medu(gag) genes required for gypsy insulator function. Embryo J 20: 2518–2527.
11. Gerasimova TI, Ghala DA, Gerasimov D, Simonova O, Corces VG (1995) A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. Cell 82: 587–597.
12. Kurfeshkova M, Maksimenko O, Golovnin A, Pulina M, Georgieva S, et al. (2007) Evolutionarily conserved Eys/2/Sus1 protein is essential for the barrier activity of Su(Hw)-dependent insulators in Drosophila. Mol Cell 27: 332–338.
13. Georgiev PG, Gerasimova TI (1989) Novel genes influencing the expression level of each gene was determined using RsaI-4B as an internal control (ΔCT). The fold change in expression of each gene relative to the wild type (CS) value was determined with the ΔΔCT method.

Supporting Information

Figure S1 Southern analysis of y-ac locus in homologous recombinant lines. Genomic DNA was isolated from ten flies, digested with EcoRV (NEB) and run on a 1% agarose gel. Flies analyzed were the parental y\textsuperscript{u}\textsuperscript{11} flies, the P\textsuperscript{y\textsubscript{1A-2}}/y\textsuperscript{1A-2\textsubscript{x}} target or P\textsuperscript{y\textsubscript{1A-2}}/y\textsuperscript{1A-2\textsubscript{x}} transgenic (TG) lines, homologous recombinants carrying the y\textsuperscript{1A-2\textsubscript{e}} gene (P\textsuperscript{y\textsubscript{1A-2\textsubscript{e}}}) and homologous recombinants deleted for y\textsuperscript{1A-2\textsubscript{i}} gene (y\textsuperscript{1A-2\textsubscript{i}-y\textsuperscript{1A-2\textsubscript{i}}}). DNAs were transferred to Nytran and hybridized with a 32P-labeled probe made with Clal to EcolI fragment of y gene (black bar, Figure 5). The probe recognizes an endogenous band of 7.6 kb in y\textsuperscript{1A-2\textsubscript{e}} flies, and transgene band of 4.5 kb. Correct recombinant events removed the endogenous band. Excision of y\textsuperscript{1A-2\textsubscript{i}} gene with Cre recombinase lead to appearance of a new EcoRV site at the LoxP element in y\textsuperscript{1A-2\textsubscript{i}} line (3 kb band). A similar event did not occur in the y\textsuperscript{1A-2\textsubscript{i}-y\textsuperscript{1A-2\textsubscript{i}}} line, therefore a smaller band is seen due to the ~1.0 kb deletion of the Su(Hw) BS (6.7 kb band).

Found at: doi:10.1371/journal.pgen.1000159.s001 (5.15 MB TIF)

Figure S2 Definition of parameters for semi-quantitative PCR analyses. Indicated volumes of cDNA were used as a template for amplification by the ac, yar, y and RsaI-4B primers for the number of cycles shown at the right. Ethidium-stained PCR products from Cre recombinase lead to appearance of a new element in LoxP (Figure 5). The probe recognizes an endogenous band of 7.6 kb in y\textsuperscript{1A-2\textsubscript{e}} flies, and transgene band of 4.5 kb. Correct recombinant events removed the endogenous band. Excision of y\textsuperscript{1A-2\textsubscript{i}} gene with Cre recombinase lead to appearance of a new EcoRV site at the LoxP element in y\textsuperscript{1A-2\textsubscript{i}} line (3 kb band). A similar event did not occur in the y\textsuperscript{1A-2\textsubscript{i}-y\textsuperscript{1A-2\textsubscript{i}}} line, therefore a smaller band is seen due to the ~1.0 kb deletion of the Su(Hw) BS (6.7 kb band).

Found at: doi:10.1371/journal.pgen.1000159.s002 (2.17 MB TIF)

Acknowledgments

We thank Lori Walrath and the Geyer laboratory for critically reading this manuscript. We thank Jinsil Kim, Bill Chen and Jonathan Zuck for technical assistance. We thank Yikang Rong for providing pHL25 and Kent Golic for providing the 70FLP and 70FR-Sed flies.

Author Contributions

Conceived and designed the experiments: AAS XL MDW PKG. Performed the experiments: AAS XL MDW PKG. Analyzed the data: AAS XL MDW PKG. Contributed reagents/materials/analysis tools: AAS PKG. Wrote the paper: AAS PKG.

Transcriptional Effects of the 1A-2 Insulator
32. Kuhn-Parnell EJ, Helou C, Marion DJ, Gilmore BL, Parnell TJ, et al. (2008) Investigation of the properties of non-gypsy Su(Hw) binding sites. Genetics in press.

33. Stefani G, Slack FJ (2008) Small non-coding RNAs in animal development. Nat Rev Mol Cell Biol 9: 219–230.

34. Barrandon C, Sulllotti B, Benasude O (2008) Non-coding RNAs regulating the transcriptional machinery. Biol Cell 100: 81–95.

35. Maitick JS, Makunin IV (2006) Non-coding RNA. Hum Mol Genet 15 Spec No 1: R17–29.

36. Parkhurst SM, Corces VG (1986) Interactions among the gypsy transposable element and the yellow and the suppressor of hairy-wing loci in Drosophila melanogaster. Mol Cell Biol 6: 47–53.

37. Golovnin A, Melnick E, Mazur A, Georgiev P (2005) Drosophila Su(Hw) insulator can stimulate transcription of a weakened yellow promoter over a distance. Genetics 170: 1133–1142.

38. Markstein M, Pitsouli C, Villalta C, Celniker SE, Perrimon N (2008) Exploring non-coding RNA function in Drosophila. In: Kedes L, Illmensee K, editors. Molecular mechanisms of gene regulation in Drosophila. Cold Spring Harbor Laboratory Press; 2008. p. 639–58.

39. Spradling AC, Mahowald AP (1979) Identification and genetic localization of the Drosophila melanogaster Hsp70 genes. Genetics 112: 519–534.

40. Gong WJ, Golic KG (2004) Genomic deletions of the Drosophila melanogaster Hsp70 genes. Genetics 168: 1467–1476.

41. Gong WJ, Golic KG (2003) CTCF mediates transcriptional regulation mediated by thyroid hormone response element 144 repressor CTCF. Mol Cell Biol 17: 1281–1290.

42. Lutz M, Barke Les J, Barreto G, Geoman F, Greb H, et al. (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Research 28: 1707–1713.

43. Awad TA, Bigler J, Ulmer JE, Hu YJ, Moore JM, et al. (1999) Negative transcriptional regulation mediated by thyroid hormone response element 144 requires binding of the multivalent factor CTCF to a novel target DNA sequence. J Biol Chem 274: 27092–27098.

44. Lutz M, Barke Les J, Barreto G, Rong YS, Golic KG; et al. (1998) Gene targeting by homologous recombination in Drosophila melanogaster. Proc Natl Acad Sci U S A 95: 14028–14033.

45. Lutz M, Barke Les J, Barreto G, Rong YS, Golic KG; et al. (1998) Gene targeting by homologous recombination in Drosophila melanogaster. Proc Natl Acad Sci U S A 95: 14028–14033.

46. Lutz M, Barke Les J, Barreto G, Rong YS, Golic KG; et al. (1998) Gene targeting by homologous recombination in Drosophila melanogaster. Proc Natl Acad Sci U S A 95: 14028–14033.

47. Lutz M, Barke Les J, Barreto G, Rong YS, Golic KG; et al. (1998) Gene targeting by homologous recombination in Drosophila melanogaster. Proc Natl Acad Sci U S A 95: 14028–14033.

48. Lutz M, Barke Les J, Barreto G, Geoman F, Greb H, et al. (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Research 28: 1707–1713.