Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome

Judith Jans, John M. Gladden, Edward J. Ralston, et al.

Genes Dev. 2009 23: 602-618
Access the most recent version at doi:10.1101/gad.1751109

Supplemental Material http://genesdev.cshlp.org/content/suppl/2009/03/05/23.5.602.DC1.html

References This article cites 41 articles, 19 of which can be accessed free at:
http://genesdev.cshlp.org/content/23/5/602.full.html#ref-list-1

Article cited in:
http://genesdev.cshlp.org/content/23/5/602.full.html#related-urls

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

To subscribe to *Genes & Development* go to:
http://genesdev.cshlp.org/subscriptions

Copyright © 2009 by Cold Spring Harbor Laboratory Press

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome

Judith Jans,[1,2] John M. Gladden,[1] Edward J. Ralston, Catherine S. Pickle, Agnès H. Michel, Rebecca R. Pferdehirt, Michael B. Eisen, and Barbara J. Meyer[3]

Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 97420, USA

In many species, a dosage compensation complex (DCC) is targeted to X chromosomes of one sex to equalize levels of X-gene products between males (1X) and females (2X). Here we identify *cis*-acting regulatory elements that target the *Caenorhabditis elegans* X chromosome for repression by the DCC. The DCC binds to discrete, dispersed sites on X of two types. *rex* sites (recruitment elements on X) recruit the DCC in an autonomous, DNA sequence-dependent manner using a 12-base-pair (bp) consensus motif that is enriched on X. This motif is critical for DCC binding, is clustered in *rex* sites, and confers much of X-chromosome specificity. Motif variants enriched on X by 3.8-fold or more are highly predictive (95%) for *rex* sites. In contrast, *dox* sites (dependent on X) lack the X-enriched variants and cannot bind the DCC when detached from X. *dox* sites are more prevalent than *rex* sites and, unlike *rex* sites, reside preferentially in promoters of some expressed genes. These findings fulfill predictions for a targeting model in which the DCC binds to recruitment sites on X and disperses to discrete sites lacking autonomous recruitment ability. To relate DCC binding to function, we identified dosage-compensated and noncompensated genes on X. Unexpectedly, many genes of both types have bound DCC, but many do not, suggesting the DCC acts over long distances to repress X-gene expression. Remarkably, the DCC binds to autosomes, but at far fewer sites and rarely at consensus motifs. DCC disruption causes opposite effects on expression of X and autosomal genes. The DCC thus acts at a distance to impact expression throughout the genome.

[*Keywords*: Dosage compensation; condensin; X chromosome; gene expression; epigenetics; *C. elegans*]

Supplemental material is available at http://www.genesdev.org.

Received October 10, 2008; revised version accepted January 20, 2009.

Gene expression in metazoans is controlled by diverse regulatory mechanisms that function over dramatically different distances. Some mechanisms act locally on individual genes, while others affect expression of genes across large subchromosomal domains (e.g., imprinted gene clusters, β-globin locus) (Palstra et al. 2008; Wan and Bartolomei 2008), or along entire chromosomes (e.g., X-chromosome dosage compensation) (Meyer 2005) in a coordinated manner. In some cases, *cis*-acting control elements also act in *trans* to regulate genes on separate chromosomes (Morris et al. 1999; Lomvardas et al. 2006; Osborne et al. 2007). The challenge presented by chromosome-wide regulatory mechanisms is to define the *cis*-acting control elements and determine whether these elements act locally on individual genes to regulate expression, or instead act over long range, perhaps by inducing changes in chromosome structure. X-chromosome dosage compensation is exemplary for such analysis because its function is essential, it distinguishes X chromosomes from autosomes, and it controls hundreds of genes simultaneously.

The strategies for dosage compensation differ (Lucchesi et al. 2005; Meyer 2005). Mammals inactivate one X chromosome in females, flies increase expression of the single X chromosome in males, and nematodes reduce expression of both X chromosomes in hermaphrodites. Nevertheless, in all cases, a regulatory complex is targeted to the X chromosome of one sex to modulate transcript levels. The dosage compensation complex (DCC) of the nematode *Caenorhabditis elegans* includes five proteins with homology with subunits of condensin, a conserved protein complex that promotes the compaction, resolution, and segregation of chromosomes during mitosis and meiosis (Chuang et al. 1994; Lieb et al. 1996, 1998; Hirano 2005; Tsai et al. 2008). In fact, the DCC shares a subunit

[1]These authors contributed equally to this work.
[2]Present address: Department of Experimental Oncology, UMC Utrecht, STR2.118, Universiteitsweg 100, Utrecht, The Netherlands.
[3]Corresponding author.
E-MAIL bjmeyer@berkeley.edu; FAX (510) 643-5584.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.1751109.

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

with the otherwise-distinct *C. elegans* mitotic condensin complex (Lieb et al. 1998; Hagstrom et al. 2002). Additional DCC subunits confer sex specificity to the dosage compensation process and recruit the condensin-like DCC subunits to X chromosomes (Davis and Meyer 1997; Dawes et al. 1999). Homology of the DCC to condensin and participation of DCC components in other chromosome behaviors, including mitotic chromosome segregation and meiotic recombination (Tsai et al. 2008), suggest the DCC may impose changes in chromosome structure to achieve chromosome-wide repression. Regulation of gene expression by condensin components extends beyond dosage compensation, to transcriptional silencing in yeast, position-effect variegation in flies, and inhibition of mammalian gene transcription during mitosis, suggesting a general function for these proteins in establishing a repressed state (Bhalla et al. 2002; Dej et al. 2004; Cobbe et al. 2006; Xing et al. 2008). Thus, lessons learned from *C. elegans* dosage compensation will apply broadly to multiple forms of gene control.

The mechanism by which the *C. elegans* DCC is targeted to X chromosomes and the relationship between DCC binding and DCC function are not understood. We address both areas. Our prior studies had surveyed large regions of X (1–5 Mb) for the ability to recruit the DCC when detached from X (Csankovszki et al. 2004). We found regions capable of robust DCC binding and regions capable of weak or no detectable binding, yet all regions showed strong DCC binding when part of an intact X. Those findings suggested a model of DCC targeting in which the DCC binds to recruitment sites distributed along X and disperses to regions lacking recruitment ability. Proof that DCC recruitment is mediated by small, discrete recruitment elements on X (*rex* sites) came from our subsequent dissection of cosmids in recruiting regions using an assay for recruitment activity in vivo (McDonel et al. 2006). The four *rex* sites discovered by that approach have two distinct, clustered DNA motifs shown by mutational analysis to contribute significantly to DCC recruitment in vivo (McDonel et al. 2006). A subsequent study by others (Ercan et al. 2007) identified and extended the more important *rex* motif in their strongest DCC-binding foci obtained by chromatin immunoprecipitation (ChIP) experiments. In addition, Ercan et al. (2007) inferred from the preferential location of DCC-binding sites in 5′ ends of genes that the DCC represses transcription of the genes it binds. However, DCC-binding foci in the Ercan et al. (2007) study were not assessed either for their DCC recruiting ability in vivo or their ability to control expression of nearby genes. Our coupled genome-wide analysis of DCC binding and gene expression evaluates that model. Unexpectedly, we found no direct correlation between DCC binding and DCC-mediated repression. While we do not rule out the possibility that the DCC might act locally on a nearby gene, we show that for numerous genes the DCC appears to act over long range.

In the current study, we identify *cis*-acting elements central for DCC recruitment and binding. To do so, we conducted a genome-wide search for DCC-binding sites and then assayed the sites in vivo for recruitment ability. The sites partition into two classes: sites that autonomously recruit the DCC when detached from X (called *rex* sites), and sites that bind the DCC in their native context on X but fail to recruit the DCC when detached from X (called dependent on X sites, *dox* sites). These findings fulfill predictions for a DCC targeting model in which the DCC binds to specific recruitment sites on X and distributes to sites lacking autonomous recruitment ability. Long distances separate *rex* and *dox* sites, implying that long-range interactions are important for DCC distribution. We also correlated DCC binding with DCC-mediated gene regulation. To do so, we identified dosage-compensated and noncompensated genes on X and compared the locations of these genes with the positions of DCC-binding sites. Many genes of both classes have DCC peaks, and many do not. Our analysis of DCC binding and its impact on gene expression was genome-wide, revealing that the DCC binds to autosomes as well as X chromosomes and influences gene expression throughout the genome. We propose the condensin-like DCC reconfigures the architecture of the chromosome to achieve DCC distribution and gene regulation over long range.

Results

Identification of cis*-acting regulatory elements that target the X chromosome for repression*

Two approaches were used to identify DCC-binding sites on X. First, all cosmids from a 2-Mb recruiting region of X were introduced into worms and assayed for their ability to recruit the DCC in vivo (Fig. 1A). Recruitment was assessed by the degree of DCC binding in transgenic animals to extrachromosomal arrays carrying multiple copies of individual cosmids (assay shown in Fig. 1C; McDonel et al. 2006). Recruitment activity was then ascribed to successively smaller DNA fragments using array assays, thereby defining *rex* sites. This approach identified 13 new *rex* sites (*rex-5* to *rex-17*) (Fig. 1A; Supplemental Table 1).

Second, DCC-binding sites were identified without regard to recruitment ability through a series of 11 biologically independent ChIP reactions performed with antibodies to four different DCC components, SDC-2, SDC-3, DPY-27, and MIX-1. The resulting DNA was hybridized to genome-wide, high-resolution Nimblegen tiling arrays (see the Materials and Methods; Supplemental Material; Supplemental Fig. 1A,B). Binding sites on X chromosomes of embryos defined by these ChIP–chip experiments were in excellent agreement (Supplemental Fig. 1A,B). Landscapes for the condensin-like subunits DPY-27 and MIX-1 matched the landscape for the hermaphrodite-specific protein SDC-2, which triggers assembly of the DCC onto X, and the zinc-finger protein SDC-3, which acts with SDC-2 to recruit the DCC to X. Although MIX-1 also participates in a mitotic condensin complex, the MIX-1 profile had few extra peaks compared with the profile of two dosage compensation-specific proteins DPY-27 and SDC-3. Either the bulk of MIX-1 binding we observe is due to its association with the

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

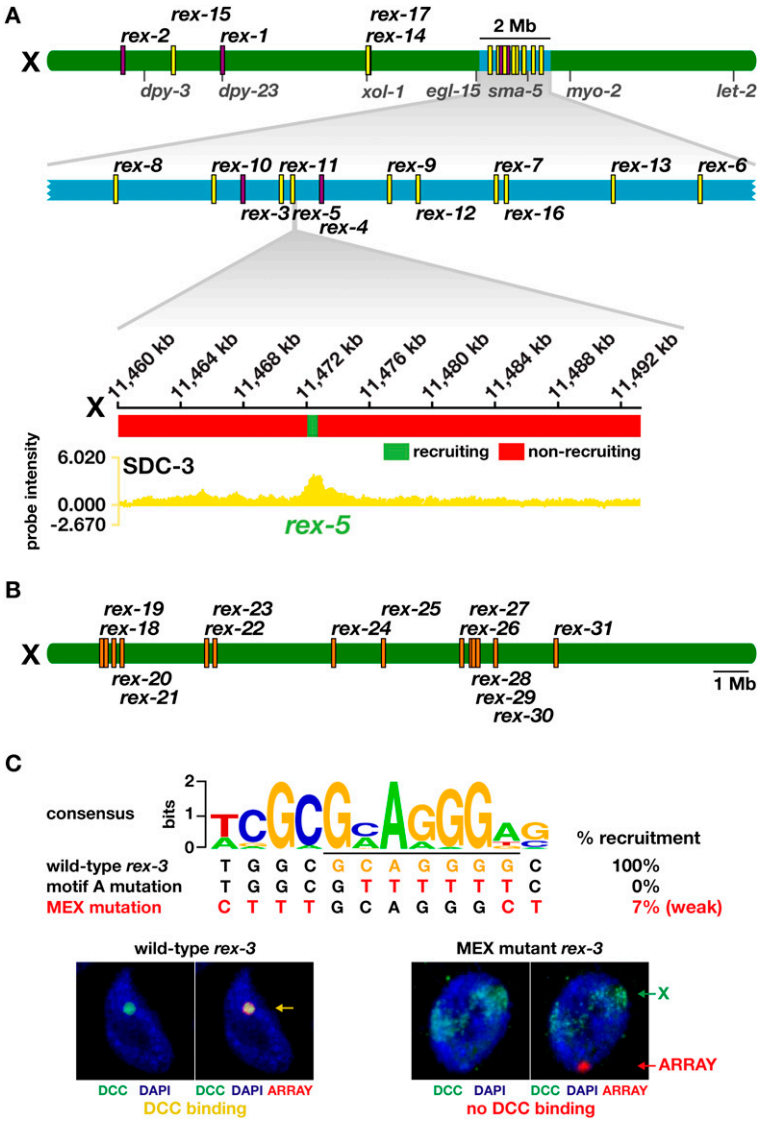

Figure 1. DCC *rex* sites share a 12-bp consensus motif that is critical for DCC recruitment. (*A*) Partial DCC recruitment map of the *C. elegans* X chromosome. Shown are the positions of *rex* sites discovered by assaying cosmids for DCC recruitment in vivo and subsequently assaying subcosmid fragments to delimit the recruitment activity. Purple, *rex* sites defined previously (McDonel et al. 2006); yellow, *rex* sites defined in this study. In blue is an enlargement of a 2-Mb region shown previously to recruit the DCC when detached from X (Csankovszki et al. 2004) and assayed here for recruiting cosmids. Juxtaposition of DCC recruitment data for the *rex-5* genomic region and SDC-3 ChIP–chip landscape show that the 421-bp *rex-5* site (light green) is coincident with an SDC-3 peak, and no DCC recruitment was found for DNA without peaks in array assays. (*B*) Positions of *rex* sites (orange) discovered from DCC ChIP–chip peaks assayed for DCC recruitment to extrachromosomal arrays. (*C*) Mutational analysis of the 115-bp *rex-3* site shows the importance of the full 12-bp consensus motif in DCC recruitment. Shown is the 12-bp consensus motif derived from *rex* sites along with wild-type and mutant *rex-3* motif sequences that were assayed for DCC recruitment. The previously defined motif A is highlighted in yellow in the wild-type *rex-3* sequence and underlined in the consensus motif. The nucleotides substituted in the mutant sites are shown in red. Confocal images show intestinal cell nuclei (DNA stained with 4,6-diamidino-2-phenylindole (DAPI), blue) from transgenic animals carrying extrachromosomal arrays (labeled with a FISH probe, red) with multiple copies of wild-type or mutant *rex-3* sequences costained with SDC-3 antibodies (green). Mutation of the 7-bp motif A within the 12-bp consensus motif of *rex-3* abolished DCC binding. Similarly, mutation of the 5-bp flanking motif A abrogated DCC binding; extrachromosomal arrays in only 11 of 168 (7%) nuclei showed patchy SDC-3 staining (not shown) and arrays in the remaining 93% of nuclei showed no SDC-3 staining (shown). The wild-type *rex-3* array titrates most of the DCC from X, rendering X staining with SDC-3 antibodies undetectable (McDonel et al. 2006). The MEX mutant array fails to titrate SDC-3 from X, which appears green in the image.

DCC, or condensin has a similar binding profile as the DCC. Concordance between the cosmid-derived *rex* sites and the DCC peaks was excellent, validating our ChIP–chip experiments (Fig. 1A). Of 11 *rex* sites defined precisely enough (≤5 kb) for comparison, all corresponded to a DCC peak (Supplemental Table 1).

To identify new *rex* sites, 2-kb fragments of DNA corresponding to the DCC ChIP–chip peaks were assessed in vivo for recruitment ability using extrachromosomal array assays (Fig. 1B). From 63 ChIP–chip peaks assayed, 14 new *rex* sites (*rex-18*–*rex-31*) were defined (Fig. 1B; Supplemental Table 1).

A DNA consensus motif critical for DCC recruitment discriminates X chromosomes from autosomes

Sequence analysis of all 31 *rex* sites using the motif-finding program wconsensus (Hertz and Stormo 1999) revealed a robust 12-base-pair (bp) motif (Fig. 1C; Supplemental Fig. 2C,D). This motif closely resembles the 12-bp motifs derived independently from the 17 *rex* sites obtained from cosmids and the 14 *rex* sites obtained from DCC ChIP–chip peaks (Supplemental Fig. 2A–D). The 12-bp motif from all *rex* sites is an extended version of both the 7-bp motif A discovered previously in our four original *rex* sites (Fig. 1C; McDonel et al. 2006) and the 10-bp version of motif A correlated with the largest ChIP–chip peaks in Ercan et al. (2007). Prior mutational analysis of *rex-1*–*rex-4* showed that A motifs are critical for DCC recruitment and that A motifs act in combination with each other and with a second motif, called motif B (8 bp), to recruit the DCC (McDonel et al. 2006). Multiple A motifs could compensate for the lack of a B motif in that analysis, indicating that the A motif is more important for recruitment. We found B motifs in all new *rex* sites but found no improvement in the B consensus sequence.

Our mutational analysis established the functional significance of the additional consensus sequences in the

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

12-bp motif. Mutation of five bases flanking motif A in *rex-3* disrupted DCC recruitment in array assays, despite the presence of an intact motif A (Fig. 1C), thereby showing that the additional sequences are important for recruitment. Each A motif in *rex-1–rex-4* forms the core of a 12-bp motif, making the 12 mer a predominant motif for recruitment. All *rex* sites have the 12-bp motif, and all but two sites have multiple copies (Supplemental Table 1), consistent with the importance of motif clustering in DCC recruitment, as found previously (McDonel et al. 2006).

In contrast to the previously identified motif A or its 10-bp variant, the 12-bp version of motif A is highly enriched on X chromosomes compared with autosomes (Fig. 2A). Hence, we named the 12-bp motif MEX (Motif Enriched on X). The MEX motif variants with the best matches to the consensus sequence ($\ln[P] \leq -15$) are enriched on X from 3.8-fold to 25-fold. Thus, at least part of the specificity in recruiting the DCC to X is conferred by the prevalence of the MEX motif on the X chromosome. The B motif alone is not significantly enriched on X.

High-scoring MEX motifs on X are predictive for rex sites

An X-enriched consensus motif central for DCC recruitment and loading should meet three further expectations. The motif should be distributed widely on X to nucleate chromosome-wide spreading. It should be coincident with DCC peaks defined by ChIP–chip experiments. The motif variants with the highest enrichment on X should be predictive for *rex* sites. All three expectations were met. For example, the 34 motifs with the best matches ($\ln[P] \leq -15$) to the MEX motif are dispersed along the entire X, and 32 are coincident with large DCC peaks (Fig. 2B). For all seven cases in which high-scoring MEX motifs cluster within 1 kb and reside in one DCC peak, the peaks are unusually large (top 1%), implying very high DCC occupancy, as expected from motif collaboration in DCC recruitment. Ten of 25 peaks correspond to known *rex* sites. Of 15 peaks not in the *rex* set, we assayed seven with a range of MEX scores and found all to recruit the DCC (*rex-32* to *rex-38*) (Fig. 2B; Supplemental Table 1). Thus, high-scoring MEX motifs on X are predictive for *rex* sites.

A total of 38 high-scoring MEX variants are distributed randomly among all five autosomes, but only three motifs correspond to DCC peaks (Fig. 2B; data not shown). The general density of 12-bp motifs on autosomes may be insufficient for DCC recruitment, or possibly some other feature, such as flanking DNA sequence or chromatin-associated proteins, prevents DCC binding to the motifs in the autosomal context. Thus, the 12-bp MEX motif confers much, but not all, of the X specificity in DCC recruitment.

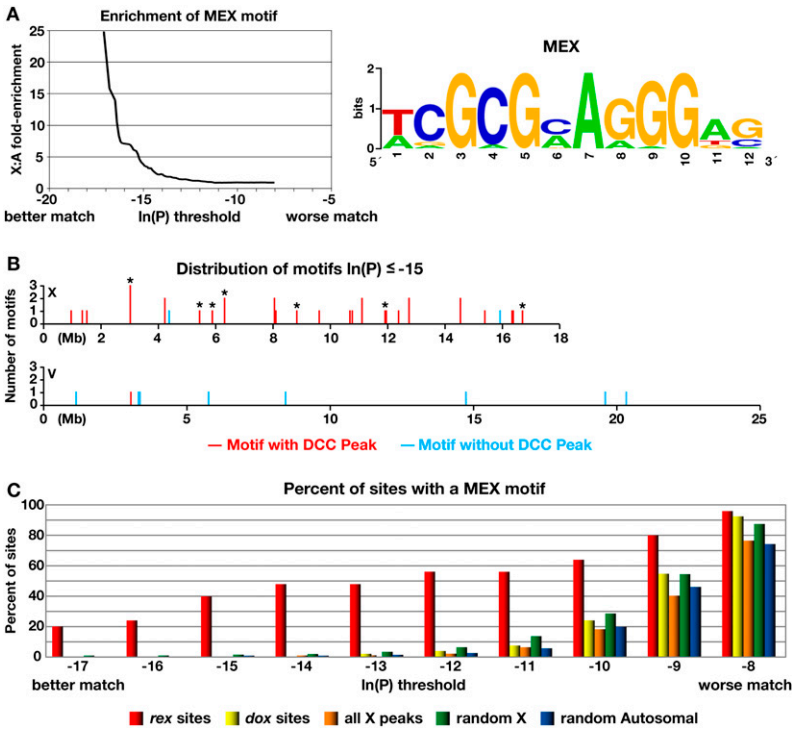

Figure 2. MEX consensus motif from *rex* sites distinguishes the X chromosome from autosomes and *rex* sites from *dox* sites. (*A*) The 12-bp MEX consensus motif from *rex* sites is enriched on the X chromosome relative to autosomes. Plot shows the fold enrichment (*Y*-axis) of MEX motif variants (*X*-axis) on X compared with autosomes. ln(P) is the natural log of the probability of finding a 12 mer that matches the MEX consensus motif matrix as calculated by the Patser program. X:A enrichment is greater for 12 mers with better matches to the MEX consensus motif (see the Materials and Methods for calculation). The plot reflects cumulative scores. For example, the 12-fold X:A enrichment at −17 is the enrichment for all motifs $\ln(P) \leq -17$. Shown also is the model describing the MEX motif based on position weight matrices. (*B*) High-scoring MEX motifs on X are predictive for *rex* sites. The distribution of MEX motifs with and without DCC peaks differs on chromosomes X and V. Shown in 1-kb bins are the number (*Y*-axis) and chromosomal positions (*X*-axis) of MEX motifs ($\ln[P] \leq -15$) with (red) or without (blue) DCC peaks for chromosomes X and V. The motifs with asterisks were selected for testing the predictive value of the MEX motif for *rex* sites. All seven tested showed DCC recruitment in vivo (*rex-32* to *rex-38*), demonstrating high predictive value. In contrast, of 38 total high-scoring MEX motifs on all five autosomes, only three are coincident with peaks. (*C*) The X-enriched MEX motif discriminates *rex* sites from *dox* sites. Shown is a histogram indicating the percentage of different classes of sites having 12-bp motifs of different probability matches to the MEX consensus motif. The plot reflects cumulative scores as in *A*. The classes of DNAs compared include 25 *rex* sites, 49 *dox* sites, 1748 SDC-3 X peaks (adjusted to 2 kb each), 457 random X sequences lacking peaks (2 kb), and 321 random autosomal sequences lacking peaks (2 kb). *rex-32* to *rex-38* were not included because the MEX motif was used to predict them.

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

Discovery of rex and dox sites fulfills predictions for a DCC targeting model involving recruitment and spreading

A recruitment and spreading model for DCC binding to X predicts two classes of sites with different DCC recruitment abilities: *rex* sites that recruit the DCC when detached from X and *dox* sites that bind the DCC only when part of an intact X. We tested the model by examining DCC binding in vivo to DNA fragments corresponding to ChIP–chip peaks using our extrachromosomal array assay (Fig. 3). Each ChIP–chip experiment had identified ~1700 peaks of DCC binding on X. Those DCC peaks varied widely in distribution along X, ranging from 900 bp to 97 kb in separation, with a mean distance of 11.5 kb and a median distance of 6.3 kb. The median peak length was 1 kb. Understanding the relative frequencies and positions of recruiting and nonrecruiting DCC peaks is key to evaluating the targeting model.

DCC peaks adjacent to the *rex-1*, *rex-14*, and *rex-17* sites were analyzed first. A large peak similar in size to *rex-1* but 5 kb away failed to recruit the DCC in the array assay, thus defining it as a *dox* site (*dox-7*) (Fig. 3A; Supplemental Table 1). Similarly, large peaks 6 kb to one side of *rex-14* and 2 kb to the other side also failed to recruit, as did the peaks adjacent to *rex-17* (Fig. 3B). These results are consistent with a model of DCC targeting involving DCC recruitment to specific *rex* sites and DCC binding to adjacent sites in a nonautonomous manner.

The generality of these results was assessed by systematically assaying DNA corresponding to all peaks in two 190-kb intervals for recruitment ability (Fig. 3C,D). In the first interval, DNA from only two of 17 peaks recruited, and in the second, DNA from two of 13 peaks recruited. Negative controls for both intervals showed that 2-kb regions lacking peaks ("flat" regions) had insignificant levels of recruitment (Supplemental Table 1). About 140 kb separate the two *rex* sites in the first interval, and 50 kb separate the *rex* sites in the second interval. *rex* and *dox* sites are separated by 4–90 kb in the two intervals. Extrapolating from these results, we estimate that only ~13% of all DCC peaks on X are *rex* sites, implying that long-range interactions facilitate DCC binding to *dox* sites.

Several different methods of approximation suggest that as few as 200–300 *rex* sites may reside on X. The first estimate comes from the two extensively studied 190-kb regions. Four *rex* sites in 380 kb implies ~190 *rex* sites on the 18-Mb X chromosome. The second estimate comes from analyzing all recruitment data, excluding those from the 190-kb regions. In total, DNA from 17 out of 17 (100%) robust peaks (scores ≥3), six out of 17 (35%) large peaks (scores 2–3), and five out of 16 (31%) medium peaks (scores 1–2) recruited (see Supplemental Table 1). A projection from the total number of peaks called in each class estimates 220 *rex* sites, a number consistent with 13% of all X peaks. This calculation does not take small peaks (scores <1) into account, because DNA to 14 out of 14 small peaks assayed failed to recruit. The relatively small number of *rex* sites on X helps explain why an extrachromosomal array containing a few hundred copies of a single *rex* site can compete effectively with the two X chromosomes of hermaphrodites for the majority of DCC in the cell (Fig. 1C; McDonel et al. 2006).

In summary, the DCC binds to discrete, dispersed *cis*-acting sites on X that partition into two categories based on their ability to recruit the DCC when detached from X: *rex* sites, which recruit the DCC in the absence of an intact X, and *dox* sites, which bind the DCC only when part of an intact X. The occurrence of both classes of binding sites fulfills the predictions of a DCC targeting model in which the DCC is recruited to X by special recruitment sites and distributed to other binding sites via a spreading mechanism. *rex* sites are present in small number relative to *dox* sites, and *rex* and *dox* sites are interspersed and separated by long distances.

MEX consensus motif distinguishes rex sites from dox sites

The finding that *rex* sites share a consensus motif that distinguishes the X chromosome from autosomes and is important for DCC recruitment raised the question of whether primary DNA sequence distinguishes *rex* sites from *dox* sites. Comparison of 49 *dox* sites with 25 *rex* sites revealed a strong discrimination between the sites based on the prevalence of MEX motifs (Fig. 2C; Supplemental Fig. 3). Unlike *rex* sites, *dox* sites lack the high-scoring motif variants ($\ln[P] \leq -15$) that are enriched on X by 3.8-fold or greater. Furthermore, only 8% of *dox* sites have a MEX variant of $\ln(P) \leq -11$, compared with 56% of *rex* sites (Fig. 2C; Supplemental Fig. 3). The distribution of MEX motifs in *dox* sites is similar to that in random autosomal or X sequences not bound by the DCC (Fig. 2C; Supplemental Fig. 3), highlighting the importance of *rex* sites in targeting the DCC to X.

Although high-scoring MEX motifs are prevalent in *rex* sites and are important for recruitment, many *rex* sites have motif scores and distributions more similar to *dox* sites (Supplemental Table 1). In those *rex* sites (e.g., *rex-2* and *rex-4*), mutation of MEX variants with $\ln(P)$ −8 or −7, which occur in all *dox* sites, seriously compromises recruitment (Supplemental Table 1; McDonel et al. 2006). Therefore, low-scoring MEX motifs are important for binding in the context of a *rex* site, but a feature other than the MEX motif must designate those *rex* sites for recruitment. Searches for additional motifs in those *rex* sites identified motifs that confer no significant X:A enrichment.

dox sites, unlike rex sites, are found preferentially in expressed genes

The positions of all DCC-binding sites were analyzed to determine whether the binding sites correlate with specific features of the genome and whether *rex* or *dox* sites share the same biases. Approximately half of DCC-binding sites were found in promoters, as noted previously (Ercan et al. 2007), with the majority of peak centers being within 1 kb upstream of the translation start codon (Fig. 4A,B). For SDC-3 sites, 51% are in

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

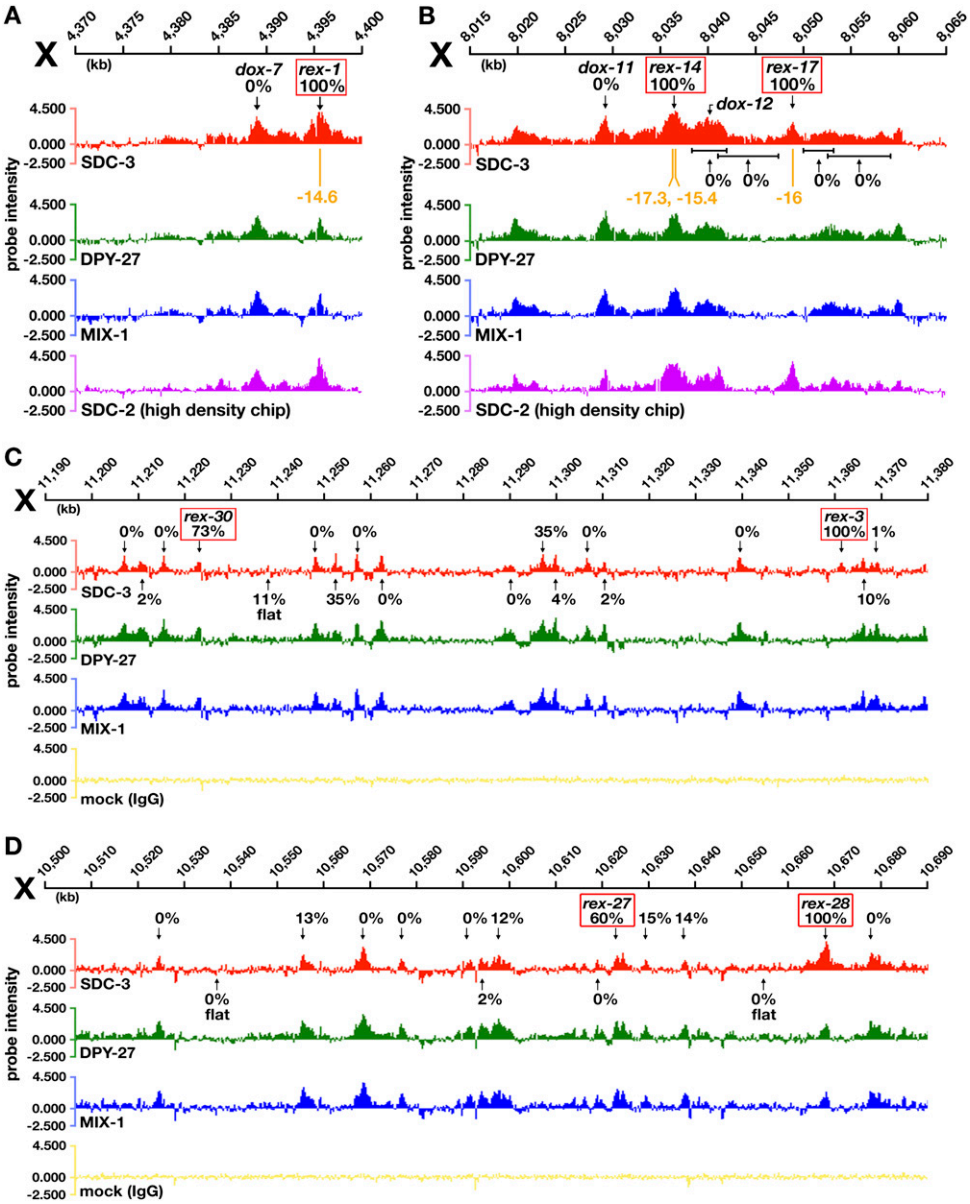

Figure 3. *dox* sites are more prevalent than *rex* sites on X chromosomes. (*A–D*) Representative landscapes from different regions of X showing ChIP–chip profiles in XX embryos for three dosage compensation components and an IgG control using a Nimblegen 380,000 feature isothermal tiling array of predominantly X sequences. The *Y*-axis is the probe intensity ratio of fluorescently labeled DNAs made from chromatin immunoprecipitated by antibodies to different DCC components to fluorescently labeled genomic DNA. The percent of DCC recruitment in extrachromosomal array assays is indicated above the ChIP–chip landscapes for 2-kb DNA fragments coincident with DCC peaks or with regions lacking peaks (flat). See Supplemental Table 1 for recruitment data summary. (*A*,*B*) *dox* sites with excellent DCC peak scores reside 2–6 kb from *rex-1*, *rex-14*, and *rex-17*, consistent with a DCC targeting model of recruitment and disbursement. For the highest scoring MEX motifs within the *rex* sites, the ln(P) (orange) of finding a match to the consensus motif, as calculated by Patser, is given. (*C*,*D*) Few *rex* sites and many *dox* sites reside in 190-kb intervals. These two regions are within the 2-Mb region surveyed with cosmids for *rex* sites. Long distances separate *rex* sites (e.g., 50 kb and 140 kb) as well as *rex* and *dox* sites (up to 90 kb), implying that long-range DNA interactions may facilitate DCC binding to *dox* sites.

promoters (start codon to 2 kb upstream), 24% reside in coding regions, and 25% are intergenic (Fig. 4A,B; Supplemental Table 2). The distribution was similar for DPY-27 peaks (Supplemental Table 2). Few peaks are near the ends of genes, except those associated with small genes (e.g., tRNA or microRNA genes) for which the gene start and stop are in close proximity (Fig. 4A,B, purple line lacks small genes). In fact, 76% of the 274 tRNA genes on X have a DCC peak center within ±1 kb of the gene start, a finding resembling the association of condensin complexes with tRNA genes in yeast mitotic cells (D'Ambrosio et al. 2008).

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

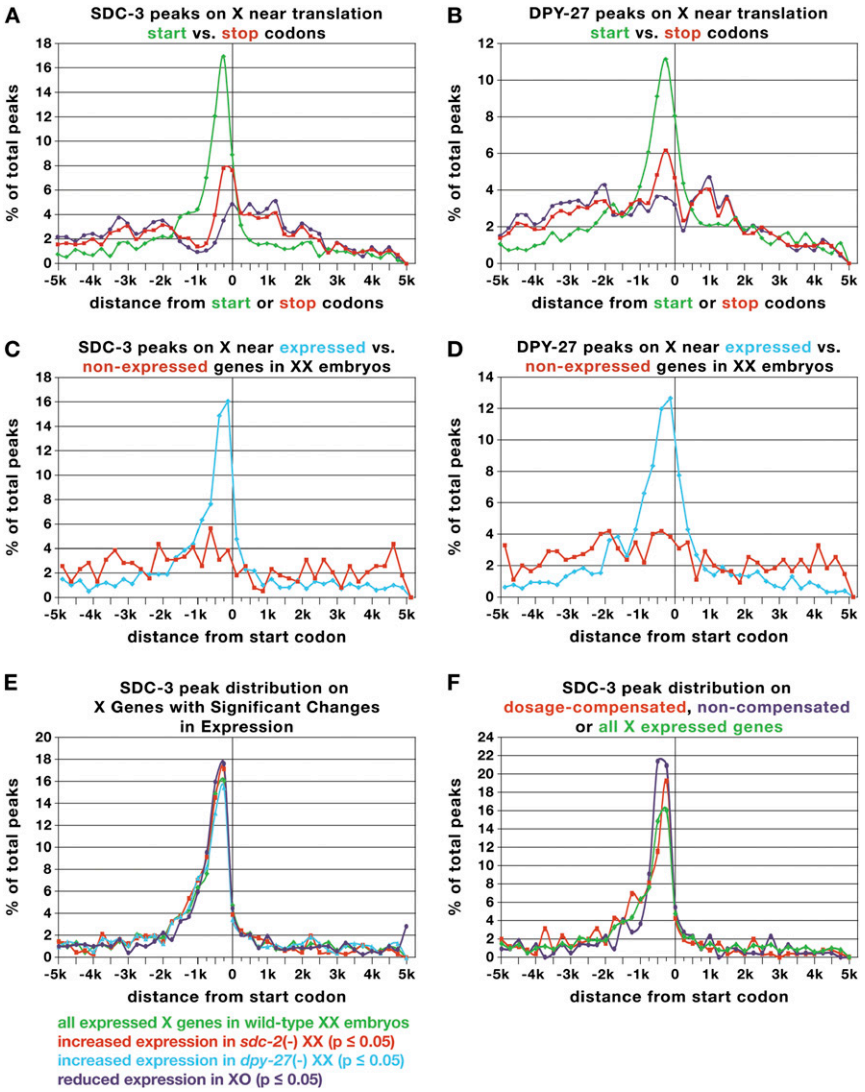

Figure 4. DCC peaks have a similar distribution in all expressed X genes regardless of whether they respond to the DCC. (*A*,*B*) More DCC peaks are near translation start than stop codons of X genes. Shown is the distribution of SDC-3 or DPY-27 peaks relative to the translation start (green line) or stop (red line) codons of genes on X. Peaks were mapped to the nearest translation start or stop codon within ±5-kb interval of the relevant codon. Peaks were then counted in 250-bp bins relative to the start or stop codon. The percentage of peaks in each bin was calculated relative to all peaks within the ±5-kb interval. Removal of genes less than 1 kb in size from the data set resulted in fewer DCC peaks correlated with stop codons (purple line). Thus, the slight enrichment of peaks near stop codons merely reflects the compact *C. elegans* genome size and the close proximity of start and stop codons. (*C*,*D*) About 50% of DCC peaks on X are found preferentially in promoters of expressed versus nonexpressed genes. Shown is the distribution of DCC peaks relative to distances from the translation start codons of expressed (blue line) or nonexpressed (red line) genes. For each expressed or nonexpressed gene, DCC peaks with centers within ±5 kb of the start codon were mapped and counted in 250-bp bins relative to the start codon. SDC-3 peaks were found in 1009 expressed and 390 nonexpressed genes. DPY-27 peaks were found in 1303 expressed and 549 nonexpressed genes. (*E*) Comparison of SDC-3 peak distribution in all expressed X genes (1506 genes), all X genes with statistically increased expression in either *sdc-2* (876 genes) or *dpy-27* (1065 genes) dosage compensation-defective embryos, and all X genes with statistically reduced expression in XO animals (876 genes). Analysis was conducted as in *C*. (*F*) Comparison of SDC-3 peak distribution in dosage-compensated, noncompensated, and all expressed X genes. Analysis was conducted as in *C*.

Not only are the DCC-binding sites biased toward promoters, the DCC sites are more prevalent in expressed than nonexpressed genes (Fig. 4C,D). In fact, DCC peak size is directly correlated with the expression level of the gene; the higher the expression level, the larger the peak (Supplemental Fig. 4). Furthermore, X-linked genes with higher expression levels have a higher proportion of DCC peaks in their promoters (Fig. 5A). However, while 76% of peaks are in genes, only 52% of embryonically expressed genes on X have DCC peaks in their promoters or coding regions (Fig. 5B).

rex sites differ from *dox* sites with regard to gene proximity (Supplemental Table 1). Only 55% (21 out of 38) of *rex* sites (≤ 5 kb) reside in genes (2 kb upstream of translation start codon to stop codon) compared with 88% (43 out of 49) of *dox* sites ($P < 0.001$). *dox* site placement resembles that of most DCC peaks, consistent with *dox* sites being more prevalent than *rex* sites. Thus, to recruit the DCC, a *rex* site need not reside in a promoter nor be associated with an expressed gene. However, the proximity of *dox* sites to expressed genes suggests that a feature of active transcription, perhaps nucleosome-free zones in promoters or the transcription machinery, may facilitate DCC distribution.

Correlation of DCC binding with gene repression

The prevalence of *dox* sites in expressed genes suggested the DCC might act directly on the genes it binds to repress transcript levels, as previously proposed (Ercan et al. 2007). To test this hypothesis, we correlated DCC binding with function. We identified genes responsive to the dosage compensation process by performing genome-wide expression analyses with Affymetrix microarrays on XX, XO, and dosage compensation-deficient XX embryos. Genes undergoing dosage compensation are expected to

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

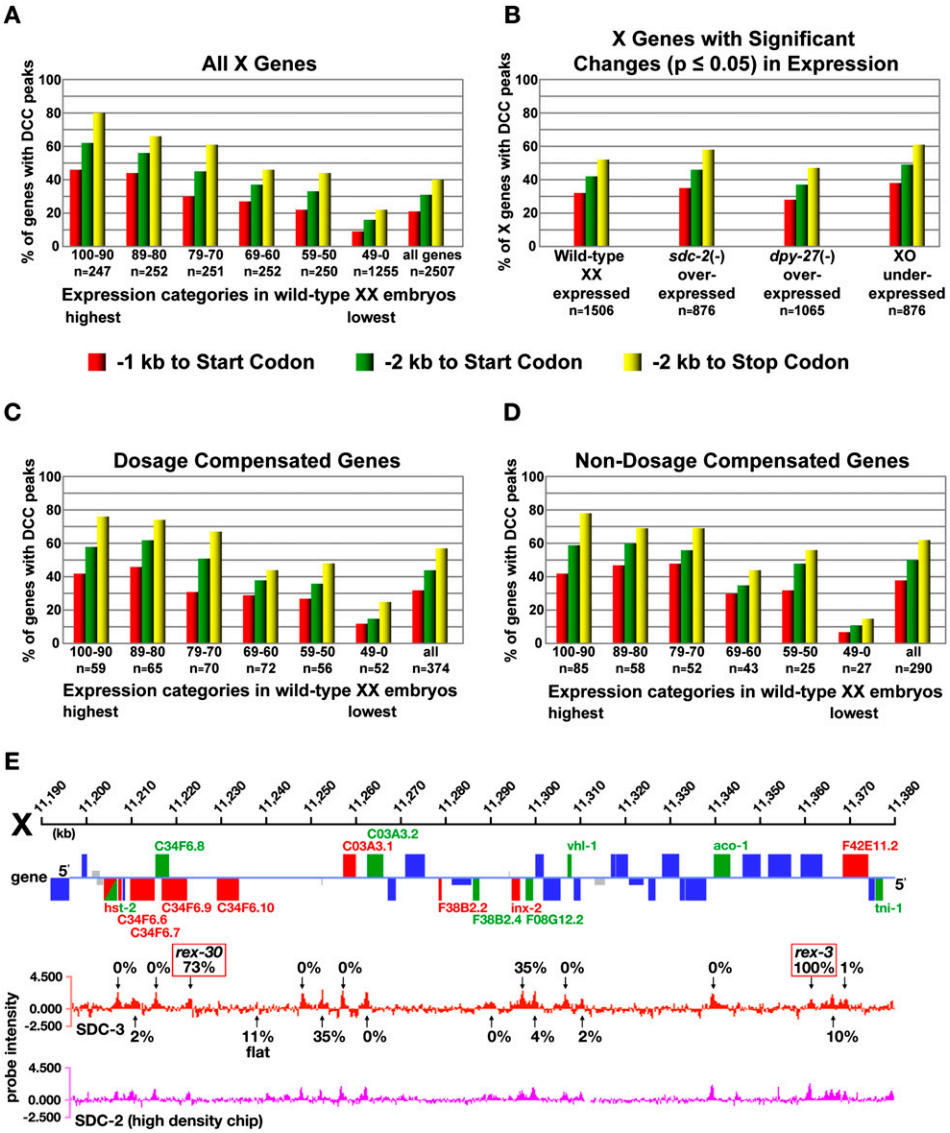

Figure 5. DCC binding and DCC-mediated repression are not directly correlated, implying action at a distance by the DCC. (*A*,*C*,*D*) Histograms show the percentage of all X genes, dosage-compensated, or noncompensated genes with or without DCC peaks plotted relative to the expression level of the gene. Peaks are quantified for three different intervals in the genes, as indicated in the legend. The bins reflect the categories of expression in wild-type XX embryos as assessed in the expression arrays: highest 10% (100–90), next 10% (89–80), etc. For compensated, noncompensated, or all X genes, the higher the expression, the higher the probability of a peak in the three intervals. Dosage-compensated and noncompensated genes have similar percentages of peaks in the three intervals measured. In *A*, "all genes" refers to the sum of X genes in all expression categories, which is the total number of X genes on the microarray. *n* refers to the number of genes in each category. (*B*) Histograms show the percentage of genes in four designated genotypes with DCC peaks in three different intervals of the gene, as indicated by the key. The categories of genes monitored for DCC peaks include all X genes expressed in wild-type animals, all X genes significantly increased in expression ($P \leq 0.05$) in *sdc-2* or *dpy-27* mutant embryos, and all X genes significantly decreased in expression ($P \leq 0.05$) in XO embryos. *n* is the number of genes in each category. The genes overexpressed in *sdc-2* and *dpy-27* mutants overlap by 656 genes. (*E*) Locations of representative dosage-compensated and noncompensated genes on X relative to each other and to *rex* and *dox* sites along the ChIP–chip landscapes. Dosage-compensated and noncompensated genes are interspersed, and peaks are found in promoters of genes from both classes. Orientation of genes is *left* to *right above* the line and *right* to *left below* the line. (Red) Dosage compensated gene; (green) noncompensated gene; (green and red) genes with increased expression in dosage compensation-defective XX embryos and decreased expression in XO embryos; (large blue rectangle) genes not classified; (small blue rectangle) genes not expressed in embryos; (gray) genes not on the expression array.

have elevated transcript levels in XX dosage compensation mutants and not to have decreased transcript levels in XO embryos. Genes insensitive to dosage compensation would have lower transcript levels in XO than XX embryos and not have increased transcript levels in XX dosage compensation mutants.

In the microarray experiments, two different dosage compensation mutants were analyzed: *sdc-2(y93,* RNAi*)*

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

and *dpy-27(y57)*. *sdc-2* triggers recruitment of the DCC to X, and the *sdc-2* treatment (RNAi into a hypomorphic mutant) caused severe dosage compensation defects and lethality of XX embryos. The *dpy-27(y57)* allele only partially disrupts DCC activity. In addition, XO embryos were transformed into phenotypic hermaphrodites by sex determination mutations to avert complications in gene expression caused by sexual dimorphism. Analysis demonstrating that our microarray expression studies were both reliable and sufficiently sensitive to small changes in transcript levels to classify the DCC responsiveness of genes is presented in the Materials and Methods and Supplemental Figures 5 and 6.

We used different regimes, involving increasingly selective criteria for comparing gene expression in wild-type and mutant embryos, to correlate the DCC responsiveness of X-linked genes with the presence or absence of nearby DCC peaks. The same conclusion was reached from all analyses: Unexpectedly, DCC binding to the promoter or coding region of a gene is neither necessary nor sufficient to elicit compensation of the gene.

In the first analysis to correlate DCC binding with gene repression, we asked the general question of whether the mean increases in expression for all active X-linked genes (~1600) in DCC-defective versus wild-type XX embryos were statistically different between genes with without DCC peaks. DCC peaks were scored in the gene interval between 2 kb upstream of the translation start codon and the stop codon. We found that the mean increases in gene expression were not different between genes with and without peaks (Student's *t*-test) (Supplemental Table 3). For *sdc-2*, the mean fold increase was 1.53 ± 0.024 SEM for genes with peaks and 1.46 ± 0.043 SEM for genes without peaks. For *dpy-27*, the mean fold increase was 1.56 ± 0.027 SEM for genes with or without peaks. If a direct correlation were to exist between DCC binding to genes and DCC-mediated repression of gene expression, the DCC-bound genes would have shown a significant increase in expression over genes without bound DCC; however, they did not.

In the second analysis, we looked more specifically at the distribution of DCC peaks in genes with significant changes in gene expression in response to perturbations in dosage compensation: (1) genes with statistically significant increases ($P \leq 0.05$) in expression in DCC-defective XX embryos (either *sdc-2* or *dpy-27* mutant) compared with wild-type XX embryos, (2) genes with significant decreases ($P \leq 0.05$) in expression in XO versus XX embryos, and (3) all expressed X genes in wild-type XX embryos (Fig. 4E). The SDC-3 peak distribution was very similar in the interval ±5 kb of the translation start codon for genes of all four genotypes. More precise quantification (Fig. 5B) showed that a similar percentage of genes in all four classes had DCC peaks in the 1-kb region upstream of the translation start codon (28%–38%), in the 2-kb region upstream of the translation start codon (37%–49%), and in the region from 2 kb upstream of the translation start codon to the stop codon (47%–61%). Together, these results indicate that DCC binding to a gene and DCC-mediated repression are not directly linked. The view is further supported by finding that the mean increase in gene expression for genes with significant increases ($P < 0.05$) in expression in *sdc-2* mutant versus wild-type XX embryos is not different for genes with DCC peaks (1.91 ± 0.027 SEM) and genes without peaks (1.92 ± 0.042 SEM) (Student's *t*-test) (Supplemental Table 3).

In the third analysis, we applied more selective criteria to classify genes by their sensitivity to dosage compensation. For a gene to be considered dosage-compensated, the total transcript level from the gene had to be reduced less than 1.5-fold and $P \leq 0.05$ in XO embryos compared with XX embryos and had to be higher by at least 1.5-fold and $P \leq 0.05$ in dosage compensation-defective XX embryos compared with wild-type XX embryos. The elevated transcript levels had to be observed in both dosage compensation-defective mutants, *sdc-2(y93,* RNAi) and *dpy-27(y57)*. To be considered an escaper from dosage compensation, the total transcript level from the gene had to be lower in XO embryos than in XX embryos by at least 1.5-fold ($P \leq 0.05$) and not increased (<1.5-fold change and $P \leq 0.05$) in either of the two dosage compensation-defective XX mutants. By these criteria, 374 X-linked genes were classified as dosage-compensated and 290 genes as noncompensated (Supplemental Tables 4, 5). Neither class of genes was represented by predominant molecular functions or participation in common biological processes. A third class of genes was found: Of 488 genes with higher transcript levels in both dosage compensation mutants, 114 also had lower transcript levels in XO embryos, indicating that genes can respond to the dosage compensation process but not have equal expression between XX and XO embryos.

A similar percentage of these 374 dosage-compensated and 290 noncompensated genes had DCC peaks in their promoters or coding regions, indicating no direct link between compensation of a gene and DCC binding to the gene (Figs. 4F, 5C,D, 6; Supplemental Figs. 7–10). In total, 57% of dosage-compensated genes, 61% of noncompensated genes, and 51% of all expressed genes on X had peaks in the interval from 2 kb upstream of the translation start codon to the stop codon (Fig. 5A,C,D). The dosage compensation status of representative genes with and without DCC peaks was confirmed by quantitative PCR (Supplemental Table 6). Furthermore, DCC peaks, if present, were of similar sizes in genes of both classes, as is evident from the graphical depictions of probe intensities from ChIP–chip experiments of representative genes (Fig. 6; Supplemental Figs. 7–10). This latter correlation did not require DCC peaks to be called by computer algorithms and confirmed that both compensated and noncompensated genes can have peaks or lack peaks within ± 5 kb of their translation start codons.

We then correlated the wild-type expression levels of these compensated and noncompensated genes with the presence or absence of DCC peaks to determine whether expression level rather than compensation status is a predictor of DCC binding. We quantified the number of compensated and noncompensated genes with a DCC peak center in each of three intervals: start codon to either 1 or 2 kb upstream and 2 kb upstream of start codon

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

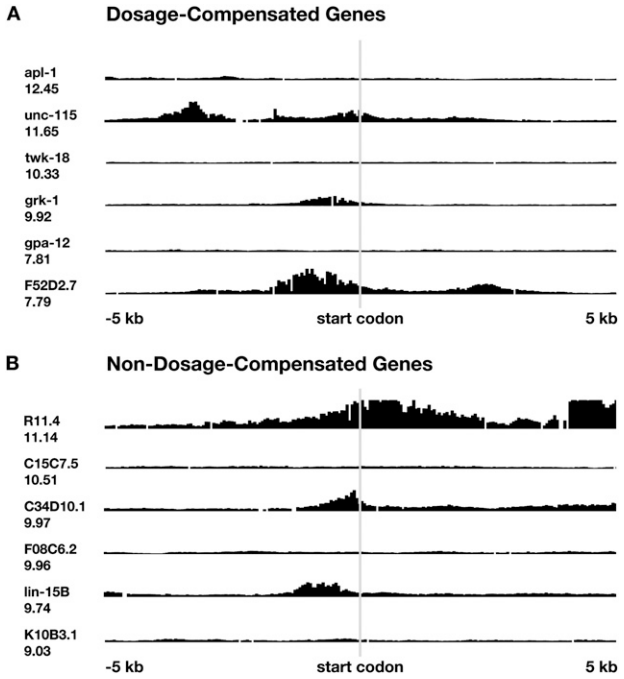

Figure 6. Dosage-compensated and noncompensated genes alike have either no DCC peaks or peaks of variable sizes. Shown are a graphical representations of probe intensities from SDC-3 ChIP–chip experiments along representative dosage-compensated and noncompensated genes in the interval ±5 kb relative to the ATG start codon (gray line). Gene names are shown to the *left*. *Below* the name is the $\log_2$ value for the gene expression level in wild-type XX embryos. Genes of both classes have no peaks or small- to large-sized peaks. Supplemental Figures 7–10 show graphical representations of probe intensities from DPY-27 and SDC-3 ChIP–chip experiment for additional compensated and noncompensated genes.

to stop codon (Fig. 5C,D). The percentage of genes with peaks in the different intervals was plotted in bins reflecting the expression level of the gene in wild-type XX embryos (Fig. 5A,C,D). A similar proportion of compensated and noncompensated genes had DCC peaks in each interval, and for both sets of genes, the percentage of genes with DCC peaks was directly correlated with the expression level of the genes: The higher the expression, the greater the percentage of genes with peaks. The same trend was observed for all X-linked genes regardless of their responsiveness to the DCC. However, the dosage compensation process clearly functions independently of the expression level of the gene: Dosage-compensated and noncompensated genes alike are expressed at high, medium, or low levels in wild-type XX embryos (Fig. 5C,D).

Thus, no correlation was found between the extent of gene dosage compensation, the proximity of a DCC peak to the gene, the DCC peak size, or the expression level of the gene. In fact, for 14% of the dosage-compensated genes, the nearest peak was >10 kb upstream of or downstream from the translation start codon. Additional factors must help define whether a gene responds to the DCC. Although we have not ruled out the possibility that DCC acts directly on some of the genes it binds, numerous examples exist in which either the DCC fails to act locally on the gene it binds or genes become compensated without a DCC-binding site nearby. These results suggest that the DCC can exert its regulatory effect at a distance.

Additional properties of dosage-compensated and noncompensated genes

Two additional conclusions can be reached about the process of dosage compensation from the analysis of dosage-compensated and noncompensated genes. First, although the DCC represses expression of X-linked genes approximately twofold in XX embryos, the extent of repression varies from gene to gene. That is, on average, the increase in gene expression of dosage-compensated genes in dosage compensation mutants was 2.2-fold ± 0.04 SEM, and the median was 1.9-fold, but the extent of increase was frequently greater. The increase ranged from 10-fold to 1.5-fold, the minimum change required. For example, 125 genes were over expressed by twofold to threefold and 39 genes by threefold to 10-fold. These results imply that the DCC does not simply achieve a precise twofold regulation (Supplemental Table 4), although the extreme changes in expression could result from secondary consequences. The same conclusion can be drawn from the noncompensated genes. For these genes, the average expression level was 1.9-fold ± 0.02 SEM lower in XO than XX embryos, and the median reduction in expression was 1.7-fold, but the range varied from 4.3-fold to 1.5-fold, the minimum change required (Supplemental Table 5).

Second, dosage-compensated genes are interspersed with genes that escape dosage compensation, and adjacent genes can belong to either class (Fig. 5E; Supplemental Figs. 11, 12). In addition, the position of a dosage-compensated gene does not influence the location of the next compensated or noncompensated gene, nor does the position of a noncompensated gene influence the location of the next noncompensated gene (Supplemental Fig. 13A–C). Distances between the closest genes of the same or different class are not different from random (Supplemental Fig. 13).

The DCC binds to sites on autosomes and regulates autosomal gene expression

Our analysis of DCC binding and its effect on gene expression was genome-wide. Unexpectedly, all tested components of the DCC (SDC-2, SDC-3, DPY-27, MIX-1) bind to discrete, dispersed sites on autosomes (Fig. 7A,B; Supplemental Fig. 14), and disruption of the DCC affects autosomal gene expression (Fig. 7C). However, only one-fifth the number of binding sites was found on each autosome compared with the X chromosome, and the autosomal peaks corresponded in size to only the small and medium peaks on X. The smaller number and size of DCC-binding sites on autosomes accounts, at least in part, for the inability to detect DCC binding to autosomes by immunofluorescence (Fig. 1C; McDonel et al. 2006). The locations of peaks on autosomes mirrored those on X: A strong bias was found for peaks to be in promoters of

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

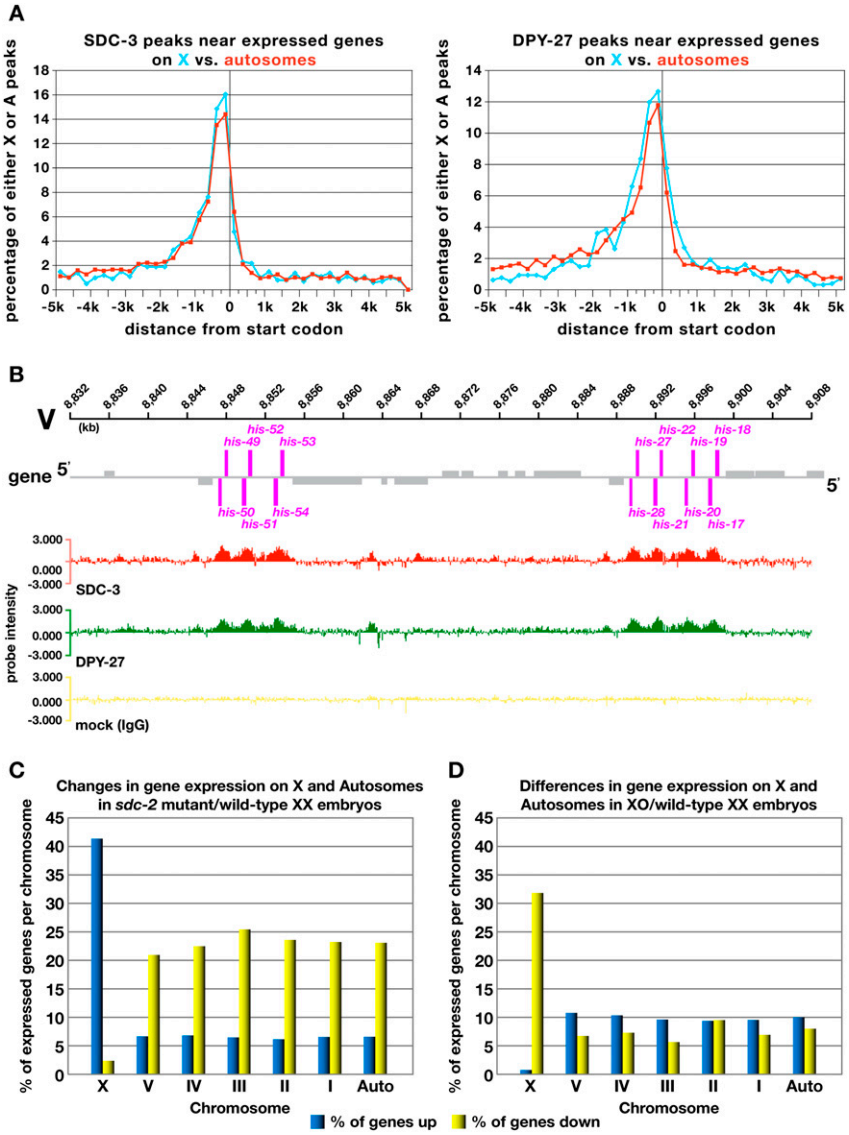

Figure 7. The DCC binds to autosomes and regulates autosomal gene expression. (*A*) Expressed autosomal genes, like expressed X genes, have DCC peaks in their promoters. Shown is the distribution of DCC peaks relative to the translation start codons of expressed genes on X (blue line) or autosomes (red line). For each expressed X or autosomal gene, DCC peaks with centers within ±5 kb of the start codon were mapped and counted in 250-bp bins relative to the start codon. The percentage of peaks in each bin was calculated relative to all peaks within the ±5-kb interval. SDC-3 peaks were found in 1009 expressed X genes and 4767 autosomal genes. DPY-27 peaks were found in 1303 expressed X genes and 4232 expressed autosomal genes. (*B*) Landscapes of SDC-3, DPY-27, and control IgG ChIP–chip experiments in the histone gene clusters on chromosome V (high-density 2.2 million feature tiling arrays). The *Y*-axis is the probe intensity ratio of fluorescently labeled DNAs made from chromatin immunoprecipitated by antibodies to SDC-3 or DPY-27 to fluorescently labeled genomic DNA. Peaks were found in the promoters of seven pairs of divergently transcribed histone genes. Expression of each histone gene was reduced in dosage compensation-defective XX mutants, but in general, genes on autosomes and X chromosomes are similar in that about half of genes affected by DCC disruption lack peaks. Magenta, genes with reduced expression in *sdc-2* XX mutants; gray rectangles, genes not changed in expression. (*C*) In XX dosage compensation-defective embryos, expression of many autosomal genes is reduced, while expression of many X genes is increased. Histogram showing the percentage of X and autosomal genes changed in expression by greater than 1.5-fold and $P \leq 0.05$ in *sdc-2* XX mutant versus wild-type XX embryos. The number of expressed genes per chromosomes in wild-type XX embryos is the following: X, 1506; V, 1903; IV, 1590; III, 1624; II, 1793; I, 1711. (*D*) In XO embryos, expression of many X genes is decreased, as expected for noncompensated genes, while expression of many autosomal genes is either increased or decreased. Histogram showing the percentage of X and autosomal genes changed in expression in XO versus XX embryos.

expressed genes (Fig. 7A,B; Supplemental Table 2; Supplemental Fig. 15A–D). About 56% of SDC-3-binding sites on autosomes are in promoters (translation start codon to 2 kb upstream), 26% are in coding regions, and 18% are intergenic (Supplemental Table 2). The peak distribution was similar for DPY-27 peaks (Supplemental Table 2). Conspicuously, DCC peaks are in the promoters of four prominent classes of genes: all autosomal genes encoding histones (68 of 74 are on autosomes) or proteins of the large and small ribosomal subunits (74 of 76 are on autosomes), 39% of autosomal genes encoding tRNAs, and 95% of autosomal genes encoding embryonically expressed miRNAs. Autosomal peaks in histone genes were among the highest scoring autosomal peaks (Fig. 7B).

Autosomal sites lack the prominent MEX consensus motif of *rex* sites. However, two consensus motifs emerged from the autosomal sites: MEA 1 (Motif Enriched on Autosomes), enriched fourfold on autosomes compared with X, and MEA 2, enriched 68-fold due predominantly to the high incidence of motifs on chromosome V (Supplemental Fig. 16). Both coincide with <10% of peaks on autosomes or the X chromosome. Unlike the X-enriched motif, both autosome-enriched motifs can be bound by the DCC regardless of whether they reside on X chromosomes or autosomes. The prevalence of MEA motifs in autosomal peaks reflects their higher occurrence on autosomes rather than their enhanced ability to attract the DCC when linked to

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

autosomes. The functional role, if any, of these motifs is not yet known.

Depleting the DCC component SDC-2 or partially disrupting DPY-27 function had a widespread effect on autosomal gene expression. In contrast to causing an increase in X-gene expression, disrupting the dosage compensation process caused a decrease in autosomal gene expression (Fig. 7C; Supplemental Figs. 17, 18). Approximately 23% of all expressed autosomal genes in *sdc-2* mutants and 30% in *dpy-27* mutants had their transcript levels reduced by at least 1.5-fold ($P < 0.05$) in contrast to only 2.5% of X genes. Among these genes, the mean reduction in *sdc-2* mutants was 1.9-fold ± 0.01 SEM, the median 1.8-fold, and the maximum 7.8-fold. Only 7% of autosomal genes had increased expression compared with 41% of X genes. Analyses presented in the Materials and Methods demonstrate that the decrease in autosomal gene expression in dosage compensation mutants is not an artifact of the microarray normalization process to compensate the increase in X-gene expression. Thus, the DCC functions genome-wide to control gene expression, but its effects are different on X chromosomes and autosomes.

The DCC-binding sites on autosomes do not correlate directly with the genes affected by mutations in dosage compensation genes. Of autosomal genes with statistically significant changes in gene expression ($P < 0.05$ with no specific degree of change imposed) in *sdc-2* mutants, 45% (1357 out of 3006 genes) with reduced expression, 43% (419 out of 980 genes) with increased expression, and 19% (2160 out of 9410 genes) with unchanged expression have a DCC peak in the gene (2 kb upstream of translational start codon through the stop codon). Two other observations corroborate these findings. In *sdc-2* mutants, the mean decrease in gene expression for genes with statistically reduced expression ($P < 0.05$) is not different for genes with DCC peaks or without peaks: 1.69 ± 0.011 SEM for both classes (Student's *t*-test). Also, for all autosomal genes expressed in wild-type XX embryos, the mean expression value in *sdc-2* mutants for genes with DCC peaks is very similar to the value for genes without peaks, 0.93 ± 0.005 SEM and 0.96 ± 0.007 SEM, respectively. Thus, the effect of the DCC on autosomal gene expression, like its effect on X-gene expression (Fig. 5C,D), is not directly correlated with DCC binding to a gene. Although the primary focus of the DCC is to adjust X-chromosome gene expression between the sexes, the DCC also impacts gene expression broadly throughout the entire genome, implying an involvement in the genome-wide balance of gene expression.

Discussion

We identified *cis*-acting regulatory elements that target the X chromosome for repression by the nematode DCC and discovered fundamental principles by which the DCC recognizes and binds X chromosomes. We also correlated DCC binding with DCC function by probing X-linked genes for their responsiveness to the dosage compensation process. Together, these approaches show the DCC binds to numerous, discrete sites along X to effect chromosome-wide repression by acting, at least in part, over long distances.

X recognition and binding by the DCC

Using a high-resolution, genome-wide approach to identify binding sites of four DCC subunits and a functional approach to assess DCC recruitment ability in vivo, we discovered two distinct classes of binding sites on X. *rex* sites recruit the DCC in an autonomous, DNA sequence-dependent manner using a 12-bp consensus MEX. The motif is critical for DCC binding and occurs in multiple copies, confirming previous studies linking motif clustering with DCC recruitment (McDonel et al. 2006). MEX variants enriched by 3.8-fold or more on X compared with autosomes are highly predictive (95%) for *rex* sites. In contrast, *dox* sites lack the MEX variants highly enriched on X, and unlike *rex* sites, cannot bind the DCC when detached from X. *dox* sites are more prevalent than *rex* sites. These findings fulfill predictions for a targeting model in which the DCC binds to recruitment sites on X and disperses to discrete sites lacking autonomous recruitment ability.

Although X-enriched DNA motifs underlie the mechanism by which the DCC distinguishes the X chromosome from autosomes, the MEX motif cannot be the sole source of X specificity. Some *rex* sites have only MEX variants with scores and distributions similar to those in *dox* sites. A feature other than the MEX motif must designate those sites for recruitment. Nonetheless, mutation of MEX motifs in such *rex* sites disrupts recruitment, showing that lower-scoring MEX variants are important for DCC binding in the context of a *rex* site.

A *rex* site need not reside in a promoter or be associated with an expressed gene to recruit the DCC. In contrast, *dox* sites occur preferentially in expressed genes, and both the level of DCC binding to a site and the probability of a gene having a site correlate directly with the expression level of the gene. The proximity of *dox* sites to expressed genes suggests that an aspect of transcription, perhaps the nucleosome-free zone in a promoter or the transcription machinery itself, facilitates DCC distribution. As of yet, only a highly G-rich 18-bp DNA motif appears to distinguish *dox* sites from random X or autosomal sequences not bound by the DCC. The motif has only a twofold enrichment on X. *rex* and *dox* sites might differ not only by utilizing different principles in DCC binding, but also by serving different functions in dosage compensation.

rex and *dox* sites are interspersed and separated by considerable distances (2–90 kb), implying that long-range communication facilitates DCC loading onto X. Communication might take place between *rex* sites and between *rex* and *dox* sites. Long-range interactions could occur by a variety of mechanisms. In a tracking model based on the enhancer-mediated activation of bacteriophage T4 late genes (Kolesky et al. 2002), *rex* sites could act as DCC loading platforms from which the DCC traverses along the chromatin fiber to *dox* site destinations. More likely, DCC binding could induce changes in chromosome structure, such as DNA looping, to bring

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

rex sites into proximity with other *rex* sites and with *dox* sites. As an example, repression of the entire bacteriophage λ genome is achieved through cooperativity in bacteriophage λ repressor binding to strong and weak sites brought together by DNA looping (Dodd et al. 2004). Furthermore, the sequential activation of genes in the β-globin locus requires DNA looping to place the *cis*-acting LCR (locus control region) next to genes it activates, while displacing 40–60 kb of intervening inactive genes (Palstra et al. 2008). LCR-stimulated genes then relocate to an active chromatin hub in the nucleus. The similarity of the DCC to condensin supports models invoking changes in chromosome structure.

Comparison of X-chromosome recognition and DCC binding in flies and nematodes

Although the DCCs of flies and nematodes are evolutionarily unrelated and regulate X chromosomes in opposite ways, similar principles appear to govern the X-chromosome targeting and binding of the two complexes. In flies, dosage compensation is achieved by the MSL (male-specific-lethal) complex, which binds to the single X chromosome of males to increase transcript levels (Lucchesi et al. 2005). As in worms, chromatin entry sites recruit the MSL complex in a DNA sequence-dependent manner using an X-enriched motif (Alekseyenko et al. 2008; Straub et al. 2008). For both flies and worms, a second mode of binding correlates with high levels of gene expression and appears sequence-independent (Kind and Akhtar 2007; Larschan et al. 2007; Kind et al. 2008), suggesting that features common to transcribed genes facilitate binding. Sequence-independent MSL binding favors the 3′ ends of active genes (Alekseyenko et al. 2006; Gilfillan et al. 2006; Legube et al. 2006), while nematode DCC binding often favors 5′ ends of active genes.

Functioning of the nematode DCC on X

To explore the mechanism of dosage compensation, we correlated DCC-binding sites with genes responsive to DCC function. We assessed the responses of genes to the DCC through a series of genome-wide expression studies using microarrays to compare gene expression in XX, XO, and DCC-defective XX embryos. We then used a range of criteria, from least selective to most selective, to judge the effect of the DCC on gene expression. Several conclusions about the dosage compensation process emerged from our analysis.

First, although DCC binding to promoters is more prevalent for highly expressed genes, a gene need not be highly expressed to be dosage-compensated. Genes of all expression levels respond to the dosage compensation process. Thus, the DCC functions independently of the expression level of the gene it controls, and the criteria for DCC binding differ from the criteria for DCC function.

Second, although the process of dosage compensation is intended to compensate for the imbalance in X-chromosome dose between XO and XX animals by equalizing expression of X-linked genes, the nematode DCC does not compensate all X genes, nor does it achieve a precise twofold repression of all genes it compensates. The dosage compensation processes in mammals and flies also fail to compensate all genes (Carrel and Willard 2005; Legube et al. 2006; Johnston et al. 2008). On average, a nematode gene that escapes compensation is expressed at half the level in XO embryos as in XX embryos, and expression of a compensated gene is increased approximately twofold in dosage compensation-defective XX mutants. However, the range in mutants is wide, consistently from twofold to threefold and up to fivefold. A reasonable conclusion is that survival of the species does not require a precise twofold regulation of all X-linked genes, and the DCC is not able to achieve this level of precision for all genes it controls.

Third, the sites of DCC binding and the targets of DCC action are often separated by long distances. Furthermore, DCC binding to a gene is neither necessary nor sufficient to invoke compensation of that gene. These conclusions were reached through a series of analyses. Using the least selective criteria to judge responsiveness of genes to the DCC, we found that the mean change in expression level for all active genes with peaks or without peaks was the same in *sdc-2* XX mutants and in *dpy-27* XX mutants. Using the most selective criteria to classify genes, we found that 43% of dosage-compensated genes lack DCC peaks, and 61% of noncompensated genes have DCC peaks. Rather than finding a strong correlation between genes with DCC peaks and genes affected by dosage compensation, we found a strong correlation between highly expressed genes and genes with peaks, those representing *dox* sites. While these cumulative results do not eliminate the possibility that the DCC might act locally to repress some genes it binds, they indicate the DCC can act over long range to repress many genes.

Our experiments evaluated the model proposed by Ercan et al. (2007) that the DCC binds to 5′ ends of genes to repress transcription of the genes it binds. Their model was inferred from the preferential location of DCC-binding sites in promoters of genes, but gene expression studies were not conducted to test the model. Unlike the prediction from their model, we found no direct correlation between DCC binding and DCC-mediated repression. The DCC-binding sites identified in our studies are virtually identical to those identified by Ercan et al. (2007) with regard to the genes we classify as compensated and noncompensated (Supplemental Figs. 7–9). We conclude that their model does not accurately describe the mechanism of DCC action.

Models for X repression by the DCC

We propose that the DCC achieves gene repression over long distance by reconfiguring the architecture of the chromosome. For example, the DCC might act in a manner similar to the *Drosophila* Polycomb group (PcG) proteins, which repress genes in the bithorax complex locus (BX-C, ~340 kb) through distant *cis*-acting Polycomb response elements (PREs) (Mateos-Langerak and Cavalli 2008). When the BX-C locus is repressed, it adopts

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

a condensed, looped structure in which the PcG-bound PRE elements and gene promoters all interact in a higher-order structure (Lanzuolo et al. 2007). As a gene becomes activated, the PRE–promoter interaction is lost and the gene loses contact with the repressive structure. Such a mechanism could account for how the DCC, a condensin-like complex with potential chromosome compaction activity, could repress a gene far removed from its primary binding site.

How a gene might escape dosage compensation can be explained by different models. A noncompensated gene could be excluded from a higher-order structure required for repression, as in the BX-C locus. Alternatively, the noncompensated gene could be insulated from repression through a mechanism resembling the regulation of imprinted genes: Gene activation by a distant enhancer is blocked by an intervening insulator site (Wan and Bartolomei 2008). Adapted for DCC regulation, two adjacent genes could respond differently to a distantly bound DCC if an intervening insulator blocks the repressive effect of the DCC. For cases in which a gene escapes compensation despite having a promoter-bound DCC, the bound DCC might be inactive, might not act locally, or might have its action blocked by a strong activator or its local chromatin environment. A relevant example is the bookmarking phenomenon in which a gene is blocked from condensin-mediated condensation, and hence repression, during mitosis by a TBP-bound phosphatase, which inactivates the promoter-bound condensin (Xing et al. 2008).

Control of autosomal gene expression by the DCC

Although the primary function of the DCC is to repress X-chromosome gene expression in hermaphrodites, and DCC disruption causes lethality only in hermaphrodites, the DCC impacts autosomal genes. The DCC binds to discrete, dispersed sites on autosomes, and DCC disruption alters expression of autosomal genes. Approximately one-quarter of autosomal genes have reduced expression in dosage compensation mutants, while about half of X-linked genes show the opposite effect, increased expression. Thus, the DCC exerts a genome-wide influence on gene activity. DCC-binding sites on autosomes, like those on X, reside preferentially in promoters of expressed genes, and the autosomal sites correlate infrequently with the autosomal genes affected by dosage compensation mutations. This result suggests that the DCC controls autosomal genes, like X genes, over long range. DCC binding to autosomal sites has not yet been linked to a primary DNA sequence, or specific recruitment sites, leaving open the question of whether binding to autosomes involves recruitment and spreading or a different mechanism. For example, interchromosomal interactions between X chromosomes and autosomes could facilitate binding to autosomes in conjunction with binding to X (Lomvardas et al. 2006; Osborne et al. 2007).

The unexpected effect of the DCC on autosomal genes suggests another model for the mechanism of DCC action. In wild-type XX animals, the DCC could reduce expression of genes on X by partially repelling a rate-limiting, genome-wide activator from X, thus making the overall level of X-bound activator, and hence X-linked gene expression, similar in XX and XO animals. In so doing, the DCC would prevent the amount of activator drawn to the two X chromosomes of XX embryos from exceeding that drawn to the single X chromosome of XO embryos, thereby ensuring that autosomes of both sexes have the same level of activator. This model explains our observation that depleting the DCC from X by disrupting SDC-2 decreases autosomal gene expression and increases X-gene expression (Fig. 7C). In *sdc-2* XX mutants, more activator would bind to X and activate those genes, leaving less activator available for autosomal genes. Hence, the difference in autosomal gene expression between XO embryos and *sdc-2* XX embryos (Fig. 7D) would be caused by their difference in X-chromosome dose. This model does not attribute functional significance to the autosomal DCC binding.

Conclusions

Prominent themes have emerged for the mechanisms underlying X-chromosome-wide gene repression by a condensin-like regulatory complex. Binding of the complex to X chromosomes involves two different mechanisms: sequence-dependent recruitment to autonomous binding sites not requiring gene expression, and nonautonomous binding to sites located preferentially in expressed genes. The complex likely acts over long range to exert at least part of its regulatory effect, perhaps by inducing changes in chromosome structure. Not only does the complex play a prominent role in controlling the sex chromosome, it also affects expression throughout the genome.

Materials and methods

ChIP–chip platform and experiments

ChIP–chip experiments were performed using two different Nimblegen chip platforms. One set of experiments, which included duplicate SDC-3 and MIX-1 IPs and a single DPY-27 IP, used a 385K feature isothermal (tm = 76°C) tiling array of predominantly X sequences with no repeat masking, using WormBase release WS158. Probes occurred every 50 bp along X and varied in length range between 45 and 73 bp. The second set of experiments, which included duplicate SDC-3, DPY-27, and SDC-2 IPs, used either of three 2.2-million-feature high-density (HD2), isothermal tiling arrays of all X and autosomal sequences based on WormBase release WS170 or WS180. For SDC-2, an anti-Flag monoclonal antibody was used to IP a Flag-tagged version of SDC-2, which was expressed from an unintegrated transgene in an *sdc-2*(*y74*null) mutant.

ChIP–chip data analysis

Ratio GFF and annotation (WS158, WS170, WS180) files were supplied by Nimblegen. Data were analyzed using NimbleScan software and viewed with SignalMap software. Peaks were called on *C. elegans* 385K whole-genome tiling arrays using the default settings, except the sliding window was 1 kb, the "min probes" and "when all probes in peak" cutoffs were eight probes. For

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

a peak to be called, at least eight of 20 probes had to be above the peak height cutoff threshold (P) within a 1-kb window. Peaks were called on *C. elegans* HD2 whole-genome tiling arrays using the default settings, except the sliding window was 1 kb, the "min probes" and "when all probes in peak" cutoffs were 11 probes. For a peak to be called, at least 11 of 25 probes had to be above the peak height cutoff threshold (P) within a 1-kb window. Peaks with an FDR ≤0.2 for HD2 arrays and all called peaks for 385K arrays were mapped to genomic features. A more detailed description of the peak calling algorithm is at http://www.nimblegen.com/products/lit/nimblescan2dot4usersguide.pdf.

Motif identification

The wconsensus version 5c program (Hertz and Stormo 1999) was used with default parameters except for the following: Number of iterations, *N*, was set to the number of DNA sequences for small samples (<50) and to one-half of the number of samples when more samples were used; the standard deviation for identifying information peaks was varied from 2 to 6; both strands were considered; and 10 matrices were printed.

X:A fold enrichment calculation for MEX motif

The Patser program (version 3e) (Hertz and Stormo 1999) was used to calculate the natural log of the probability (ln[P]) of finding a match to the MEX motif at all positions along each chromosome. For each threshold value, the ln(P) values less than this value were summed for X and for autosomes. The autosomal value was divided by the total number of autosomal base pairs to find the number of motifs per base pair. Similarly, the number of motifs per base pair was calculated for X. The final X to A ratio was calculated by dividing the motifs per base pair for X by the motifs per base pairs for the autosomes.

Gene expression arrays and data analysis

For gene expression analysis using Affymetrix *C. elegans* Genome Microarrays (whole-genome GeneChip array) or qRT–PCR, RNA was prepared as in the Supplemental Material. For microarray analysis, the nematode strains and number of experiments was as follows: wild-type XX embryos (six microarrays), XO *her-1(hv1y101); xol-1(y9) sdc-2(y74) unc-9(e101)* embryos (eight microarrays), *dpy-27(y57)* XX embryos (three microarrays), and *sdc-2(y93,* RNAi*)* XX embryos (three microarrays). Arrays were normalized by RMA (Robust Multichip Average) using Affymetrix Expression Consol software (Irizarry et al. 2003). Quantile normalization is further described at http://www.bea.ki.se/staff/reimers/Web.Pages/Affymetrix.Normalization.htm. Data were analyzed by ArrayStar. Statistical analysis was performed using a moderated *t*-test with FDR multiple testing correction (Benjamini-Hochberg method). For each experimental strain, a separate list of genes was complied that met different sets of criteria explained in the results. Comparison of changes in gene expression between different strains was performed by comparing the lists of genes.

To determine the number of genes expressed in embryos on each chromosome, data from the six wild-type XX embryo arrays were normalized by MAS5 using Expression Consol software. A gene was considered expressed if it was called "present" on three or more arrays probing RNA from wild-type XX embryos and nonexpressed if called "absent" on two or fewer arrays. For genes with multiple probe sets, any gene with conflicting present/absent calls (<1%) was discarded form the data set. *C. elegans* Genome Microarrays were hybridized at The Stanford Genome Technology Center.

Evaluation of microarray normalization

Two analyses showed that the decrease in autosomal gene expression in dosage compensation-defective mutants reflects autosome gene regulation by the DCC rather than the array normalization process compensating for the increase in X-gene expression.

The first analysis evaluated the RMA normalization. Comparing the sum of all probe intensities of RMA normalized chips for wild-type XX embryos to the sum for *sdc-2* mutant XX embryos (12,810,292 vs. 12,554,506, respectively) showed a 2% change, indicating successful normalization. Summed X probe intensities for wild-type XX embryos compared with *sdc-2* mutant XX embryos (1,263,736 vs. 1,740,042, respectively) showed a 40% increase. However, the overall decrease in summed autosomal probe intensities *for sdc-2* mutant versus wild-type XX embryos was only 6.5%: 11,566,556 versus 10,814,465, respectively. Thus, the normalization procedure could in principle cause all autosomal genes to be decreased by 6.5% in *sdc-2* mutants. However, in our *sdc-2* data set, the criterion that the expression level of a gene be statistically different ($P \leq 0.05$) from the expression level in wild-type XX embryos results in genes whose expression is decreased by >10%. The decrease was even greater when we used a more selective criterion (expression decrease of ≥1.5-fold and $P \leq 0.05$) for defining autosomal genes with reduced expression in *sdc-2* mutants. Using that criterion, 20%–23% of genes on each autosome had decreased expression. Therefore, the decrease in autosomal gene expression in *sdc-2* mutants is significant and not a consequence of the normalization procedure.

Furthermore, if the RMA normalization procedure were problematic, we would have seen a compensating increase in autosomal gene expression in XO embryos, given that 30% of X genes in XO embryos showed decreased expression (at least 1.5-fold and $P \leq 0.05$) compared with genes in XX embryos. We did not. Instead, the number of autosomal genes with increased expression (10%) was roughly the same as the number with decreased expression (8%).

The second analysis used the dChip normalization procedure to evaluate expression changes in XO or dosage compensation-defective XX embryos compared with wild-type XX embryos. dChip software finds an invariant set of probes and uses it for normalizing chips (Li and Wong 2001). Using dChip normalization, we found a similar number of autosomal genes with statistically reduced gene expression in dosage compensation mutants (either *P* value ≤0.05 with no fold change cutoff or $P \leq 0.05$ and a fold decrease of ≥1.5) as found using the RMA normalization (Supplemental Fig. 18A,B).

Evaluation of microarray reliability

The reliability of our microarray expression assays for defining genes responsive to dosage compensation is demonstrated by the overlap in genes identified as overexpressed in different DCC-defective XX mutants. Of 1583 X probe sets representing actively expressed genes in wild-type XX embryos, 766 showed statistically increased fluorescence ($P < 0.05$) when probed with RNA from *sdc-2*(*y93*, RNAi) XX mutants and 925 when probed with RNA from *dpy-27(y57)* XX mutants; the overlap is 631. A chi-square test (χ^2) rejects the hypothesis that the data sets are independent ($P < 0.001$), and the Cohen's κ coefficient (κ) of 0.46 (95% confidence interval [CI]: 0.41–0.51) indicates a positive association. The κ increased to 0.57 (95% CI: 0.52–0.62), indicating greater agreement, when more selective criteria (≥1.5-fold increase and $P < 0.05$) were used to judge a probe set as increased in intensity when probed with RNA from DCC-defective mutants.

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

In contrast, probe sets with decreased fluorescence when probed with XO embryo RNA compared with XX embryo RNA represent noncompensated genes and should not overlap with probe sets showing increased fluorescence when probed with dosage compensation XX mutant RNA. In fact, they do not. Of 1583 X probe sets, 583 had statistically decreased fluorescence ($P < 0.05$) with XO embryo RNA, and only 117 overlap with the 631 probe sets with increased fluorescence with dosage compensation mutant RNA. The χ^2 test rejects the independence of the data sets ($P < 0.001$), but a κ of -0.31 (95% CI: -0.36 to -0.26) indicates a negative association, and hence, discordance between genes with increased expression in XX DCC-defective embryos and decreased expression in XO embryos.

Further evidence for the reliability of the microarrays is the concordance in autosomal genes changed in expression in DCC-defective XX embryos. Of 12,012 autosomal probe sets representing actively expressed genes in wild-type XX embryos, 3231 had reduced fluorescence ($P < 0.05$) when probed with *sdc-2*(*y93*, RNAi) RNA and 3914 when probed with *dpy-27(y57)* RNA; the overlap was 1970. The χ^2 test ($P < 0.001$) coupled with κ of 0.36 (95% CI: 0.35–0.38) shows that the data sets are positively associated. Similarly, of 12,012 probe sets, 1433 had increased fluorescence ($P < 0.05$) when probed with *sdc-2*(*y93*, RNAi) RNA and 2724 when probed with *dpy-27(y57)* RNA; the overlap was 894. The data sets are positively associated (χ^2, $P < 0.001$), and κ of 0.32 (95% CI: 0.31–0.34).

Nematode culture

Embryos were used for microarray analysis and obtained from bleached gravid adults of several genotypes. Wild-type XX and XO *her-1(hv1y101); xol-1(y9); sdc-2(y74); unc-9(e101)* animals were grown on egg plates (Chu et al. 2006). *sdc-2*$^-$ XX animals were prepared by growing *sdc-2(y93)* XX animals on Ahringer feeding library bacteria (Kamath et al. 2003) bearing an *sdc-2* plasmid. The wild-type control for the *sdc-2* RNAi was N2 worms fed Ahringer library feeding bacteria bearing a plasmid lacking an insert. Feeding bacteria was prepared by seeding 1L LB cultures with a single colony, growing overnight at 37°C, inducing for 2 h with 5 mM IPTG, pelleting, and resuspending in 1 vol (w/v) of LB with 20% glycerol.

Accession number

ChIP–chip and gene expression data in this manuscript can be accessed in the NCBI GEO public repository with the accession number GSE14640.

Topics included in the Supplemental Material

The following information is provided in the Supplemental Material: protocol for RNA preparation, protocol for ChIP from embryo lysates, and references.

Acknowledgments

We thank S. MacArthur for initial computational analysis of ChIP–chip experiments; N. Wang, T. Speed, and M. Ralston for help with statistical analysis; S. Uzawa for microscopy; J. Wilhelmy for expression microarray processing; J. Gunther for figure design; and M. Botchan, T. Cline, and A. Wood for critical comments on the manuscript. R.R.P. was supported by training grant T32GM07127, J.J. was supported by Nederlandse Wetenschaps Organisatie, and A.H.M. was supported by Swiss National Science Foundation Grant PBGEA-119304. Research was supported by NIH grant R01-GM30702 to B.J.M. M.B.E. and B.J.M. are Investigators of the Howard Hughes Medical Institute.

References

Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. 2006. High-resolution ChIP–chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. *Genes & Dev.* **20:** 848–857.

Alekseyenko, A.A., Peng, S., Larschan, E., Gorchakov, A.A., Lee, O.K., Kharchenko, P., McGrath, S.D., Wang, C.I., Mardis, E.R., Park, P.J., et al. 2008. A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome. *Cell* **134:** 599–609.

Bhalla, N., Biggins, S., and Murray, A.W. 2002. Mutation of *YCS4*, a budding yeast condensin subunit, affects mitotic and nonmitotic chromosome behavior. *Mol. Biol. Cell* **13:** 632–645.

Carrel, L. and Willard, H.F. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* **434:** 400–404.

Chu, D.S., Liu, H., Nix, P., Wu, T.F., Ralston, E.J., Yates 3rd, J.R., and Meyer, B.J. 2006. Sperm chromatin proteomics identifies evolutionarily conserved fertility factors. *Nature* **443:** 101–105.

Chuang, P.T., Albertson, D.G., and Meyer, B.J. 1994. DPY-27: A chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through association with the X chromosome. *Cell* **79:** 459–474.

Cobbe, N., Savvidou, E., and Heck, M.M. 2006. Diverse mitotic and interphase functions of condensins in *Drosophila*. *Genetics* **172:** 991–1008.

Csankovszki, G., McDonel, P., and Meyer, B.J. 2004. Recruitment and spreading of the *C. elegans* dosage compensation complex along X chromosomes. *Science* **303:** 1182–1185.

D'Ambrosio, C., Schmidt, C.K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., and Uhlmann, F. 2008. Identification of *cis*-acting sites for condensin loading onto budding yeast chromosomes. *Genes & Dev.* **22:** 2215–2227.

Davis, T.L. and Meyer, B.J. 1997. SDC-3 coordinates the assembly of a dosage compensation complex on the nematode X chromosome. *Development* **124:** 1019–1031.

Dawes, H.E., Berlin, D.S., Lapidus, D.M., Nusbaum, C., Davis, T.L., and Meyer, B.J. 1999. Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science* **284:** 1800–1804.

Dej, K.J., Ahn, C., and Orr-Weaver, T.L. 2004. Mutations in the *Drosophila* condensin subunit dCAP-G: Defining the role of condensin for chromosome condensation in mitosis and gene expression in interphase. *Genetics* **168:** 895–906.

Dodd, I.B., Shearwin, K.E., Perkins, A.J., Burr, T., Hochschild, A., and Egan, J.B. 2004. Cooperativity in long-range gene regulation by the λ CI repressor. *Genes & Dev.* **18:** 344–354.

Ercan, S., Giresi, P.G., Whittle, C.M., Zhang, X., Green, R.D., and Lieb, J.D. 2007. X chromosome repression by localization of the *C. elegans* dosage compensation machinery to sites of transcription initiation. *Nat. Genet.* **39:** 403–408.

Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. 2006. Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes & Dev.* **20:** 858–870.

Hagstrom, K.A., Holmes, V.F., Cozzarelli, N.R., and Meyer, B.J. 2002. *C. elegans* condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid

Downloaded from genesdev.cshlp.org on April 20, 2010 - Published by Cold Spring Harbor Laboratory Press

segregation during mitosis and meiosis. *Genes & Dev.* **16:** 729–742.

Hertz, G.Z. and Stormo, G.D. 1999. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15:** 563–577.

Hirano, T. 2005. Condensins: Organizing and segregating the genome. *Curr. Biol.* **15:** R265–R275. doi: 10.1016/j.cub.2005.03.037.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4:** 249–264.

Johnston, C.M., Lovell, F.L., Leongamornlert, D.A., Stranger, B.E., Dermitzakis, E.T., and Ross, M.T. 2008. Large-scale population study of human cell lines indicates that dosage compensation is virtually complete. *PLoS Genet.* **4:** e9. doi: 10.1371/journal.pgen.0040009.

Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., et al. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421:** 231–237.

Kind, J. and Akhtar, A. 2007. Cotranscriptional recruitment of the dosage compensation complex to X-linked target genes. *Genes & Dev.* **21:** 2030–2040.

Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. 2008. Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* **133:** 813–828.

Kolesky, S.E., Ouhammouch, M., and Geiduschek, E.P. 2002. The mechanism of transcriptional activation by the topologically DNA-linked sliding clamp of bacteriophage T4. *J. Mol. Biol.* **321:** 767–784.

Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F., and Orlando, V. 2007. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell Biol.* **9:** 1167–1174.

Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. 2007. MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell* **28:** 121–133.

Legube, G., McWeeney, S.K., Lercher, M.J., and Akhtar, A. 2006. X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in *Drosophila*. *Genes & Dev.* **20:** 871–883.

Li, C. and Wong, W.H. 2001. Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc. Natl. Acad. Sci.* **98:** 31–36.

Lieb, J.D., Capowski, E.E., Meneely, P., and Meyer, B.J. 1996. DPY-26, a link between dosage compensation and meiotic chromosome segregation in the nematode. *Science* **274:** 1732–1736.

Lieb, J.D., Albrecht, M.R., Chuang, P.T., and Meyer, B.J. 1998. MIX-1: An essential component of the *C. elegans* mitotic machinery executes X chromosome dosage compensation. *Cell* **92:** 265–277.

Lomvardas, S., Barmea, G., Pisapia, D.J., Medelsohn, M., and Kirkland, J. 2006. Interchromosomal interaction and olfactory receptor choice. *Cell* **126:** 403–413.

Lucchesi, J.C., Kelly, W.G., and Panning, B. 2005. Chromatin remodeling in dosage compensation. *Annu. Rev. Genet.* **39:** 615–651.

Mateos-Langerak, J. and Cavalli, G. 2008. Polycomb group proteins and long-range gene regulation. *Adv. Genet.* **61:** 45–66.

McDonel, P., Jans, J., Peterson, B.K., and Meyer, B.J. 2006. Clustered DNA motifs mark X chromosomes for repression by a dosage compensation complex. *Nature* **444:** 614–618.

Meyer, B.J. 2005. X-Chromosome dosage compensation. In *WormBook* (ed. The *C. elegans* Research Community), *WormBook*. doi: 10.1895/wormbook.1.8.1. http://www.wormbook.org.

Morris, J.R., Geyer, P.K., and Wu, C.T. 1999. Core promoter elements can regulate transcription on a separate chromosome in trans. *Genes & Dev.* **13:** 253–258.

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. 2007. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol.* **5:** e192. doi: 10.1371/journal.pbio.005092.

Palstra, R.J., de Laat, W., and Grosveld, F. 2008. β-Globin regulation and long-range interactions. *Adv. Genet.* **61:** 107–142.

Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. 2008. The chromosomal high-affinity binding sites for the *Drosophila* dosage compensation complex. *PLoS Genet.* **4:** e1000302. doi: 10.1371/journal.pgen.10000302.

Tsai, C.J., Mets, D.G., Albrecht, M.R., Nix, P., Chan, A., and Meyer, B.J. 2008. Meiotic crossover number and distribution are regulated by a dosage compensation protein that resembles a condensin subunit. *Genes & Dev.* **22:** 194–211.

Wan, L.B. and Bartolomei, M.S. 2008. Regulation of imprinting in clusters: Noncoding RNAs versus insulators. *Adv. Genet.* **61:** 207–223.

Xing, H., Vanderford, N.L., and Sarge, K.D. 2008. The TBP–PP2A mitotic complex bookmarks genes by preventing condensin action. *Nat. Cell Biol.* **10:** 1318–1323.