Published in final edited form as:
Nat Genet. 2007 March ; 39(3): 403–408. doi:10.1038/ng1983.

X-chromosome repression by localization of the *C. elegans* dosage compensation machinery to sites of transcription initiation

Sevinc Ercan[1], **Paul G. Giresi**[1], **Christina M. Whittle**[1], **Xinmin Zhang**[2], **Roland D. Green**[2], and **Jason D. Lieb**[1,*]

[1]Department of Biology and Carolina Center for the Genome Sciences CB #3280, 202 Fordham Hall University of North Carolina at Chapel Hill Chapel Hill, NC 27599-3280

[2]Nimblegen Systems, Inc. 1 Science Court Madison, WI 53711

Introductory paragraph

Among organisms with chromosome-based mechanisms of sex determination, failure to equalize expression of X-linked genes between the sexes is typically lethal. In *C. elegans*, XX hermaphrodites halve transcription from each X chromosome to match the output of XO males[1]. Here, we mapped the binding location of the condensin homolog DPY-27 and the zinc finger protein SDC-3, two components of the *C. elegans* dosage compensation complex (DCC) [2,3]. Strong foci of DCC binding were observed on X, around which broader regions of localization were centered. Binding foci, but not adjacent regions of localization, were distinguished by clusters of a stereotypic 10-bp DNA sequence, suggesting a recruitment-and-spreading mechanism for X recognition. In contrast to the Drosophila DCC, the *C. elegans* DCC was bound preferentially upstream of genes, suggesting modulation of transcriptional initiation and polymerase-coupled spreading. A mechanism for tuning DCC activity at specific loci was indicated by stronger DCC binding upstream of genes with high transcriptional activity. These data provide a basis for understanding how proteins involved in higher-order chromosome dynamics can regulate transcription at individual loci.

Main Text

To compensate for differences in X-linked gene dosage between XY males and XX females, mammals inactivate most genes on one of the two female X chromosomes[4]. In contrast, *C. elegans* XX hermaphrodites dosage compensate by reducing transcription from each X chromosome by a factor of two to match the expression of XO males[1]. This mechanism is remarkable in that the subtle two-fold downregulation is imposed upon X-linked genes expressed over a large dynamic range[5,6]. The dosage compensation complex (DCC) required for *C. elegans* X-repression is composed of proteins encoded by the genes sdc-1, sdc-2, sdc-3, dpy-21, dpy-26, dpy-27, dpy-28, dpy-30 and *mix-1*[7]. DPY-26, DPY-27, DPY-28 and MIX-1 are homologous to members of the condensin complex, which is required for chromosome condensation and segregation in organisms ranging from bacteria to humans[8]. While the

*Correspondence should be addressed to: Jason D. Lieb jlieb@bio.unc.edu (919) 843-3228.

AUTHOR CONTRIBUTIONS

This study was designed by SE and JDL. SE conducted the experiments. SE, PGG, CMW, and JDL conducted the data analysis. XZ and RDG designed, manufactured, and hybridized the DNA microarrays. SE and JDL wrote the paper.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

molecular parallels to mitotic chromosome condensation broadly suggest a mechanism for reducing X-linked transcription[9,10], the features of X that control DCC binding have proven more difficult to investigate. Four loci on X (*rex-1-4*) sufficient for DCC recruitment were identified recently by using confocal microscopy to detect DCC binding to multiple-copy extrachromosomal transgenic DNA[11,12]. Here, we take an alternative approach that comprehensively maps the distribution of DCC binding along the natural X chromosomes of wild-type animals, identifies DNA sequence elements enriched at sites of DCC recruitment on chromosomes *in vivo*, and determines the relationship between DCC binding and genomic organization at high resolution.We performed ChIP (Chromatin ImmunoPrecipitation) of the zinc-finger protein SDC-3 and the condensin subunit homolog DPY-27 and hybridized the enriched DNA fragments to microarrays consisting of 50-bp oligonucleotide probes tiled across the genome with 36-bp spacing ("ChIP-chip", Methods). To confirm that our reagents and detection methods could identify DCC binding sites, we first examined *rex-1,* a locus previously shown to recruit the DCC to extrachromosomal DNA *in vivo*[11]. Enrichment at *rex-1* was consistently strong for SDC-3 (Figure 1A) and DPY-27 (Figure 1B), while the identical procedure performed in the absence of an antibody revealed no enrichment (Figure 1C). As a second validation, we examined the *her-1* locus on Chromosome V, the lone autosomal region known to be bound by the DCC[13]. DCC binding to *her-1* is required for the repression of *her-1* transcription in XX animals[13]. SDC-3 exhibited strong binding at the promoter and second intron of *her-1* (Figure 1D-E), consistent with published data showing binding to these regions[13,14]. Finally, the previously identified *rex-1* site[12] and two new sites discovered from our microarray data were confirmed by locus-specific PCR (Figure 1F). Based on these results, we interpret the ChIP enrichment measured by our high-resolution DNA microarray hybridizations to reflect the binding location and relative local abundance of the *C. elegans* dosage compensation machinery. To our knowledge, this is the first successful application of ChIP-chip methodology to this important model system.

The DCC exhibits a striking preference for binding to the X chromosome relative to autosomes (Figure 2; $p < 10^{-16}$ for SDC-3 and DPY-27; $p = 0.88$ for no antibody control), consistent with earlier results obtained by *in situ* microscopy[2,3]. For both DPY-27 and SDC-3, we noted a "baseline" level of enrichment on X, punctuated by peaks of greater enrichment (Figure 2A). This led us to speculate that the binding and distribution of the DCC on X involves at least two distinct mechanisms, one that leads to the appearance of distinct peaks, and one that leads to a more uniform distribution. Furthermore, the "baseline" level of binding was higher for DPY-27 than SDC-3 (Figure 2A and B). This, coupled with the observation that SDC-2 can bind to X independently of other DCC components[15] and that SDC-2 and SDC-3 bind to X early in the DCC assembly process[3], suggests that a protein complex containing SDC-2 and SDC-3 is critical for recruitment of the DCC to specific locations, while a more uniform distribution of the DCC, or possibly a subcomplex containing DPY-27 but not SDC-3, spreads from those points. The baseline binding and putative spreading of the DCC helps explain previous observations in which large regions of X were unable to recruit the DCC as extrachromosomal DNA, but were dosage compensated and bound by the DCC in their natural chromosomal context[11].

This recruitment-and-spreading hypothesis predicts that at recruitment sites, peaks of DPY-27 and SDC-3 would be largely coincident. We identified 1193 binding peaks for SDC-3 and 1499 for DPY-27 on X, compared to a total of 9 sites of enrichment for each protein on the five autosomes (Methods, Supplementary Table)[16]. To identify potential recruitment sites, we designated a class of exceptionally high-amplitude peaks as "foci" of DCC binding (Figure 3A-C, Supplementary Figure 1A-B). Fifty-five foci were identified for SDC-3, and 44 for DPY-27 (Supplementary Table). Whereas fewer than two foci would be expected to overlap under the null expectation, 35 of the SDC-3 and DPY-27 foci were coincident ($p < 10^{-100}$).

Three of the four previously discovered *rex* sites[12] were classified as foci, confirming that these sites act as recruitment centers on natural chromosomes.

We further hypothesized that the baseline mode of binding (represented in part by the smaller peaks) would emanate from putative recruitment sites, which themselves are distributed randomly along X. Indeed, foci of SDC-3 or DPY-27 binding did not cluster ($p = 0.15$ and 0.14 respectively; Methods; Supplementary Figure 1C-D), but smaller peaks were strongly clustered (SDC-3 $p < 10^{-25}$; DPY-27 $p < 10^{-4}$). Furthermore, small peaks were clustered around foci. For example, 45% of SDC-3 peaks were within 50 kb of a focus (22% expected by chance; $p = 2\times10^{-6}$; Supplementary Figure 1E-F). At foci, SDC-3 binding was high but diminished rapidly to a low baseline, while the signal from DPY-27 maintained a higher baseline along the X (Figure 2A-B, Figure 3D, Supplementary Note). These data further support the hypothesis that complexes at foci are of distinct composition relative to complexes that spread from the foci.

We reasoned that if foci were DCC recruitment centers, they may be specified by DNA sequence motifs. When used independently as input for motif discovery algorithms, DNA sequence defined by SDC-3 and DPY-27 foci yielded nearly identical 10-bp motifs (Figure 4A, Methods). In contrast, no motifs were derived using sequence from the smaller peaks as input, consistent with the hypothesis that foci are DNA-sequence based recruitment centers. The motif we identify using the natural DCC binding data expands one of the two motifs derived from extrachromosomal DNA assays[12] by 4 bp and helps explain focused tests of function on the *rex* sites (Figure 4B, Supplementary Figure 2). For example, at *rex-1* a 33-bp fragment sufficient for recruitment includes the motif, and mutations that abolished recruitment activity[12] specifically alter the motif identified here (Figure 4B). In addition, the second intron of *her-1* contains the motif (Figure 1D). Finally, one of the strongest newly discovered sites of SDC-3 and DPY-27 binding along the X is centered on two closely spaced motifs upstream of the male-specific *xol-1* gene[17] (Figure 4C). In XX hermaphrodites, *xol-1* is repressed both transcriptionally and post-transcriptionally[18,19], but the potential contribution of the DCC in directly repressing *xol-1* was previously unknown.

It is clear, however, that the identified motif does not specify DCC binding completely. At the most strict motif definition (Matrixscan $p \leq 10^{-6}$, Methods), 47% of X-linked motif occurrences were within an SDC-3 binding focus, and 42% of foci on X contained a motif (Supplementary Figure 3A-B). Furthermore, the motif occurs throughout the genome and is only moderately X-enriched (Supplementary Figure 3C-D). Consistent with previous observations[12] and the occurrence of two closely spaced motifs upstream of *xol-1*, clustering of motif occurrences appears to be important for DCC recruitment along X (Supplementary Figure 3E).

We sought to determine the relationship between DCC binding and gene location (Figure 5A). We found that DCC binding was strongest at a stereotypic distance upstream of the translation start and interpreted the signal to emanate from the transcription start site. Consistent with this interpretation, DCC binding peaks occurred preferentially in intergenic regions, specifically regions annotated to be promoters (Supplementary Figure 4A-B). The *C. elegans* genome contains numerous operons, in which multiple genes are encoded by a single transcript. If the signal indeed emanates from transcription start sites, signal at operons is predicted to occur upstream of the first translation initiation site, but not upstream of internal translation start sites. Indeed, a DCC binding peak is observed only upstream of genes that are at the start of an operon, supporting our interpretation (Figure 5B). Localization to transcription start sites suggests a mechanism of repression that acts at transcriptional initiation, in contrast to the Drosophila DCC (Supplementary Note). Furthermore, yeast studies have shown that the homologous condensins and the related cohesin complexes are pushed along genes by polymerases[20-22], suggesting a spreading mechanism coupled directly to polymerase

movement. This could explain the "baseline" X-association of the DCC that occurs in the absence of DNA sequence signals, and the preferential occurrence of foci downstream of convergently transcribed genes (Supplementary Figure 4B).

The DCC uniformly down-regulates genes transcribed at diverse frequencies, but how this is achieved is unknown. We asked whether genes associated with abundant transcripts required more DCC binding to achieve down-regulation. Indeed, a positive correlation between DCC binding and the transcript level of downstream genes during embryogenesis was observed (Figure 5C, Supplementary Figure 4C). Furthermore, DCC binding to promoters is likely to be dynamic because embryonic DCC binding was not correlated to adult RNA levels (Figure 5D). This "tuning" of the degree of DCC binding for different levels of gene expression presents a more refined depiction of DCC action than previous models, which suggested a more uniform global condensation or "overwinding" of the entire X[10]. A possible mechanism for directing the DCC to sites of active transcription is suggested by DPY-30, which is required for the localization of the DCC to X[23]. DPY-30 is a homolog of *S. cerevisiae* Saf19, a member of the histone methyltransferase Set1 complex responsible for H3K4 methylation, a mark associated with active transcription[24].

Our comprehensive DCC binding data reveal that a non-uniform distribution of a condensin-like protein complex applies a constant, subtle, two-fold level of repression across an entire chromosome. The emerging model involves cooperation of DNA sequence and chromatin signals that target the DCC to the promoters of individual genes, possibly affecting transcriptional initiation. Recruitment, which is tuned to the level of expression, may be followed by a short-range polymerase-coupled spreading mechanism. Detailed tests of the hypotheses suggested here are required to determine the relative contribution of DNA sequence, transcriptional activity, and chromatin modifications to DCC localization, whether DCC composition varies across the X, and the precise mechanism of DCC-mediated repression.

METHODS

Antibodies

Recombinant proteins containing amino acids 1067-1340 of SDC-3 and 1-409 of DPY-27 were cloned and expressed with Novagen's Pet30-EkLIC system. Rabbit polyclonal antibodies were produced at Covance Immunology Services. DNA encoding the SDC-3 and DPY-27 epitopes were cloned into pGEX-5X-2 vector (Amersham Biosciences) and used to affinity purify antibodies.

Chromatin Immunoprecipitation

For extract preparations, N2 adults were grown in standard S-liquid media and embryos were obtained by bleaching[25]. Embryos were fixed with 2% formaldehyde for 30 minutes at room temperature and washed with 100 mM Tris pH 7.5 once, M9 buffer twice and 10 ml FA buffer (50 mM HEPES/KOH pH 7.5, 1 mM EDTA, 1% Triton X-100, 0.1 % sodium deoxycholate; 150 mM NaCl) with protease inhibitors (Calbiochem) and frozen at −80°C. 500 μl of packed embryos were thawed and resuspended in 2 ml of FA buffer and dounce homogenized on ice with 30 strokes. To shear chromatin (size range 200-800 bp), samples were sonicated with a Branson sonifier at 30% amplitude for 7 cycles of 12 pulses (0.9 sec on, 0.1 sec off), with cooling in dry ice/ethanol bath for 2 seconds between cycles. Cellular debris was removed by microfuge centrifugation at 13,000 g for 15 minutes at 4°C. Immunoprecipitation reactions contained approximately 3 mg of total protein with 3-5 μg of purified antibodies in 500 μl total volume, with 1% sarkosyl. Prior to addition of the antibody, 10% of the material was taken as input. Immunocomplexes were collected with Protein A-sepharose beads (Amersham Biosciences) and washed with 1 ml of the following buffers : FA buffer 2 × 5 min, FA-1M

NaCl (FA buffer with 1 M NaCl) 5 min, FA-500 mM NaCl 10 min, TEL buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0) 10 min., TE pH 8.0 2 × 5 min. Complexes were eluted in 1% SDS in TE with 250 mM NaCl at 65°C 30 minutes. Samples and inputs were treated with Proteinase K for 1 hour and crosslinks were reversed at 65°C overnight. DNA was purified with Qiagen PCR purification columns and amplified by LM-PCR (Supplementary Methods). Locus-specific ChIP PCRs were performed by diluting amplified material 50-fold and querying with locus-specific primers (Supplementary Methods).

Microarrays and Data Extraction

The *C. elegans* genome tiling arrays were designed and constructed by NimbleGen System Inc. using maskless array synthesis technology, with approximately 650,000 oligonucleotides per array[26]. Two arrays covered the *C. elegans* genome at 86-bp resolution as described in the text. We performed three biologically independent experiments for SDC-3, DPY-27 and No Antibody control ChIPs. Amplified samples were labeled and hybridized to high-resolution microarrays by the NimbleGen Service Laboratory as described previously[27]. Two ChIPs (ChIP 1 and 3) were labeled with Cy5 and the reference materials (Input DNA) were labeled with Cy3, and for one (ChIP 2), the dyes were swapped. Raw intensity data was extracted from scanned images using NimbleScan 2.3 extraction software (NimbleGen Systems Inc.). The peak finding algorithm ChIPOTle[16], was used to identify regions of SDC-3 or DPY-27 binding, with a sliding window of 500 bp and 86 bp steps. For details, including data normalization and visualization, see Supplementary Methods.

Motif Discovery

MDscan[28] and BioProspector[29], were used to identify motifs with the sequence composition of the *C. elegans* genome used as a background model. Motif widths of up to 15 bp were explored. The 10-bp motif described was most predictive of DCC binding. We used the MatrixScan module from BioProspector to scan the genome for the presence of the 10-bp sequence motif at three cutoffs ($p \leq 10^{-4}$, $p \leq 10^{-5}$, $p \leq 10^{-6}$).

Transcription Analysis

RNA levels in embryos and adults were obtained from the literature[6]. The transcription level of genes during embryogenesis was determined by averaging the last three time points from a published study[5]. The average DCC binding z score for each gene was calculated by taking the average value of probes that fall between 500 bp upstream and 200 bp downstream of the translation start site. The list of genes and the translation start coordinates were obtained from UCSC WormBase Genes from the sangerGene table (http://genome.ucsc.edu/). For genes with splice variants, the coordinates for the longest splice variant were used. WormMart (http://www.wormbase.org) was used to determine if a gene was located within or at the start of an operon.

Accession numbers

Raw data are available from GEO (Gene Expression Omnibus), accession number GSE6739. Note: Supplementary information is available on the Nature Genetics website.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Andrew B. Nobel (Professor of Statistics, University of North Carolina) for providing statistical guidance, especially regarding the clustering of DCC binding peaks. This work was supported by the Carolina Center for Genome Sciences and by a grant awarded to J.D.L. from The V Foundation for Cancer Research.

REFERENCES

1. Meyer BJ, Casson LP. Caenorhabditis elegans compensates for the difference in X chromosome dosage between the sexes by regulating transcript levels. Cell 1986;47:871–81. [PubMed: 3779843]
2. Chuang PT, Albertson DG, Meyer BJ. DPY-27:a chromosome condensation protein homolog that regulates C. elegans dosage compensation through association with the X chromosome. Cell 1994;79:459–74. [PubMed: 7954812]
3. Davis TL, Meyer BJ. SDC-3 coordinates the assembly of a dosage compensation complex on the nematode X chromosome. Development 1997;124:1019–31. [PubMed: 9056777]
4. Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B. Xist RNA and the mechanism of X chromosome inactivation. Annu Rev Genet 2002;36:233–78. [PubMed: 12429693]
5. Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP. Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development 2003;130:889–900. [PubMed: 12538516]
6. Jiang M, et al. Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans. Proc Natl Acad Sci U S A 2001;98:218–23. [PubMed: 11134517]
7. Meyer, BJ.; Community, T.C.e.R.. X-Chromosome dosage compensation.. WormBook. 2005. WormBook doi/10.1895/wormbook.1.8.1, http://www.wormbook.org
8. Hirano T. Condensins: organizing and segregating the genome. Curr Biol 2005;15:R265–75. [PubMed: 15823530]
9. Hagstrom KA, Holmes VF, Cozzarelli NR, Meyer BJ. C. elegans condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis. Genes Dev 2002;16:729–42. [PubMed: 11914278]
10. Lieb JD, Albrecht MR, Chuang PT, Meyer BJ. MIX-1: an essential component of the C. elegans mitotic machinery executes X chromosome dosage compensation. Cell 1998;92:265–77. [PubMed: 9458050]
11. Csankovszki G, McDonel P, Meyer BJ. Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes. Science 2004;303:1182–5. [PubMed: 14976312]
12. McDonel P, Jans J, Peterson BK, Meyer BJ. Clustered DNA motifs mark X chromosomes for repression by a dosage compensation complex. Nature 2006;444:614–8. [PubMed: 17122774]
13. Chu DS, et al. A molecular link between gene-specific and chromosome-wide transcriptional repression. Genes Dev 2002;16:796–805. [PubMed: 11937488]
14. Li W, Streit A, Robertson B, Wood WB. Evidence for multiple promoter elements orchestrating male-specific regulation of the her-1 gene in Caenorhabditis elegans. Genetics 1999;152:237–48. [PubMed: 10224257]
15. Dawes HE, et al. Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. Science 1999;284:1800–4. [PubMed: 10364546]
16. Buck MJ, Nobel AB, Lieb JD. ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. Genome Biol 2005;6:R97. [PubMed: 16277752]
17. Miller LM, Plenefisch JD, Casson LP, Meyer BJ. xol-1: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in C. elegans. Cell 1988;55:167–83. [PubMed: 3167975]
18. Hodgkin J, Zellan JD, Albertson DG. Identification of a candidate primary sex determination locus, fox-1, on the X chromosome of Caenorhabditis elegans. Development 1994;120:3681–9. [PubMed: 7821230]
19. Nicoll M, Akerib CC, Meyer BJ. X-chromosome-counting mechanisms that determine nematode sex. Nature 1997;388:200–4. [PubMed: 9217163]

20. Wang BD, Eyre D, Basrai M, Lichten M, Strunnikov A. Condensin binding at distinct and specific chromosomal sites in the Saccharomyces cerevisiae genome. Mol Cell Biol 2005;25:7216–25. [PubMed: 16055730]

21. Lengronne A, et al. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature 2004;430:573–8. [PubMed: 15229615]

22. Glynn EF, et al. Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. PLoS Biol 2004;2:E259. [PubMed: 15309048]

23. Hsu DR, Chuang PT, Meyer BJ. DPY-30, a nuclear protein essential early in embryogenesis for Caenorhabditis elegans dosage compensation. Development 1995;121:3323–34. [PubMed: 7588066]

24. Nagy PL, Griesenbeck J, Kornberg RD, Cleary ML. A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. Proc Natl Acad Sci U S A 2002;99:90–4. [PubMed: 11752412]

25. Lewis JA, Fleming JT. Basic culture methods. Methods Cell Biol 1995;48:3–29. [PubMed: 8531730]

26. Singh-Gasson S, et al. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. Nat Biotechnol 1999;17:974–8. [PubMed: 10504697]

27. Selzer RR, et al. Analysis of chromosome breakpoints in neuroblastoma at subkilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chromosomes Cancer 2005;44:305–19. [PubMed: 16075461]

28. Liu XS, Brutlag DL, Liu JS. An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. Nat Biotechnol 2002;20:835–9. [PubMed: 12101404]

29. Liu X, Brutlag DL, Liu JS. BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. Pac Symp Biocomput 2001:127–38. [PubMed: 11262934]

30. Workman CT, et al. enoLOGOS: a versatile web tool for energy normalized sequence logos. Nucleic Acids Res 2005;33:W389–92. [PubMed: 15980495]

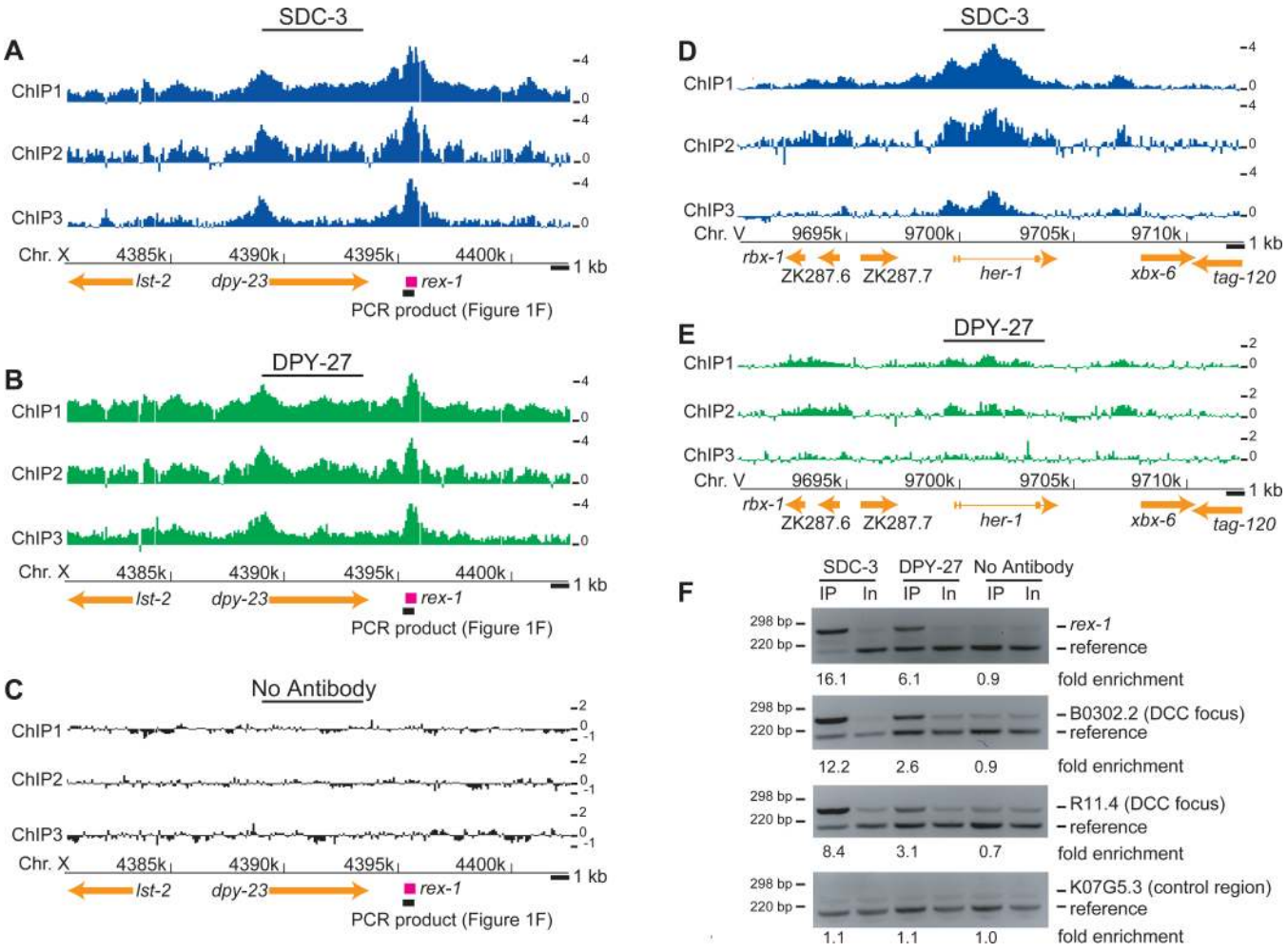

Figure 1. High-resolution localization of the DCC authenticates binding at putative recruitment elements and reveals new targets

(A) Raw $\log_2$ ratios (ChIP/Input) from each of three SDC-3 ChIP replicates plotted along 25 kb of the X chromosome, with gene annotations below. Arrows indicate direction of transcription. A previously identified recruitment element on X (*rex-1*)[12] downstream of *dpy-23* is indicated in pink. **(B)** Same as (a), but a DPY-27 ChIP **(C)** Same as (a), but a no antibody control ChIP. **(D)** SDC-3 localization at the *her-1* locus. Introns are indicated by thin lines connecting exons. **(E)** Same as (d), but for DPY-27. Lower levels of DPY-27 localization are observed at *her-1*, consistent with the lower requirement indicated by genetic data. **(F)** Sequence-specific ChIP analysis of *rex-1*[12], two newly identified DCC binding foci on X (near B0302.2 and R11.4) and a reference region on chromosome I (K07G5.3).

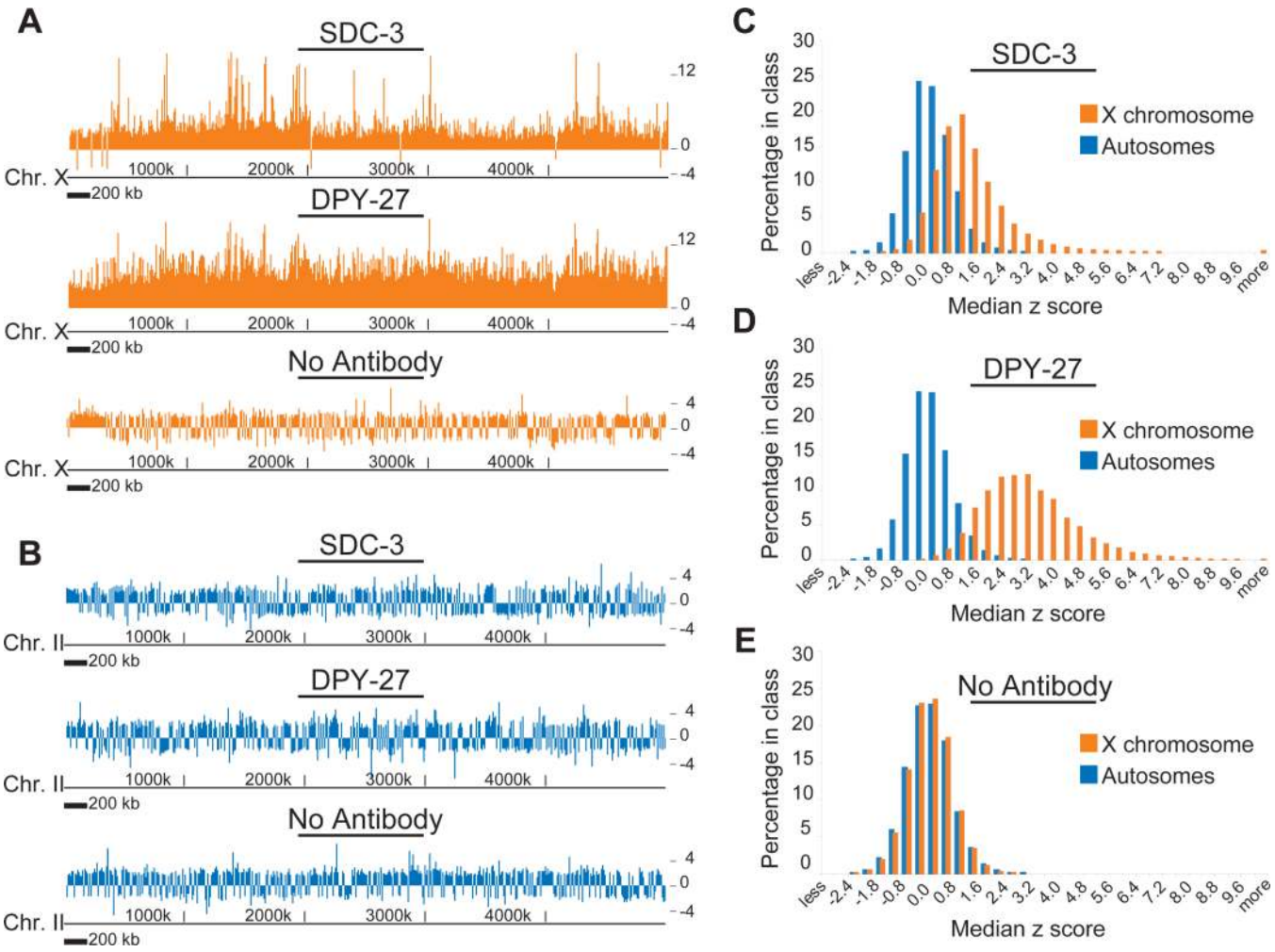

Figure 2. SDC-3 and DPY-27 bind specifically to the X with distinct modes of distribution

(A) Median z scores of enrichment (calculated from ChIP/Input signals) following ChIP using antibodies to SDC-3, DPY-27 and a no antibody control, plotted along a 5 MB region from the left end of X. **(B)** Same as (a), plotted along a 5 MB region of chromosome II. **(C-E)** Histograms of the distribution of z-score ChIP enrichment values of individual probes for autosomes and X chromosomes. **(C)** SDC-3 ChIP, **(D)** DPY-27 ChIP and **(E)** a no antibody control ChIP.

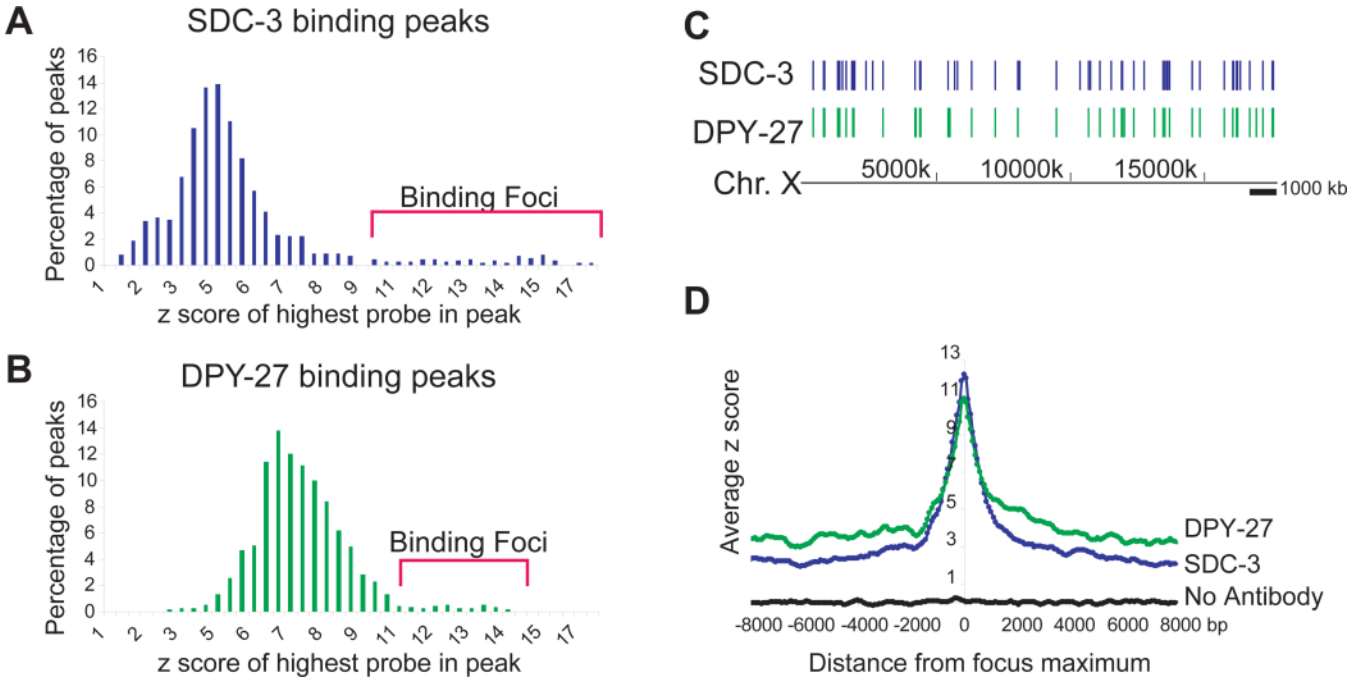

Figure 3. A distinct class of DCC localization foci along X

(A) The distribution of maximum amplitudes (z-score) for SDC-3 ChIP peaks. Peaks classified as "Foci" are indicated by brackets (> 2 standard deviations from mean of the distribution). **(B)** Same as (a) for DPY-27. **(C)** Distribution of SDC-3 foci and DPY-27 foci along X. **(D)** Peaks classified as foci were aligned by their maxima, and the rate of signal decline was measured by sliding a window (width of 3 probes, step size of 1 probe) away from the maximum in both directions.

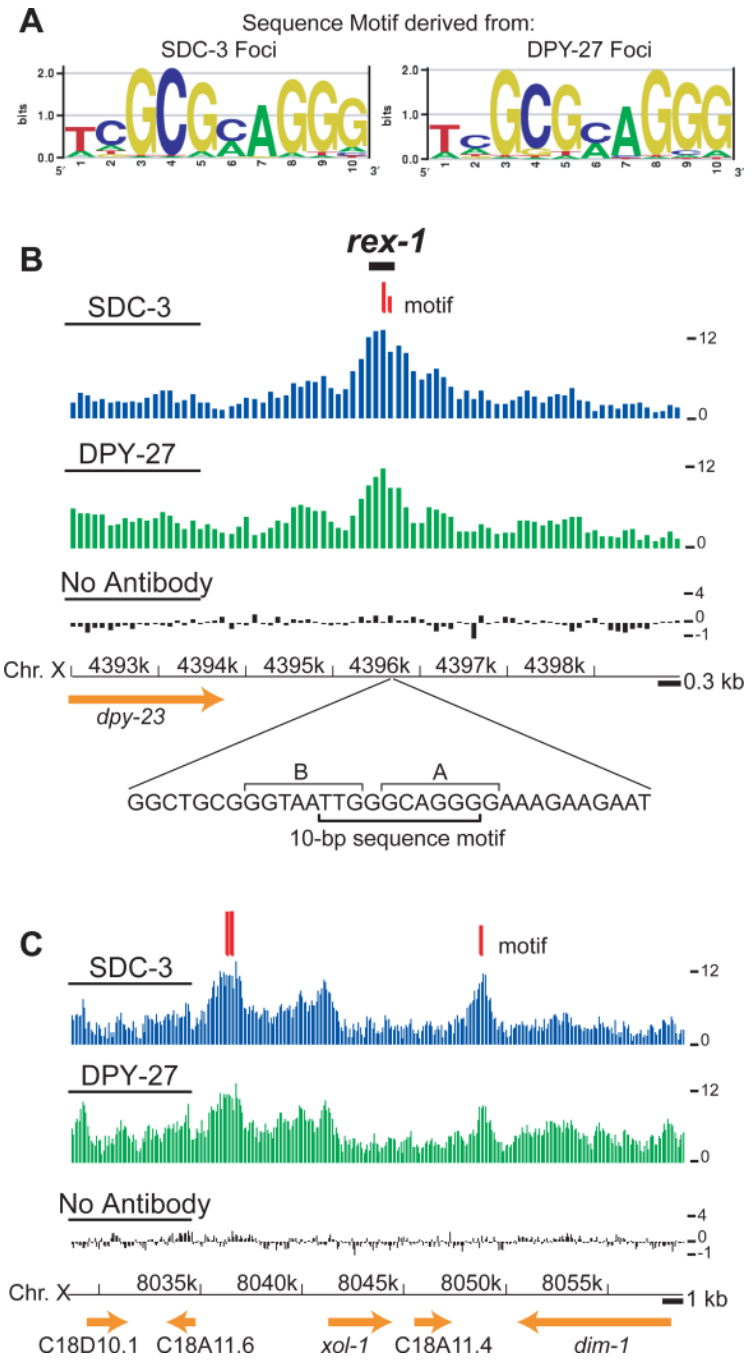

Figure 4. A stereotypic 10 bp DNA sequence motif is enriched in foci of DCC binding

(A) DNA sequence motifs derived from foci of SDC-3 and DPY-27 binding, depicted by sequence logos[30]. **(B)** DCC binding and motif occurrence at *rex-1*. Median z scores are plotted with respect to the chromosome coordinates for SDC-3, DPY-27, and no antibody control ChIPs. Locations of the 10-bp motifs are depicted with red tick marks, with longer marks indicating better matches to the consensus motif (MatrixScan $p \leq 10^{-5}$) than shorter marks (MatrixScan $p \leq 10^{-4}$). The location and sequence of a 33-bp fragment of *rex-1* competent for DCC recruitment is shown below. In this fragment, the motif we identify spans two previously characterized[12] motifs (7-bp A and 8-bp B), which were derived from four *rex* sites. The ChIP data suggest the previously reported "B" motif (TGTAATTG)[12] plays a weak role, if any, in

DCC recruitment on natural chromosomal substrates (Supplementary Figure 5). (C) DCC binding and motif occurrence near *xol-1*. Longer marks indicate better matches to the consensus motif (MatrixScan $p \leq 10^{-6}$) than shorter marks (MatrixScan $p \leq 10^{-5}$).

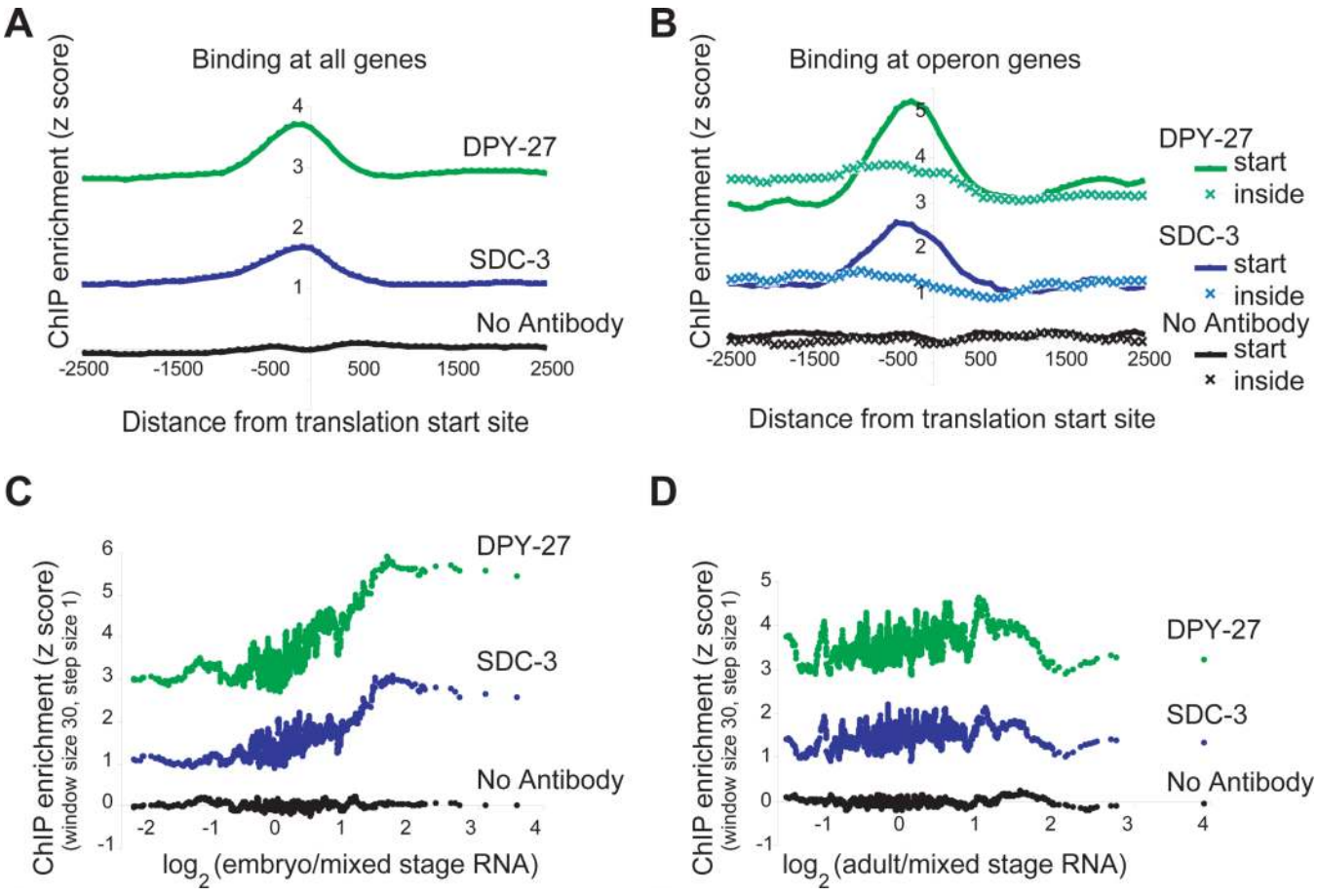

Figure 5. The DCC preferentially binds upstream of genes and is positively correlated with transcriptional activity

(A) Data was centered (coordinate 0) according to the predicted translation start sites of each X-linked gene, and the average z score of the probes in a sliding window (width 3 probes, step size 1 probe) was plotted, taking into account the direction of transcription. Translation start sites were used as a proxy for transcription start sites because they are much better-annotated. **(B)** Translation start sites of X-linked genes located at the start (solid, 56 genes) or inside (hatched, 69 genes) of an operon were centered. As in panel A, DCC binding is plotted as a function of distance from translation start site. **(C)** A moving average of DCC ChIP enrichment (performed in embryos) plotted as a function of embryonic transcript level of the downstream gene[6]. The positive correlation between embryonic DCC binding and embryonic transcription persisted when RNA measurements from a second publication were utilized[5] (Supplementary Figure 4C). **(D)** As in panel C, except adult RNA levels are used. Additionally, as expected, DCC binding was not correlated with male- or hermaphrodite-specific expression (Supplementary Figure 4D).