# X-Chromosome Target Specificity Diverged Between Dosage Compensation Mechanisms of Two Closely Related *Caenorhabditis* Species

Qiming Yang[1,2◆], Te-Wen Lo[1,2◆+], Katjuša Brejc[1,2], Caitlin Schartner[1,2●], Edward J. Ralston[1,2], Denise M. Lapidus[1,2], and Barbara J. Meyer[1,2*]

[1]Howard Hughes Medical Institute,
United States

[2]Department of Molecular and Cell Biology
University of California at Berkeley
Berkeley, United States

◆Q.Y and ◆T.-W.L. contributed equally to this work.

*For correspondence:
bjmeyer@berkeley.edu

Present addresses:
+Department of Biology
Ithaca College
Ithaca, NY 14850

●Roche Diagnostics
2881 Scott Blvd.
Santa Clara, CA 95050

Competing interests: The authors declare that no competing interests exist.

ORCID identifiers: Q.Y., 0000-0003-1419-868X; T.-W.L., 0000-0002-1231-5531; K.B., 0000-0002-4562-6109; B.J.M., 0000-0002-6530-4588

## Abstract

An evolutionary perspective enhances our understanding of biological mechanisms. Comparison of sex determination and X-chromosome dosage compensation mechanisms between the closely related nematode species *C. briggsae* (*Cbr*) and *C. elegans* (*Cel*) revealed that the genetic regulatory hierarchy controlling these processes is conserved, but both the X-chromosome target specificity and mode of binding for the specialized condensin dosage compensation complex (DCC) controlling X gene expression have diverged. We identified two motifs within *Cbr* DCC recruitment sites that are highly enriched on X: 13-bp MEX and 30-bp MEX II. Mutating either MEX or MEX II in an endogenous recruitment site with multiple copies of one or both motifs reduced binding, but only removing all motifs eliminated binding *in vivo*. Hence, DCC binding to *Cbr* recruitment sites appears additive. In contrast, DCC binding to *Cel* recruitment sites is synergistic: mutating even one motif *in vivo* eliminated binding. Although all X-chromosome motifs share the sequence CAGGG, they have otherwise diverged so that a motif from one species cannot function in the other. This functional divergence was demonstrated *in vivo* and *in vitro*. A single nucleotide position in *Cbr* MEX can act as a critical determinant for whether *Cel* DCC binds. The rapid divergence of DCC target specificity could have contributed to nematode speciation and contrasts dramatically with the conservation of target specificity for transcription factors that control developmental processes such as body-plan specification from fruit flies to mice.

## Introduction

Comparative studies have shown that different facets of metazoan development exhibit remarkably different degrees of conservation across species (Carroll, 2008). At one extreme, homeobox-containing *Hox* genes and *Wnt* pathway signaling genes play conserved roles in body plan formation (*Hox*) as well as cell-fate determination, neural patterning, or organogenesis (*Wnt*) across clades diverged by at least 600 million years (MYR) (Malicki *et al.*, 1990; De Kumar and Darland, 2021; Rim *et al.*, 2022). Distant orthologous genes within these ancestral pathways can often substitute for each other. For example, both the mouse *Small eye* (*Pax-6*) gene (Hill *et al.*, 1991) and the fruit fly *eyeless* (*ey*) gene (Quiring *et al.*, 1994; Halder *et al.*, 1995) control eye morphogenesis and encode a transcription factor that includes a paired domain and a homeodomain. Ectopic expression of mouse *Pax-6* in diverse fruit fly imaginal disc primordia induced morphologically normal ectopic compound eye structures on fruit fly wings, legs, and antennae (Halder *et al.*, 1995). Hence, at a deep level, eye morphogenesis is under related genetic and molecular control in vertebrates and insects, despite profound differences in eye morphology and mode of development.

At the other extreme are aspects of development related to sex. For example, chromosomal strategies to determine sexual fate in mice, fruit flies, and nematodes (XY or XO males and XX females or hermaphrodites) and the mechanism needed to compensate for the consequent difference in X-chromosome dose between sexes have diverged greatly. To balance X gene expression between sexes, female mice randomly inactivate one X chromosome (Yin *et al.*, 2021; Loda *et al.*, 2022), male fruit flies double expression of their single X chromosome (Samata and Akhtar, 2018; Rieder *et al.*, 2019), and hermaphrodite worms halve expression of both X chromosomes (Meyer, 2022a; Meyer, 2022b)

The divergence in these pathways is so great that comparisons among animals of the same genus can provide useful evolutionary context for understanding the developmental mechanisms that distinguish the sexes. Therefore, we determined the genetic and molecular specification of sexual fate and X-chromosome dosage compensation in the nematode *Caenorhabditis briggsae* (*Cbr*) and

compared it to the wealth of knowledge amassed about these processes in *Caenorhabditis elegans* (*Cel*). These two species have diverged by 15-30 MYR (Cutter, 2008).

In *C. elegans*, the sex determination and dosage compensation pathways are linked by genes that coordinately control both processes. For example, in XX embryos, the switch gene *sdc-2* sets the sex determination pathway to the hermaphrodite mode and triggers binding of a dosage compensation complex (DCC) onto both X chromosomes to reduce X gene expression by half and thereby match X expression with that from XO males (Meyer, 2022a). The DCC shares subunits with condensin, a protein complex that controls the structure, resolution, and segregation of mitotic and meiotic chromosomes from yeast to humans (Yatskevich *et al.*, 2019; Meyer, 2022a).

We determined the extent to which the sex-specific gene regulatory hierarchy is conserved between *C. elegans* and *C. briggsae* and the extent to which subunits of the *C. briggsae* DCC correspond to those of the *C. elegans* DCC. We also defined the *cis*-acting regulatory sites that confer X-chromosome specificity and recruit the *C. briggsae* DCC. We found that the DCC itself and the regulatory hierarchy that determines sex and directs the DCC to X have been conserved, but remarkably, both the X-chromosome target specificity of the *C. briggsae* DCC and its mode of binding to X have diverged.

## Results

### Conservation between *C. briggsae* and *C. elegans* of the core dosage compensation machinery and genetic hierarchy that regulates dosage compensation

The pivotal hermaphrodite-specific regulatory protein that controls both sex determination and dosage compensation in *C. elegans* is SDC-2, a 350 kDa protein with no known homologs outside of nematodes and only a coiled-coil domain as a predicted structural feature (Meyer, 2022a). It directs the DCC to both X chromosomes of XX embryos to achieve dosage compensation and also activates the hermaphrodite program of sexual differentiation (Chuang *et al.*, 1996; Dawes *et al.*, 1999; Chu *et al.*, 2002; Pferdehirt *et al.*, 2011). Loss of *Cel sdc-2* causes XX-specific lethality due to excessive X-chromosome gene expression and masculinization of escaper animals (Nusbaum and Meyer, 1989;

Kruesi *et al.*, 2013). To assess the conservation of gene function, we deployed genome-editing technology in *C. briggsae* to knockout *sdc-2*, which shares 31% amino acid identity and 44% similarity with *C. elegans* (Figure 1—figure supplement 1). Using a PCR-based molecular strategy to identify insertions and deletions induced by DNA repair following directed mutagenesis with zinc finger nucleases, we recovered several independent *Cbr sdc-2* mutant lines (Figure 1—figure supplement 2). Homozygous *Cbr sdc-2* mutations caused extensive XX-specific lethality, consistent with a defect in dosage compensation and the conservation of gene function (Figure 1A). Nearly all *Cbr sdc-2* hermaphrodites died as embryos or young larvae; rare XX survivors exhibited slow growth and masculinization. *Cbr sdc-2* males were viable (Figure 1A) and had wild-type body morphology.

To determine whether the hermaphrodite-specific lethality of *Cbr sdc-2* mutants was caused by defects in dosage compensation, we first identified components of the *C. briggsae* DCC and then asked whether DCC binding to X is disrupted by a *Cbr sdc-2* mutation, as it is by a *Cel sdc-2* mutation. In *C. elegans*, five of the ten known DCC proteins are homologous to subunits of condensin, an evolutionarily conserved protein complex required to restructure and resolve chromosomes in preparation for cell divisions in mitosis and meiosis (Figure 1B) (Chuang *et al.*, 1994; Lieb *et al.*, 1996; Lieb *et al.*, 1998; Chan *et al.*, 2004; Tsai *et al.*, 2008; Csankovszki *et al.*, 2009; Mets and Meyer, 2009; Yatskevich *et al.*, 2019; Meyer, 2022a). The evolutionary time scale over which condensin subunits became coopted for dosage compensation in nematodes had not been explored.

Several lines of evidence indicate that a condensin complex mediates dosage compensation in *C. briggsae* as well. First, BLASTP searches revealed *C. briggsae* orthologs of all known *C. elegans* DCC condensin subunits (Figure 1B). Immunofluorescence experiments using antibodies raised against *Cbr* DPY-27, the SMC4 ortholog of the sole *Cel* DCC condensin subunit (*Cel* DPY-27) not associated with mitotic or meiotic condensins (Chuang *et al.*, 1994), revealed X chromosome-specific localization in hermaphrodites, but not males, indicating conservation of function (Figure 1C and Figure 2A,B). Second, mass-spectrometric analysis of proteins that co-immunoprecipitated with *Cbr* DPY-27 identified *Cbr* MIX-1 (Figure 1—table supplement 1), the SMC2 condensin subunit ortholog (Lieb *et al.*, 1998). Immunofluorescence experiments using *Cbr* MIX-1 antibodies revealed co-localization of *Cbr* MIX-1 with

*Cbr* DPY-27 on hermaphrodite X chromosomes (Figure 1D). Third, disruption of *Cbr dpy-27* conferred hermaphrodite-specific lethality, with XX escaper animals exhibiting a dumpy phenotype, like disruption of *Cel dpy-27* (Figure 1F). *Cbr* MIX-1 protein did not localize to X in *Cbr dpy-27(y436)* mutant animals, consistent with the two proteins participating in a complex (Figure 1B,E). These data demonstrate that condensin subunits play conserved roles in the dosage compensation machinery of both *C. briggsae* and *C. elegans*.

Evidence that DCC binding defects underlie the XX-specific lethality caused by *Cbr sdc-2* mutations is our observation that neither *Cbr* DPY-27 (Figure 2C) nor *Cbr* MIX-1 (not shown) localizes to X chromosomes in *Cbr sdc-2* mutant hermaphrodites. Thus, the role for *sdc-2* in the genetic hierarchies that activate dosage compensation is also conserved.

We next explored why maternally supplied DCC subunits fail to bind to the single X chromosome of *C. briggsae* males. In *C. elegans* XO embryos, the master switch gene *xol-1* (XO lethal) represses the hermaphrodite-specific *sdc-2* gene required for DCC binding to X and thereby prevents other DCC subunits from functioning in males (Miller *et al.*, 1988; Rhind *et al.*, 1995; Dawes *et al.*, 1999; Meyer, 2022a). Loss of *Cel xol-1* activates *Cel sdc-2* in XO embryos, causing DCC binding to X, reduction in X-chromosome gene expression, and consequent death. We isolated a null mutant allele of *Cbr xol-1(y430)* by PCR screening of a *C. briggsae* deletion library (Supplementary File 1). We found that the *Cbr xol-1* mutation caused inappropriate binding of the DCC to the single X of XO embryos (Figure 2D) and fully penetrant male lethality (Figure 3B), as expected from the disruption of a gene that prevents the DCC machinery from functioning in *C. briggsae* males. *Cbr xol-1* mutant XX hermaphrodites appeared wild type.

To investigate the hierarchical relationship between *Cbr xol-1* and *Cbr sdc-2*, we asked whether a *Cbr sdc-2* mutation could suppress the male lethality caused by a *Cbr xol-1* mutation. Both genes are closely linked in *C. briggsae*, prompting us to use genome editing technology to introduce *de novo* mutations in *cis* to pre-existing lesions without relying on genetic recombination between closely linked genes. If *Cbr xol-1* controls *Cbr sdc-2*, then mutation of *Cbr sdc-2* should rescue the male lethality of *Cbr xol-1* mutants (Figure 2E). This prediction proved to be correct. XO males were observed among F1

progeny from mated *Cbr xol-1* hermaphrodites injected with ZFNs targeting *Cbr sdc-2* (Figure 3A,B,D). Insertion and deletion mutations were found at the *Cbr sdc-2* target site in more than twenty tested F1 males (examples are in Figure 1—figure supplement 2C,D). Quantification of male viability in four different *xol-1 sdc-2* mutant lines revealed nearly full rescue (Figure 3B), with concomitant absence of DCC binding on the single X chromosome (Figure 2E). Therefore, *Cbr xol-1* functions upstream of *Cbr sdc-2* to repress it and thereby prevents DCC binding to the male X chromosome. In summary, not only is the core condensin dosage compensation machinery conserved between *Caenorhabditis* species, so also are the key features of the genetic hierarchy that confers sex-specificity to the dosage compensation process.

## Conservation between *C. briggsae* and *C. elegans* of the genetic hierarchy that regulates sex determination

Mechanisms controlling sex determination and differentiation are dynamic over evolutionary time; major differences can exist even within individual species. For example, males within the house fly species *Musca domestica* can utilize one of many different male-determining factors on autosomes and sex chromosomes to determine sex depending on a factor's linkage to other beneficial traits (Meisel *et al.*, 2016). Within the *Caenorhabditis* genus, differences and similarities have been observed in the genetic pathways governing the later stages of sexual differentiation (Haag, 2005). In fact, hermaphroditism evolved independently for *C. elegans* and *C. briggsae*; sperm production differs between them (Yin *et al.*, 2018; Cutter *et al.*, 2019). However, the earlier stages of sex determination had not been analyzed in *C. briggsae*. Therefore, we asked whether *xol-1* and *sdc-2* control sexual fate as well as dosage compensation in *C. briggsae*, as they do in *C. elegans*.

Our analysis of *Cbr sdc-2* XX mutant phenotypes revealed intersexual tail morphology in the rare animals that survived to the L3/L4 stage (Figure 3C), indicating a role for *Cbr sdc-2* in sex determination. Sexual transformation to the male fate is unlikely to have resulted from a disruption in dosage compensation, since such transformation was never observed in *Cbr dpy-27* XX mutants (Figure 3C). Analysis of sexual phenotypes in double mutant strains confirmed that *Cbr sdc-2* controls sex

determination. Specifically, *Cbr xol-1 Cbr sdc-2* double mutant XO animals develop as males, whereas *Cbr dpy-27*; *Cbr xol-1* double mutant XO animals develop as hermaphrodites (Figure 3C,D). That is, both *Cbr sdc-2* and *Cbr dpy-27* mutations suppress the XO lethality caused by a *xol-1* mutation, but only *Cbr sdc*-2 mutations also suppress the sexual transformation of XO animals into hermaphrodites. These results show that both *sdc-2* and *dpy-27* function in *C. briggsae* dosage compensation, but only *sdc-2* also functions in sex determination. Thus, the two master regulatory genes that control both sex determination and X-chromosome dosage compensation are conserved between *C. briggsae* and *C. elegans*.

## DCC recruitment sites isolated from *C. briggsae* X chromosomes fail to bind the *C. elegans* DCC

Discovery that the dosage compensation machinery and the gene regulatory hierarchy that control sex determination and dosage compensation are functionally conserved between *C. briggsae* and *C. elegans* raised the question of whether the *cis*-acting regulatory sequences that recruit dosage compensation proteins to X chromosomes are also conserved. In *C. elegans*, the DCC binds to recruitment elements on X (*rex*) sites and then spreads across X to sequences lacking autonomous recruitment ability (Csankovszki *et al.*, 2004; Jans *et al.*, 2009; Pferdehirt *et al.*, 2011; Albritton *et al.*, 2017; Anderson *et al.*, 2019). Within *rex* sites, combinatorial clustering of three DNA sequence motifs directs synergistic binding of the DCC (Fuda *et al.*, 2022). To compare X-recruitment mechanisms between species, DNA binding sites for the *Cbr* DCC recruitment protein SDC-2 and the C*br* DCC condensin subunit DPY-27 were defined by chromatin immuno-precipitation experiments followed by sequencing of captured DNA (ChIP-seq experiments) (Figure 4A). SDC-2 sites were obtained with anti-FLAG antibodies from a genome-engineered *Cbr* strain encoding a FLAG-tagged version of endogenous SDC-2. DPY-27 sites were obtained from either a wild-type *Cbr* strain with DPY-27 antibodies or from a genome-engineered strain encoding endogenous FLAG-tagged DPY-27 with anti-FLAG antibodies.

A consistent set of twelve large, overlapping SDC-2 ChIP-seq peaks and DPY-27 ChIP-seq peaks emerged from the studies (Figure 4A), representing less than one-fourth the number of DCC peaks than on the *C. elegans* X chromosome, which is smaller (17.7 Mb for *Cel* vs. 21.5 Mb for *Cbr*). Confirmation that each of these twelve *Cbr* peaks represented an actual DCC binding site resulted from DCC recruitment assays conducted *in vivo* (Figure 4B). Embryos carrying extrachromosomal arrays composed of multiple copies of DNA from a single ChIP-seq peak were stained with DPY-27 antibodies and a FISH probe to the array. Typically, DPY-27 localized to 80-90% of extrachromosomal arrays carrying DNA from an individual peak (Figure 4C and 4E and Figure 4—table supplement 1A). In contrast, extrachromosomal arrays made from three regions of X lacking DCC binding in ChIP-seq experiments showed minimal recruitment (0-6% of nuclei with arrays) (Figure 4E and Figure 4—table supplement 1A). In strains with arrays comprised of *Cbr* DCC binding sites, the X chromosomes rarely exhibited fluorescent signal, because the arrays titrated the DCC from X (Figure 4C). The titration was so effective that brood sizes of array-bearing hermaphrodites were very low, and hermaphrodite strains carrying arrays could not be maintained. *Cbr* DCC binding sites that were confirmed by array assays were named recruitment elements on X (*rex* sites) (Table 1), like the *C. elegans* DCC binding sites, due to their ability to recruit the DCC when detached from X.

To determine whether *rex* sites from *C. briggsae* and *C. elegans* had functional overlap in DCC binding specificity, we asked whether a *rex* site from one species could recruit the DCC from the other. We made extrachromosomal arrays in *C. elegans* with DNA from *C. briggsae rex* sites and extrachromosomal arrays in *C. briggsae* with DNA from *C. elegans rex* sites. Arrays in *C. elegans* with *C. briggsae rex* sites failed to recruit the *Cel* DCC or to titrate the *Cel* DCC from *Cel* X chromosomes (Figure 4C, *Cbr rex-8*), indicating evolutionary divergence in *rex* sites between the two *Caenorhabditis* species. Reciprocally, extrachromosomal arrays made in *C. briggsae* with *Cel rex* sites failed to bind the *Cbr* DCC or titrate it from the *Cbr* X, confirming divergence in *rex* sites (*Cel rex-33* in Figure 4D; *Cel rex-33* and *Cel rex-4* in Figure 4—table supplement 1B). In contrast, controls showed that 100% of extrachromosomal arrays made in *C. elegans* with DNA from either *Cel rex-33* or *Cel rex-4* recruited the *Cel* DCC (Figure 4—table supplement 1B).

Because X chromosomes and extrachromosomal arrays have different topologies, histone modifications, DNA binding proteins, and positions within nuclei, we devised a separate assay to assess the divergence of *rex* sites between species in a more natural chromosomal environment. We inserted six *Cbr rex* sites with a range of ChIP-seq scores into a location on the endogenous *Cel* X chromosome that lacked DCC binding (15, 574, 674 bp) (Figure 5 and Table 1). Proof of principle for the experiment came from finding that insertion of *Cel rex-32*, a high-affinity *Cel* DCC binding site, into the new location on X resulted in DCC binding that was not significantly different from binding at its endogenous location on X ($p = 0.2$, Figure 5). All *Cbr rex* sites except *rex-1*, which will be discussed later, failed to recruit the *Cel* DCC when inserted into *Cel* X chromosomes, confirming the divergence of *rex* sites between species.

## Identification of motifs on *Cbr* X chromosomes that recruit the *Cbr* DCC

To understand the mechanisms underlying the selective recruitment of the *Cbr* DCC to X chromosomes, but not autosomes, and the basis for the divergence in X-chromosome targeting between *Caenorhabditis* species, we searched for DNA sequence motifs that are enriched in the twelve *Cbr rex* sites (Figure 6—figure supplement 1A) using the website-based program called Multiple Em fo Motif Elicitation (MEME) (Version 5.4.1) (Bailey and Elkan, 1994; Bailey *et al.*, 2015) and compared them to motifs in *C. elegans rex* sites important for recruiting the *Cel* DCC to X (Figure 6A,B). We found two motifs enriched within *Cbr rex* sites that are highly enriched on *Cbr* X chromosomes compared to autosomes (Figure 6A; Figure 7A,B; Table 1). A 13-bp motif named MEX (Motif Enriched on X) is enriched up to 12-fold on X chromosomes versus autosomes, and a 30-bp motif named MEX II is enriched up to 30-fold on X versus autosomes (Figure 7A,B). The similarity of a motif to the consensus motif is indicated by the ln(P) score, which is the natural log of the probability that the 13-mer for MEX or the 30-mer for MEX II matches the respective consensus motif matrix as calculated by the Patser program (Hertz and Stormo, 1999). The lower the score, the better the match. For both MEX and MEX II, the lower the ln(P) score, and hence the better the match to the consensus sequence, the more highly enriched is the motif on X chromosomes compared to autosomes (Figure 7A,B).

Our analysis revealed that only the *Cbr* MEX (Figure 7C) or MEX II (Figure 7D) motifs on X that are located within *rex* sites are bound by SDC-2. Negligible SDC-2 binding was found at single, isolated MEX (Figure 7C) or MEX II (Figure 7D) motifs on X that are not in *rex* sites. These results implicate MEX and MEX II as important elements for *Cbr* DCC recruitment to *rex* sites.

Neither of the *Cbr* motifs is enriched on the X chromosomes of *C. elegans*, indicating motif divergence between species (Figure 7A,B). No additional enriched *C. briggsae* motif candidates were found when the sequences of the two motifs in the twelve *rex* sites were eliminated from the search by converting them to N's and searches for potential motifs were conducted again. In addition, motif analysis of DNA from SDC-2 and DPY-27 ChIP-seq peaks with intermediate or low levels of DCC binding (*i.e.*, lower than for *rex-2*) (Figure 6—figure supplement 1B) revealed no motif candidates that correlate with DCC binding.

In *C. elegans*, two motifs are highly enriched on X chromosomes relative to autosomes: a 12-bp motif also called MEX and a 26-bp motif called MEX II (Figure 6B) (Fuda *et al.*, 2022). These *C. elegans* X-enriched motifs are not enriched on *C. briggsae* X chromosomes (Figure 6B and Figure 7--figure supplement 7A,B). *Cbr* MEX as well as *Cel* MEX and *Cel* MEX II share a common core sequence of CAGGG (Figure 6), which is necessary but not sufficient for DCC binding in *C. elegans* (Fuda *et al.*, 2022). The core is likely indicative of a common evolutionary history between species. However, the *Cbr* and *Cel* motifs diverged sufficiently that the motifs from one species are not enriched on the X chromosomes of the other species. Moreover, the *Cbr* MEX motif has a nucleotide substitution that would render the *Cel* MEX motif incapable of binding the *Cel* DCC. Predominantly, the *C. elegans* consensus MEX motif has a cytosine nucleotide located two nucleotides 5' to the core CAGGG sequence: 5'-TCG<u>C</u>GCAGGGAG-3' (Figure 6B). Mutational analysis in *C. elegans* demonstrated that replacing that nucleotide with a guanine greatly reduced DCC binding both *in vivo* and *in vitro* (Fuda *et al.*, 2022). The consensus *Cbr* MEX motif has a guanine at that critical location, and in principle, the *Cbr* MEX motif would not function as a *Cel* DCC binding motif (Figure 6), thereby offering insight into the divergence of X-chromosome binding sites between species.

In *C. elegans*, a 9-bp motif called Motif C also participates in *Cel* DCC recruitment to X but lacks enrichment on X (Figure 6B) (Fuda *et al.*, 2022). Sequences between the clustered Motif C variants within a *Cel rex* site are also critical for DCC binding (Fuda *et al.*, 2022). Evidence that *C. elegans* Motif C fails to participate in *Cbr* DCC recruitment to *Cbr* X chromosomes is our finding that *Cbr* SDC-2 binding is negligible at *Cel* Motif C variants on *Cbr* X, except in the case of rare variants (0.26% of all *Cel* Motif C variants on X) that are within *bona fide* MEX or MEX II motifs in *Cbr rex* sites (Figure 7—figure supplement 1C). The likely reason that *Cbr rex-1* recruits the *Cel* DCC when inserted into *Cel* X chromosomes (Figure 5) is that each of the four *Cbr* MEX motifs includes a strong match to the consensus *Cel* Motif C (Figure 5 legend), and DNA sequences surrounding the *Cel* Motif C variants in *Cbr rex-1* are highly conserved with the syntenic region of *C. elegans*, which includes *Cel rex-34*. Both *Cel rex-34* and *Cbr rex-1* are within coding regions of orthologous *pks-1* genes. In contrast, *Cbr rex-7* also contains Motif C variants but lacks necessary surrounding sequences to permit *Cel* DCC binding when inserted on the *Cel* X (Figure 5).

### Mutational analysis of motifs on endogenous *C. briggsae* X chromosomes showed that combinatorial clustering of motifs in *rex* sites facilitates *Cbr* DCC binding but some binding can still occur with only a single motif in a *rex* site

To assess further the importance of the *Cbr* motifs and the divergence of motifs between species, we performed mutational analyses of the two *Cbr* X-enriched motifs. Initial demonstration that both *Cbr* MEX and *Cbr* MEX II motifs participate in DCC binding at *Cbr rex* sites in *C. briggsae* came from analysis using extrachromosomal arrays carrying wild-type and mutant forms of *Cbr rex-1* (Figure 8—figure supplement 1). Eighty-nine percent of *C. briggsae* nuclei carrying extrachromosomal arrays composed of wild-type *rex-1* sequences recruited the DCC and titrated it away from X. In contrast, only 24% of nuclei carrying arrays with mutant copies of *rex-1* lacking MEX II recruited the DCC, demonstrating the importance of MEX II. Only 38% of nuclei carrying arrays with mutant copies of *rex-1* lacking all four MEX motifs recruited the DCC, demonstrating the importance of MEX. DCC binding was

reduced to 6% of arrays carrying mutant copies of *rex-1* lacking both MEX II and the four MEX motifs. Hence, both motifs contribute to DCC binding.

This conclusion was reinforced by using genome editing to mutate the MEX II sequence or all MEX II and MEX sequences in the endogenous *rex-1* site on *C. briggsae* X chromosomes and then assaying DCC binding (Figure 8). ChIP-seq analysis revealed significant reduction in DPY-27 binding at *rex-1* lacking MEX II sequences and negligible DPY-27 binding at *rex-1* lacking both MEX and MEX II sequences. Hence, clustering of motifs in the endogenous *rex-1* on X is important for DCC binding (Figure 8).

To evaluate more precisely the participation of different *Cbr* motifs in DCC binding, we used genome editing at three endogenous *rex* sites to evaluate the interplay between MEX and MEX II motifs, only MEX II motifs, or only MEX motifs. Eliminating either MEX or MEX II in *rex-4* reduced binding significantly, but binding was evident at the remaining motif (Figure 9 and Figure 9—figure supplement 1). Binding was dramatically reduced when both motifs were mutated. This result demonstrates that an individual MEX or MEX II motif can confer significant DCC binding at a *rex* site, but both motifs are needed for full DCC binding.

Equivalent results were found by mutating either of the two MEX II motifs in *rex-3* or combinations of the three MEX motifs in *rex-7*. For *rex-3*, DCC binding was reduced significantly when one of the two MEX II motifs was mutated, but significant binding occurred at either of the remaining MEX II motifs (Figure 10 and Figure 10—figure supplement 1). Binding was greatly reduced when both motifs were mutated. For *rex-7*, DCC binding at the endogenous site lacking the MEX motif with the best match to the consensus sequence (-18.22) was not significantly different from binding at the wild-type site. In contrast, mutating different combinations of two motifs (-18.72 and -12.26 or -18.7 and -12.58) reduced binding significantly (Figure 11 and Figure 11—figure supplement 1). Mutating all three motifs reduced binding severely. Results with the four *Cbr rex* sites, *rex-1, rex-3, rex-4,* and *rex-7* demonstrate that combinatorial clustering of motifs achieves maximal DCC binding at *Cbr rex* sites, but significant binding can occur at a single motif.

These results contrast with results in *C. elegans*. Mutating individual motifs, either MEX, MEX II, or Motif C, at an endogenous *C. elegans rex* site with multiple different motifs dramatically reduced DCC binding *in vivo* to nearly the same extent as mutating all motifs, demonstrating synergy in DCC binding (Fuda *et al.*, 2022). Hence, not only have the motifs diverged between species, the mode of binding to motifs has also changed.

### Functional divergence of motifs demonstrated by *Cel* DCC binding studies *in vivo* and *in vitro* to a *Cel rex* site with *Cbr* MEX and MEX II motifs replacing *Cel* motifs

To explore the divergence in motifs between species in greater detail, we replaced each of the two MEX II motifs of the endogenous *Cel rex-39* site on X with a copy of MEX II from *Cbr rex-3* and assayed the level of *Cel* SDC-3 binding *in vivo* by ChIP-qPCR (Figure 12A,B). SDC-3 binding *in vivo* was negligible at the *Cel rex-39* site with the *Cbr* MEX II motifs and indistinguishable from binding at the *Cel rex-39* site with two scrambled MEX II motifs, thus demonstrating the high degree of functional divergence between MEX II motifs of different species (Figure 12B).

We performed a similar analysis for MEX motifs. We replaced the three MEX motifs in endogenous *Cel rex-33* with the three *Cbr* MEX motifs from endogenous *Cbr rex-7* (Figure 12D). SDC-3 binding *in vivo* was negligible at the *Cel rex-33* site with the *Cbr* MEX motifs and indistinguishable from binding at the *Cel rex-33* site with three scrambled MEX motifs, demonstrating the functional divergence between MEX motifs of different species (Figure 12E).

As a second approach, we conducted DCC binding studies *in vitro*. We assayed *Cel* DCC binding *in vitro* to a *Cel rex-39* site with two *Cbr* MEX II motifs (Figure 12C) and to the *Cel rex-33* site with the three *Cbr* MEX motifs (Figure 12F). Our prior studies developed a robust *in vitro* assay using *C. elegans* embryo extracts to evaluate the role of motif combinations toward *Cel* SDC-2 binding at *rex* sites (Fuda *et al.*, 2022). The advantage of this assay is that *Cel* SDC-2 is capable of binding to a single motif on an *in vitro* template, perhaps because that DNA lacks the competing binding of nucleosomes and general transcription factors that occurs *in vivo* (Fuda *et al.*, 2022). If either of the *Cbr* MEX II motifs inserted into the *Cel rex-39* site were functional or if any of the three *Cbr* MEX motifs inserted into the *Cel rex-33* site were functional, we would detect *Cel* SDC-2 binding to the template *in vitro*.

The *in vitro* assay demonstrated robust binding of *Cel* SDC-2 to the wild-type *Cel rex-39* template (Figure 12C) and to the wild-type *Cel rex-33* template (Figure 12F), as shown previously (Fuda *et al.*, 2022). However, *Cel* SDC-2 binding at the *Cel rex-39* site with substituted *Cbr* MEX II motifs was indistinguishable from binding to the mutant *Cel rex-39* template with two scrambled *Cel* MEX II motifs or to the negative control template made from *Cel* X DNA at a site lacking *Cel* DCC binding *in vivo* (Figure 12C). Similarly, *Cel* SDC-2 binding at the *Cel rex-33* site with substituted *Cbr* MEX motifs was indistinguishable from binding to the mutant *Cel rex-33* template with three scrambled *Cel* MEX motifs or to the negative control template (Figure 12F). Thus, the *in vitro* assay demonstrates that substituting *Cbr* MEX II or MEX motifs for *Cel* MEX II or MEX motifs in a *Cel rex* site eliminates *Cel* DCC binding.

## A single nucleotide position in the consensus *Cbr* MEX motif acts as a critical determinant for whether the *Cel* DCC can bind to *Cbr* MEX

In contrast to the many nucleotide changes that mark the difference between MEX II motifs in *C. briggsae* versus *C. elegans*, the MEX motifs are strikingly similar in nucleotide composition and core CAGGG sequence between species (Figure 6). The most significant change between the consensus MEX motifs is the substitution in *Cbr* MEX of a guanine for the cytosine in *Cel* MEX located two nucleotides 5' from the CAGGG core of both motifs (Figure 13A). That C4G transversion was not found in a functional *Cel* MEX motif *in vivo* or *in vitro*, and creating a C4G change in either the MEX motif of endogenous *Cel rex-1* or in an *in vitro Cel* DNA template reduced binding (Fuda *et al.*, 2022). In principle, that single cytosine-to-guanine transversion could be a critical evolutionary change in MEX motifs that render the motifs incapable of binding the DCC from the other species. To test this hypothesis, we made the C4G transversion in each of the three MEX motifs within the endogenous *Cel rex-33* site (Figure 13B). *Cel* SDC-3 binding *in vivo* to the C4G-substituted *Cel rex-33* site was reduced to the same level of binding as that at the *Cel rex-33* site with all three *Cel* MEX motifs scrambled, confirming the functional significance of the nucleotide substitution between species (Figure 13B). Our *in vitro* assay comparing *Cel* SDC-2 binding to the C4G-substituted and the MEX-scrambled *Cel rex-33* DNA templates produced the same result (Figure 13C).

If the evolutionary transversion of that C to G between *Cel* and *Cbr* MEX motifs represents an important step in the divergence of motif function, then making a G-to-C change within the *Cbr* MEX motifs (G7C) inserted into *Cel rex-33* might enhance *Cel* DCC binding. Indeed, the G7C change to *Cbr* MEX within *Cel rex-33* increased the *Cel* SDC-3 binding *in vivo* by 4.2-fold and increased the specific *Cel* SDC-2 binding *in vitro* by 4.3-fold. The G7C change increased *Cel* SDC-3 binding *in vivo* to 18% of its binding at wild-type *Cel rex-33* (Figure 13B) and increased *Cel* SDC-2 binding *in vitro* to 44% of its the specific binding at the wild-type *Cel rex-33* template (Figure 13C). Hence, the cytosine-to-guanine transversion between MEX motifs of *C. elegans* versus *C. briggsae* is important for the functional divergence in motifs.

## Discussion

Comparison of X-chromosome dosage compensation mechanisms between the closely related *Caenorhabditis* species *C. briggsae* and *C. elegans* revealed that both the dosage compensation machinery and the regulatory hierarchy that directs it to hermaphrodite X chromosomes have been conserved, but remarkably, the X-chromosome target specificity of the *C. briggsae* machinery and its mode of binding to X have diverged, as well as the density of DCC recruitment sites. The extent of evolutionary changes in dosage compensation mechanisms between species diverged by only 15-30 MYR is in striking contrast to changes in developmental processes such as body-plan specification and eye morphogenesis from fruit flies to mice, which instead utilize highly conserved transcription factors and *cis*-acting DNA regulatory sequences (Malicki *et al.*, 1990; Halder *et al.*, 1995).

Central to the dosage compensation machinery of both species is a specialized condensin complex. Here we identified two *C. briggsae* dosage compensation proteins (DPY-27 and MIX-1) that are orthologs of the SMC (structural maintenance of chromosome) subunits of condensin and bind to hermaphrodite X chromosomes. As in *C. elegans* (Chuang *et al.*, 1994; Lieb *et al.*, 1998), mutation of *dpy-27* causes hermaphrodite-specific lethality in *C. briggsae*, and MIX-1 fails to bind to X in the absence of DPY-27, consistent with both proteins acting in a complex. We also found that the hermaphrodite-specific *Cbr sdc-2* gene triggers binding of the condensin subunits to X and activates the

hermaphrodite mode of sexual differentiation, as in *C. elegans.* Mutation of *Cbr sdc-2* causes XX-specific lethality, and rare XX animals that escape lethality develop as masculinized larvae. SDC-2 and condensin subunits are prevented from binding to the single X of males by the action of *xol-1*, the master sex-determination gene that controls both sex determination and dosage compensation and triggers the male fate by repressing *sdc-2* expression. Mutation of *xol-1* kills XO animals because the DCC assembles on the single male X, thereby reducing gene expression inappropriately. Mutations in *sdc-2* or *dpy-27* suppress the XO-specific lethality caused by *xol-1* mutations, but only mutations in *sdc-2* permit the rescued animals to develop as males. Just as in *C. elegans*, XO animals rescued by *dpy-27* mutations develop as hermaphrodites, consistent with *dpy-27* controlling only dosage compensation and *sdc-2* controlling both sex determination and dosage compensation. Hence, the two master regulators that control sexual fate and dosage compensation are functionally conserved between the two *Caenorhabditis* species, as is the condensin dosage compensation machinery. In both species, SDC-2 recruits the condensin DCC subunits to X and is the likely protein to interact directly with X DNA. These 350 kDa proteins lack homology to proteins outside of *Caenorhabditis,* and the only predicted structural feature is a coiled-coil region.

Although the two species have conserved DCC machinery, the DCC binding sites have diverged, as has their density on X. ChIP-seq analysis of *C. briggsae* SDC-2 and DPY-27 revealed twelve sites of binding on X that were validated by functional analysis *in vivo* as being strong autonomous recruitment (*rex*) sites. Even though the X chromosome of *C. briggsae* (21.5 Mb) is larger than the X of *C. elegans* (17.7 Mb), it has only one-fourth the number of recruitment sites. The *C. briggsae* sites are sufficiently strong that extrachromosomal arrays carrying multiple copies of a single site can titrate the DCC from X and cause dosage-compensation-defective phenotypes in XX animals, including death, as in *C. elegans*. In contrast, extrachromosomal arrays of *C. briggsae rex* sites made in *C. elegans* fail to recruit the *C. elegans* DCC, and *vice versa*, indicating that *rex* sites have diverged between the two species. As a more rigorous test of divergence, we inserted individual *C. briggsae rex* sites in single copy into *C. elegans* X chromosomes and assayed binding. The *C. elegans* DCC failed to bind to the five *C. briggsae rex* sites inserted into *C. elegans* X chromosomes.

Not only have the *rex* sites diverged, the mechanism by which the Cbr DCC binds to X motifs differs from that of the *Cel* DCC. We identified two motifs within *C. briggsae rex* sites that are highly enriched on X, the 13-bp MEX motif and the 30-bp MEX II motif. Mutating one copy of either motif in endogenous *rex* sites with multiple motifs reduced binding, but significant binding still occurred at the sites. Binding was eliminated only when all motifs were removed. Hence, DCC binding to motifs in *C. briggsae rex* sites appears additive. In contrast, mutating one motif in *C. elegans rex* sites that have multiple different combinations of motifs reduced binding to nearly the same extent as mutating all motifs, indicating synergy in *C. elegans* DCC binding (Fuda *et al.*, 2022).

Additional factors, such as yet-unidentified DNA binding proteins might alter the specificity of DCC binding between species as well as aid DCC binding at *Cbr rex* sites. Precedent exists in the homeodomain family of Hox DNA binding proteins that have remarkably similar DNA specificities for DNA binding *in vitro* but a wide range of specificities *in vivo* due to interactions with heterologous cofactors required for functional specificity, such as Pbx-Hox complexes (Chang *et al.*, 1996).

The need for synergy in DCC binding to *Cel rex* sites is likely caused by competition between DCC binding and nucleosome formation, since nucleosomes preferentially bind to *rex* sites when DCC binding is precluded by mutations (Fuda *et al.*, 2022). The status of nucleosomes on *C. briggsae* X chromosomes remains to be determined. Although a single MEX or MEX II motif enables some DCC binding to a *Cbr rex* site, equivalent motifs on X that are not in *rex* sites appear to lack DCC binding. Nucleosome formation may preclude DCC binding at those motifs. The X may have a paucity of DNA-binding proteins that interact with core histones and open compacted chromatin to enable DCC binding.

Although the X-chromosome motifs of both species share the core consensus sequence CAGGG, the motifs have diverged such that they function in only *C. elegans* or *C. briggsae*. This functional divergence was demonstrated through DCC binding studies *in vivo* and *in vitro* to *C. elegans rex* sites engineered with *C. briggsae* motifs substituted for *C. elegans* motifs. We replaced the two MEX II motifs in the endogenous C*. elegans rex-39* site with *C. briggsae* MEX II motifs and the three MEX motifs in *Cel rex-33* with *Cbr* MEX motifs while maintaining motif spacing appropriate for *C. elegans*. We found negligible *C. elegans* DCC binding *in vivo* and in vitro. A feature of the *in vitro* assay

is that *Cel* SDC-2 is capable of binding to a single motif on a DNA template, likely because the DNA lacks competing binding of nucleosomes that occurs in vivo. If either *Cbr* MEX II or MEX motif were functional in *C. elegans* we would have detected binding.

While the MEX II motif has diverged sufficiently that evolutionary tracing is difficult, divergence of MEX motifs provides important insight into their evolution. A major difference in MEX motifs between the two species is the preference for a guanine instead of a cytosine two nucleotides 5' of the conserved CAGGG sequence. We demonstrated that converting that C to G in the three *Cel* MEX motifs of *Cel rex-33* eliminated DCC binding *in vitro*. Conversely, replacing the G nucleotide in each *Cbr* MEX motif inserted into *Cel rex-33* with a C nucleotide restored *Cel* DCC binding *in vivo* and *in vitro*, indicating that the single nucleotide change can be important in the evolutionary divergence of this motif. The evolutionary C-to-G substitution in the *Cbr* MEX motif is sufficient to prevent it from functioning in the closely related *C. elegans* species.

Like many developmental regulatory proteins, such as *Drosophila* Dorsal (Schloop *et al.*, 2020), the DCC controls hundreds of genes through its action on *cis*-acting target sites. However, the DCC acts in a single developmental process, the control of X-chromosome dosage compensation, while the more typical developmental regulators participate in multiple, unrelated developmental processes. For such multi-purpose proteins, target site specificity is evolutionarily constrained: protein function is changed far more by changes in the number and location of conserved *cis*-acting target sequences than by changes in the target sequences themselves (Carroll, 2008; Nitta *et al.*, 2015). Hence, it was of considerable interest to know whether the nematode DCC complex with multiple targets but lacking the constraints of such pleiotropy, would exhibit a different pattern of evolution. Our results show that it does: clearly the target specificity of the DCC has changed dramatically over the 15-30 MYR that separate *C. elegans* and *C. briggsae*. This divergence of X-chromosome target sites could have been an important factor for nematode speciation.

## Materials and methods

### Procedures for mutant isolation

Procedures for *sdc-2* mutant isolation were described previously in (Wood *et al.*, 2011). *xol-1(y430)*, *dpy-27(y436)*, and *mix-1(y435)* were isolated from a *C. briggsae* deletion library provided by E. Haag using primers listed in Supplementary File 2. Resulting strains are listed in Supplementary File 1.

### Protein Sequence Alignment of SDC-2

Sequence alignment of the *Caenorhabditis elegans* SDC-2 (UniProtKB G5EBL3) and *Caenorhabditis briggsae* SDC-2 (Uniprot A8XQT3) was generated using Clustal Omega (Madeira *et al.*, 2022) and ESPript 3.0 server (https://espript.ibcp.fr) (Robert and Gouet, 2014). The coiled-coil annotations were predicted using the web server version of DeepCoil (Ludwiczak *et al.*, 2019), part of the MPI Bioinformatics Toolkit (Zimmermann *et al.*, 2018; Gabler *et al.*, 2020).

### Preparation of FISH probes

Chromosome FISH probes were prepared from 1 mg of total DNA, which included multiple *C. briggsae* BACs listed in Supplementary File 3 (BACPAC Resources Center, CHORI, Oakland, CA). BACs were purified using the QIAGEN midiprep kit (catalog number 12243). Chromosomal FISH probes were made with the Invitrogen DNA FISH-tag kit. X-chromosome probes (10 BACS covering approximately 5% of the chromosome) were labeled with AlexaFluor 594 (Molecular Probes, F32949), and chromosome III probes (3 BACS covering approximately 1% of the chromosome) were labeled with AlexaFluor 488 (Molecular Probes, F32947).

### Preparation of gut nuclei for FISH and immunofluorescence

Adult worms were dissected in 4 µl egg buffer (25 mM HEPES, pH 7.4, 118 mM NaCl, 48 mM KCl, 0.2 mM $CaCl_2$, 0.2 mM $MgCl_2$) on a 18 mm X 18 mm coverslip. Four µl of 4% formaldehyde (in egg buffer) were added, and the solution was mixed by tapping the coverslip before it was placed onto a Superfrost/Plus glass slide (Fisherbrand, 12-550-15). Fixed samples were incubated for 5 min at room temperature in a humid chamber, then frozen in liquid nitrogen for at least 1 min. Coverslips were

removed quickly with a razor blade, and slides were placed immediately into PBS-T (PBS with 1mM EDTA and 0.5% Triton X-100). Slides were subjected to three 10-min washes in PBS-T at room temperature. Slides were dehydrated in 95% ethanol for 10 min at room temperature followed by either the FISH or immunofluorescence protocol below.

### FISH

Following dehydration of the slides, excess ethanol was removed, 15 µl of hybridization solution (50% formamide, 3X SSC, 10% dextran sulfate, 10 ng labeled DNA probe in water) were added, and a coverslip was placed on each slide. Slides were placed into a slide chamber, and the FISH incubation protocol was conducted in a PCR machine overnight (80 ºC for 10 min, 0.5 ºC/second to 50 ºC, 50 ºC for 1 min, 0.5 ºC/second to 45 ºC, 45 ºC for 1 min, 0.5 ºC/second to 40 ºC, 40 ºC for 1 min, 0.5 ºC/second to 38 ºC, 38 ºC for 1 min, 0.5 ºC/second to 37 ºC, 37 ºC overnight). After overnight incubation at 37 °C, slides were washed at 39 °C using the following regime: three times (15 min each) in 2X SSC (0.3 M NaCl and 30 mM $Na_3C_6H_5O_7$) in 50% formamide, three times (10 min each) in 2X SSC in 25% formamide, three times (10 min each) in 2X SSC, and three times (1 min each) in 1X SSC. Samples were incubated in PBS-T for 10 min at room temperature, and immunofluorescence staining was performed as described below.

### Immunofluorescence of gut nuclei

Following dehydration of slides subjected to immunofluorescence only or to PBS-T treatment (after FISH protocol), the excess liquid was removed (either ethanol from dehydration step, or PBS-T from FISH protocol) and 20 µl of affinity-purified primary antibodies (*Cbr*-DPY-27 and *Cbr*-MIX-1 peptide antibodies [Covance, Inc.]) in PBS-T were added at 1:200 dilution. Samples were incubated in a humid chamber for between 4h and overnight. Slides were washed three times (10 min each) in PBS-T at room temperature and then incubated in secondary antibodies for 3-6 h. Slides were washed three times (10 min each) in PBS-T at room temperature before Prolong (Molecular Probes, P36934) with DAPI (1 µg/ml) was added, and the samples were imaged using a Leica TCS SP2 AOBS. Antibodies used: anti-DPY-27 rabbit antibody raised to *Cbr*-DPY-27 C-terminal peptide

DVQSEAPSAGRPVETDREGSYTNFD, anti-DPY-27 guinea pig antibody raised to the same *Cbr*-DPY-27 peptide, anti-MIX-1 rabbit antibody raised to *Cbr*-MIX-1 C-terminal peptide EATKKPSKKSAKKAVQNTDDEME, Alexa Flour 488 goat anti-rabbit antibody (Molecular Probes, A11034), Alexa Flour 488 goat anti-guinea pig antibody (Molecular Probes, A11073), and Alexa Flour 594 goat anti-rabbit antibody (Molecular Probes, A11037).

### Immunofluorescence of embryos

Embryos were picked into 4 µl of water on poly-lysine-treated slides. After adding a coverslip, slides were frozen in liquid nitrogen for at least 1 min. Coverslips were removed rapidly with a razor blade and samples were dehydrated in 95% ethanol for 10 min. Next, 40 µl of fix solution (2% paraformaldehyde in egg buffer) were added and slides were incubated in a humid chamber for 10 min. Slides were washed three times (10 min each) in PBS-T at room temperature. Antibody staining was performed as described above for gut nuclei.

### Immunoprecipitation analysis

Immunoprecipitation with DPY-27 antibodies followed by MALDI mass spectrometry of trypsinized protein bands excised from an SDS-PAGE gel was performed according to (Mets and Meyer, 2009).

### Calculation of viability for *C. briggsae sdc-2* mutants

XX animals: *sdc-2* (X) / + hermaphrodites were crossed to JU935 males, which carry a *gfp* transgene integrated on the X chromosome, and the hermaphrodite cross progeny (*sdc-2* + / + *gfp*) were moved to individual plates. Three classes of genotype are expected among the self-progeny of *sdc-2* + / + *gfp* hermaphrodites. Two classes, (+ *gfp* / + *gfp* and *sdc-2* + / + *gfp*) express GFP, whereas the third class, (*sdc-2* + / *sdc-2* +) does not. If -*sdc-2* + / *sdc-2* + animals are 100% viable, the expected proportion of non-green animals among the self-progeny of *sdc-2* + / + *gfp* hermaphrodites is 25%. In each case, the expected number of viable non-green adult progeny is shown in parentheses, and the observed proportion is depicted in the chart as a percentage of the expected number. Wild-type XX viability was calculated among the self-progeny of + + / + *gfp* animals.

XO animals: *sdc-2* + / + *gfp* hermaphrodites were crossed with + + / O (wild-type) males. Successfully mated hermaphrodites were identified by the presence of a copulatory plug and then moved to individual plates. Two classes of genotype were expected among the progeny of this cross. One class (+ *gfp* / O) expresses GFP, whereas the other (*sdc-2* + / O) does not. If *sdc-2* + / O animals are 100% viable, the expected proportion of non-green animals among the male progeny is 50%. In each case, the expected number of non-green animals is shown in parentheses, and the observed proportion is depicted in the chart as a percentage of the expected number. Wild-type XO viability was calculated among the male cross-progeny of + + / + *gfp* hermaphrodites and + + / O males.

### Calculation for rescue of *xol-1* XO-specific lethality in *C. briggsae* by an *sdc-2* mutation

The percent viability of wild-type XO animals and mutant XO animals carrying combinations of *xol-1* and *sdc-2* mutations was calculated by formulae that follow. For wild-type XO or *xol-1(y430)* XO progeny from crosses of wild-type or *xol-1(y430)* hermaphrodites mated with wild-type males, the formula is [(number of F1 males)/(total F1 progeny/2)] x 100, a calculation that assumes successful mating and the potential for 50% male cross progeny among the F1. For *xol-1 sdc-2* XO double mutants, *xol-1 -sdc-2 / xol-1* hermaphrodites were mated with wild-type males. Given that *xol-1* XO progeny are inviable, *xol-1 sdc-2* F1 males should make up 1/3 of viable F1s. Thus, % XO rescue is calculated as [(number of males)/(total progeny/3)] x 100.

### Genome Editing Using CRISPR-Cas9

The *Cbr rex-1* (Figure 8), *Cbr rex-3* (Figure 10), *Cbr rex-4* (Figure 9) and *Cbr rex-7* (Figure 11) mutations, as well as *Cel* site 2 insertions (Figure 5) and substitutions of *Cbr* MEX motifs into *Cel rex-33* and substitution of *Cbr* MEX II motifs into *Cel rex-39* (Figure 12) were made with the CRISPR-Cas9 co-conversion technique using Cas9 RNP injections and species-appropriate co-injection markers (Farboud *et al.*, 2019). *C. elegans* editing utilized the *dpy-10* roller marker, and *C. briggsae* editing utilized the *ben-1* marker. The tracrRNA and crRNA guides (Dharmacon) were resuspended in 600 μM of nuclease-free water (Ambion AM9937). The Cas9 RNP mixture for injections included 5 μl Cas9 protein (UC Berkeley QB3 MacroLab, 10 mg/ml), 1.15 μl 2M HEPES, pH 7.5, 0.35 μl 0.5 M KCl, 0.5 μl

600 μM *dpy-10* crRNA, 1 μl target crRNA (Table S3), 5μl tracrRNA, and 7 μl nuclease-free water. The Cas9 RNP mix was incubated at 37 °C for 15 min, and 1 μl of the resulting Cas9 RNP mix was combined with 0.5 μl 10 μM *dpy-10* repair oligo (IDT), 0.5 μl 10 μM *rex* repair oligo (IDT), and 8 μl nuclease-free water. After centrifuging at 16,100 x g for 10 min, the Cas9 RNP mix was injected into gonads of adult hermaphrodites. The target-specific sequences for Cas9 guide RNAs are listed in Supplementary File 4. The DNA sequences for the repair templates are listed in Supplementary File 5.

For *C. elegans*, injected adults were placed on NGM plates. After 3 days of growth at 25 °C, progeny with the roller phenotype were picked to individual plates and allowed to lay embryos. The roller parents were picked into lysis buffer, and the edited site was amplified and sequenced to identify the worms that were edited. The homozygous progeny from properly edited worms were backcrossed twice to wild-type (N2) worms before usage in experiments. For *C. briggsae*, mutants were isolated as published (Farboud *et al.*, 2019). The homozygous progeny from those were backcrossed twice to AF16 worms before usage in experiments. Primers used for genotyping are listed in Supplementary File 2.

### *C. briggsae* ChIP Extract Preparation

Mixed-stage animals were grown on MYOB agar plates with concentrated HB101 bacteria at 20 °C. Animals were cross-linked with 2% formaldehyde for 10 min and quenched with 100 mM Tris-HCl, pH 7.5. Cross-linked animals were resuspended in 1 ml of FA Buffer (150 mM NaCl, 50 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM DTT, and protease inhibitor cocktail [Calbiochem, #539134]) for every 1 gram of animals. This mixture was frozen in liquid nitrogen and then ground under liquid nitrogen by mortar and pestle for 3 min. Once thawed, the mixture was then homogenized with 50 strokes in a Dounce homogenizer. The chromatin was sheared using the Covaris S2 (20% duty factor, power level 8, 200 cycles per burst) for a total of 30 min processing time (60 sec ON, 45 sec OFF, 30 cycles). The concentration of protein in each extract was quantified using the BCA assay (Thermo Fisher, #23228).

### *C. elegans* ChIP-seq Extract Preparation

Mixed-stage embryos were harvested from hermaphrodites grown on MYOB agar plates with concentrated HB101 bacteria at 20 °C. Embryos were cross-linked with 2% formaldehyde for 10 min and quenched with 100 mM Tris-HCl, pH 7.5. Cross-linked embryos were resuspended in 1 ml of FA Buffer (150 mM NaCl, 50 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM DTT, and protease inhibitor cocktail [Calbiochem, #539134]) for every 1 gram of embryos and homogenized with 50 strokes in a Dounce homogenizer. The chromatin was sheared using the Covaris S2 (20% duty factor, power level 8, 200 cycles per burst) for a total of 30 min processing time (60 sec ON, 45 sec OFF, 30 cycles). The concentration of protein in each extract was quantified using the BCA assay (Thermo Fisher, #23228).

### *C. briggsae* ChIP Reactions

To perform the ChIP reactions, a 50 µl bed volume of protein A Dynabeads (Thermo Fisher, #10001D) was re-suspened in 1 ml of FA Buffer (150 mM NaCl, 50 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM DTT, and protease inhibitor cocktail [Calbiochem, #539134]). The beads were incubated in a microcentrifuge tube with 5 µg of anti-FLAG antibodies (Sigma-Aldrich, #F1804) and 5 µg of rabbit anti-mouse IgG antibodies (Jackson ImmunoResearch, #315-005-003), or 5 µg of mouse IgG (Sigma-Aldrich, #I5381) and 5 µg of rabbit anti-mouse IgG antibodies (Jackson ImmunoResearch, #315-005-003), for 90 min at room temperature. Tubes with incubated beads were placed on a magnetic rack, and the liquid was discarded.

Extracts containing 2 mg of protein ChIPs were increased in volume to 1 ml with FA buffer and then added to each tube of Dynabeads for a for 90-min incubation. The Dynabead-extract mixture was washed at room temperature twice with FA Buffer (150 mM NaCl), once with FA Buffer (1 M NaCl), once with FA Buffer (500 mM NaCl), once with TEL buffer (10 mM Tris-HCl, pH 8.0, 250 mM LiCl, 1% IGEPAL CA-630 [Sigma-Aldrich, #I3021], 1% sodium deoxycholate, 1 mM EDTA), and twice with TE Buffer (10 mM Tris, pH 8.0, 1 mM EDTA). Protein and DNA were eluted with 250 µl of buffer (1% SDS, 250 mM NaCl, 1mM EDTA) at 65 °C for 20 min.

### *C. elegans* ChIP Reactions

To perform the ChIP reactions, a 25 µl bed volume of protein A Dynabeads (Thermo Fisher, #10001D) was re-suspened in 1 ml of FA Buffer (150 mM NaCl, 50 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM DTT, and protease inhibitor cocktail [Calbiochem, #539134]). The beads were incubated in a microcentrifuge tube with 3 µg rabbit anti-SDC-3 (lab stock), or 3 µg rabbit IgG (Jackson Immunoresearch, #301-005-003) for 90 min at room temperature. Tubes with incubated beads were placed on a magnetic rack and liquid was discarded. Protocols for the incubation of extract with beads and elution of protein and DNA from beads were the same as those described for *C. briggsae* ChIP reactions.

### ChIP-seq, Illumina Sequencing, and Data Processing

Sequencing libraries were prepared with the eluted materials from ChIP reactions as published (Zhong *et al.*, 2010) with minor changes: sequencing adapters were obtained from Bioo (NEXTflex), and adapters were ligated using the NEB Quick Ligation Kit (M2200). Libraries were sequenced on the Illumina HiSeq 4000 platforms. After barcode removal, reads were aligned uniquely to the *C. briggsae* CB4 genome using the default settings in Bowtie version 2.3.4.3. To account for read depth, ChIP signal was normalized to the total number of reads that uniquely aligned to the genome.

### *C. elegans* qPCR

To perform qPCR reactions, protein and DNA from a *C. elegans* ChIP reaction or from 50% of a control extract (1 mg protein) were de-crosslinked at 65 °C for at least 4 h with 150 µg/ml Proteinase K (Sigma, #3115887001). DNA from each ChIP reaction or from control extract was isolated using the Qiagen PCR purification kit and diluted to a final volume of 200 µl with (10 mM Tris-HCl, pH 8.5). For quantitative PCR, the immunoprecipitated DNAs were quantified by comparing their threshold cycle to the standard curve from control DNA (10% and 3 serial 10-fold dilutions). For the site 2 insertions, the DCC levels at each inserted *rex* site were calculated for each biological replicate as a ratio of the average DCC level at 5 control *rex* sites (*rex-8*, *rex-16*, *rex-32*, *rex-35*, and *rex-48*). For all experiments

involving endogenous *Cel rex-39* in Figure 12B or involving endogenous *Cel rex-33* in Figure 12E, the DCC levels at each inserted *rex* site were calculated for each biological replicate as a ratio of the average DCC level at 7 control *rex* sites (*rex-8*, *rex-14, rex-16*, *rex-32*, *rex-35*, *rex-36,* and *rex-48*). Primers used for qPCR are listed in Supplementary File 2.

### *C. briggsae* qPCR

To perform the qPCR reactions, protein and DNA from a *C. briggsae* ChIP reaction or from 50% of a control extract (1 mg protein) were de-crosslinked at 65 °C for at least 4 h with 150 μg/ml Proteinase K (Sigma, #3115887001). DNA from each ChIP reaction or from control extract was isolated using the Qiagen PCR purification kit and diluted to a final volume of 400 μl with (10 mM Tris-HCl, pH 8.5). For quantitative PCR, the immunoprecipitated DNAs were quantified by comparing their threshold cycle to the standard curve from control DNA (10% and 3 serial 10-fold dilutions). For the endogenous *rex* site mutations, the DCC levels at each inserted *rex* site were calculated for each biological replicate as a ratio of the average DCC level at 4 control *rex* sites (*rex-1*, *rex-2*, *rex-5*, and *rex-9*). Primers used for qPCR are listed in Supplementary File 2.

### Identification of *C. briggsae* DCC binding motifs

The 500 bp DNA sequence centered on each *C. briggsae* SDC-2 ChIP-seq peak location for the 12 *Cbr rex* sites was isolated from the CB4 reference genome. Motif candidates were obtained by inputting twelve 500 bp sequences onto MEME on the MEME-suite website (Bailey and Elkan, 1994; Bailey *et al.*, 2015). The settings used to identify motif candidates were the classic mode and any number of repetitions (anr). The X:A enrichment was calculated for motif candidates. The two motif candidates enriched on the *Cbr* X chromosomes were named *Cbr* MEX for the 13-bp motif and *Cbr* MEX II for the 30-bp motif (Figure 6).

### X:A Fold Motif Enrichment Calculation

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 *Cbr* MEX motif, *Cbr* MEX II, *Cel* MEX motif, and *Cel* MEX II motif at all positions along each chromosome, as explained in (Fuda *et al.*, 2022). For each threshold value, the number of motifs with ln[P] values less than the value (better match) was summed for X and for autosomes. The number of autosomal motifs was divided by the total number of autosomal base pairs to find the number of motifs per base pair. The number of motifs per base pair of X was calculated similarly. The final X:A ratio was calculated by dividing the motifs per base pair for X by the motifs per base pair for the autosomes.

### *C. elegans* DCC binding assay performed *in vitro*

The *in vitro Cel* DCC binding assays (Figure 12 and Figure 13) were performed as described previously in (Fuda *et al.*, 2022). The sequences for the DNA templates used in the *in vitro* DCC binding assays are listed in Supplementary File 6.

#### Data

The accession number for the ChIP-seq data reported in this paper is GEO GSE214714.

## Acknowledgements

We are grateful to E. Haag and his laboratory for generously providing expertise and reagents to make and screen *C. briggsae* deletion pools, A. Wood for initiating ZFN mutagenesis of *C. briggsae sdc-2*, D. Stalford for figure preparation, T. Cline and laboratory members for valuable discussions, and the QB3 Genomics Facility (RRID:SCR_022170) for DNA sequencing. This work was supported in part by NIH Grant R35 GM131845 (to B.J.M.). B.J.M. is an investigator of the Howard Hughes Medical Institute.

## References


Albritton SE, Kranz AL, Winterkorn LH, Street LA, Ercan S. 2017. Cooperation between a hierarchical set of recruitment sites targets the X chromosome for dosage compensation. *Elife* **6**. doi:10.7554/eLife.23645.

Anderson EC, Frankino PA, Higuchi-Sanabria R, Yang Q, Bian Q, Podshivalova K, Shin A, Kenyon C, Dillin A, Meyer BJ. 2019. X Chromosome Domain Architecture Regulates Caenorhabditis elegans Lifespan but Not Dosage Compensation. *Developmental Cell* **51**:192-207 e196. doi:10.1016/j.devcel.2019.08.004.

Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology* **2**:28-36.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. *Nucleic Acids Research* **43**:W39-49. doi:10.1093/nar/gkv416.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**:25-36. doi:10.1016/j.cell.2008.06.030.

Chan RC, Severson AF, Meyer BJ. 2004. Condensin restructures chromosomes in preparation for meiotic divisions. *Journal of Cell Biology* **167**:613-625. doi:10.1083/jcb.200408061.

Chang CP, Brocchieri L, Shen WF, Largman C, Cleary ML. 1996. Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. *Molecular and Cellular Biology* **16**:1734-1745. doi:10.1128/MCB.16.4.1734.

Chu DS, Dawes HE, Lieb JD, Chan RC, Kuo AF, Meyer BJ. 2002. A molecular link between gene-specific and chromosome-wide transcriptional repression. *Genes and Development* **16**:796-805. doi:10.1101/gad.972702.

Chuang PT, Albertson DG, Meyer BJ. 1994. DPY-27:a chromosome condensation protein homolog that regulates C. elegans dosage compensation through association with the X chromosome. *Cell* **79**:459-474.

Chuang PT, Lieb JD, Meyer BJ. 1996. Sex-specific assembly of a dosage compensation complex on the nematode X chromosome. *Science* **274**:1736-1739.

Csankovszki G, Collette K, Spahl K, Carey J, Snyder M, Petty E, Patel U, Tabuchi T, Liu H, McLeod I, Thompson J, Sarkeshik A, Yates J, Meyer BJ, Hagstrom K. 2009. Three distinct condensin complexes control C. elegans chromosome dynamics. *Current Biology* **19**:9-19. doi:10.1016/j.cub.2008.12.006.

Csankovszki G, McDonel P, Meyer BJ. 2004. Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes. *Science* **303**:1182-1185. doi:10.1126/science.1092938.

Cutter AD. 2008. Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. *Molecular Biology and Evolution* **25**:778-786. doi:10.1093/molbev/msn024.

Cutter AD, Morran LT, Phillips PC. 2019. Males, Outcrossing, and Sexual Selection in Caenorhabditis Nematodes. *Genetics* **213**:27-57. doi:10.1534/genetics.119.300244.

Dawes HE, Berlin DS, Lapidus DM, Nusbaum C, Davis TL, Meyer BJ. 1999. Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science* **284**:1800-1804.

De Kumar B, Darland DC. 2021. The Hox protein conundrum: The "specifics" of DNA binding for Hox proteins and their partners. *Developmental Biology* **477**:284-292. doi:10.1016/j.ydbio.2021.06.002.

Farboud B, Severson AF, Meyer BJ. 2019. Strategies for Efficient Genome Editing Using CRISPR-Cas9. *Genetics* **211**:431-457. doi:10.1534/genetics.118.301775.

Fuda NJ, Brejc K, Kruesi WS, Ralston EJ, Bigley R, Shin A, Okada M, Meyer BJ. 2022. Combinatorial clustering of distinct DNA motifs directs synergistic binding of Caenorhabditis elegans dosage compensation complex to X chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **119**:e2211642119. doi:10.1073/pnas.2211642119.

Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Soding J, Lupas AN, Alva V. 2020. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr Protoc Bioinformatics* **72**:e108. doi:10.1002/cpbi.108.

Haag ES. 2005. The evolution of nematode sex determination: C. elegans as a reference point for comparative biology. *WormBook: The Online Review of C Elegans Biology*:1-14. doi:10.1895/wormbook.1.120.1.

Halder G, Callaerts P, Gehring WJ. 1995. Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. *Science* **267**:1788-1792. doi:10.1126/science.7892602.

Hertz GZ, Stormo GD. 1999. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**:563-577. doi:10.1093/bioinformatics/15.7.563.

Hill RE, Favor J, Hogan BL, Ton CC, Saunders GF, Hanson IM, Prosser J, Jordan T, Hastie ND, van Heyningen V. 1991. Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354**:522-525. doi:10.1038/354522a0.

Jans J, Gladden JM, Ralston EJ, Pickle CS, Michel AH, Pferdehirt RR, Eisen MB, Meyer BJ. 2009. A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. *Genes and Development* **23**:602-618. doi:10.1101/gad.1751109.

Kruesi WS, Core LJ, Waters CT, Lis JT, Meyer BJ. 2013. Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. *Elife* **2**:e00808. doi:10.7554/eLife.00808.

Lieb JD, Albrecht MR, Chuang PT, Meyer BJ. 1998. MIX-1: an essential component of the C. elegans mitotic machinery executes X chromosome dosage compensation. *Cell* **92**:265-277.

Lieb JD, Capowski EE, Meneely P, Meyer BJ. 1996. DPY-26, a link between dosage compensation and meiotic chromosome segregation in the nematode. *Science* **274**:1732-1736.

Loda A, Collombet S, Heard E. 2022. Gene regulation in time and space during X-chromosome inactivation. *Nature Reviews: Molecular Cell Biology*. doi:10.1038/s41580-021-00438-7.

Ludwiczak J, Winski A, Szczepaniak K, Alva V, Dunin-Horkawicz S. 2019. DeepCoil-a fast and accurate prediction of coiled-coil domains in protein sequences. *Bioinformatics* **35**:2790-2795. doi:10.1093/bioinformatics/bty1062.

Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*. doi:10.1093/nar/gkac240.

Malicki J, Schughart K, McGinnis W. 1990. Mouse Hox-2.2 specifies thoracic segmental identity in Drosophila embryos and larvae. *Cell* **63**:961-967. doi:10.1016/0092-8674(90)90499-5.

Meisel RP, Davey T, Son JH, Gerry AC, Shono T, Scott JG. 2016. Is Multifactorial Sex Determination in the House Fly, Musca domestica (L.), Stable Over Time? *Journal of Heredity* **107**:615-625. doi:10.1093/jhered/esw051.

Mets DG, Meyer BJ. 2009. Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure. *Cell* **139**:73-86. doi:10.1016/j.cell.2009.07.035.

Meyer BJ. 2022a. Mechanisms of sex determination and X-chromosome dosage compensation. *Genetics* **220**. doi:10.1093/genetics/iyab197.

-. 2022b. The X chromosome in C. elegans sex determination and dosage compensation. *Current Opinion in Genetics and Development* **74**:101912. doi:10.1016/j.gde.2022.101912.

Miller LM, Plenefisch JD, Casson LP, Meyer BJ. 1988. xol-1: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in C. elegans. *Cell* **55**:167-183.

Nitta KR, Jolma A, Yin Y, Morgunova E, Kivioja T, Akhtar J, Hens K, Toivonen J, Deplancke B, Furlong EE, Taipale J. 2015. Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. *Elife* **4**. doi:10.7554/eLife.04837.

Nusbaum C, Meyer BJ. 1989. The Caenorhabditis elegans gene sdc-2 controls sex determination and dosage compensation in XX animals. *Genetics* **122**:579-593.

Pferdehirt RR, Kruesi WS, Meyer BJ. 2011. An MLL/COMPASS subunit functions in the C. elegans dosage compensation complex to target X chromosomes for transcriptional regulation of gene expression. *Genes Dev* **25**:499-515. doi:10.1101/gad.2016011.

Quiring R, Walldorf U, Kloter U, Gehring WJ. 1994. Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans. *Science* **265**:785-789. doi:10.1126/science.7914031.

Rhind NR, Miller LM, Kopczynski JB, Meyer BJ. 1995. xol-1 acts as an early switch in the C. elegans male/hermaphrodite decision. *Cell* **80**:71-82.

Rieder LE, Jordan WT, 3rd, Larschan EN. 2019. Targeting of the Dosage-Compensated Male X-Chromosome during Early Drosophila Development. *Cell Reports* **29**:4268-4275 e4262. doi:10.1016/j.celrep.2019.11.095.

Rim EY, Clevers H, Nusse R. 2022. The Wnt Pathway: From Signaling Mechanisms to Synthetic Modulators. *Annual Review of Biochemistry* **91**:571-598. doi:10.1146/annurev-biochem-040320-103615.

Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**:W320-324. doi:10.1093/nar/gku316.

Samata M, Akhtar A. 2018. Dosage Compensation of the X Chromosome: A Complex Epigenetic Assignment Involving Chromatin Regulators and Long Noncoding RNAs. *Annual Review of Biochemistry* **87**:323-350. doi:10.1146/annurev-biochem-062917-011816.

Schloop AE, Bandodkar PU, Reeves GT. 2020. Formation, interpretation, and regulation of the Drosophila Dorsal/NF-kappaB gradient. *Current Topics in Developmental Biology* **137**:143-191. doi:10.1016/bs.ctdb.2019.11.007.

Tsai CJ, Mets DG, Albrecht MR, Nix P, Chan A, Meyer BJ. 2008. Meiotic crossover number and distribution are regulated by a dosage compensation protein that resembles a condensin subunit. *Genes and Development* **22**:194-211. doi:10.1101/gad.1618508.

Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ. 2011. Targeted genome editing across species using ZFNs and TALENs. *Science* **333**:307. doi:10.1126/science.1207773.

Yatskevich S, Rhodes J, Nasmyth K. 2019. Organization of Chromosomal DNA by SMC Complexes. *Annual Review of Genetics* **53**:445-482. doi:10.1146/annurev-genet-112618-043633.

Yin D, Schwarz EM, Thomas CG, Felde RL, Korf IF, Cutter AD, Schartner CM, Ralston EJ, Meyer BJ, Haag ES. 2018. Rapid genome shrinkage in a self-fertile nematode reveals sperm competition proteins. *Science* **359**:55-61. doi:10.1126/science.aao0827.

Yin H, Wei C, Lee JT. 2021. Revisiting the consequences of deleting the X inactivation center. *Proceedings of the National Academy of Sciences of the United States of America* **118**. doi:10.1073/pnas.2102683118.

Zhong M, Niu W, Lu ZJ, Sarov M, Murray JI, Janette J, Raha D, Sheaffer KL, Lam HY, Preston E, Slightham C, Hillier LW, Brock T, Agarwal A, Auerbach R, Hyman AA, Gerstein M, Mango SE, Kim SK, Waterston RH, Reinke V, Snyder M. 2010. Genome-wide identification of binding sites defines distinct functions for Caenorhabditis elegans PHA-4/FOXA in development and environmental response. *PLoS Genet* **6**:e1000848. doi:10.1371/journal.pgen.1000848.

Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Soding J, Lupas AN, Alva V. 2018. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology* **430**:2237-2243. doi:10.1016/j.jmb.2017.12.007.

Figure 1

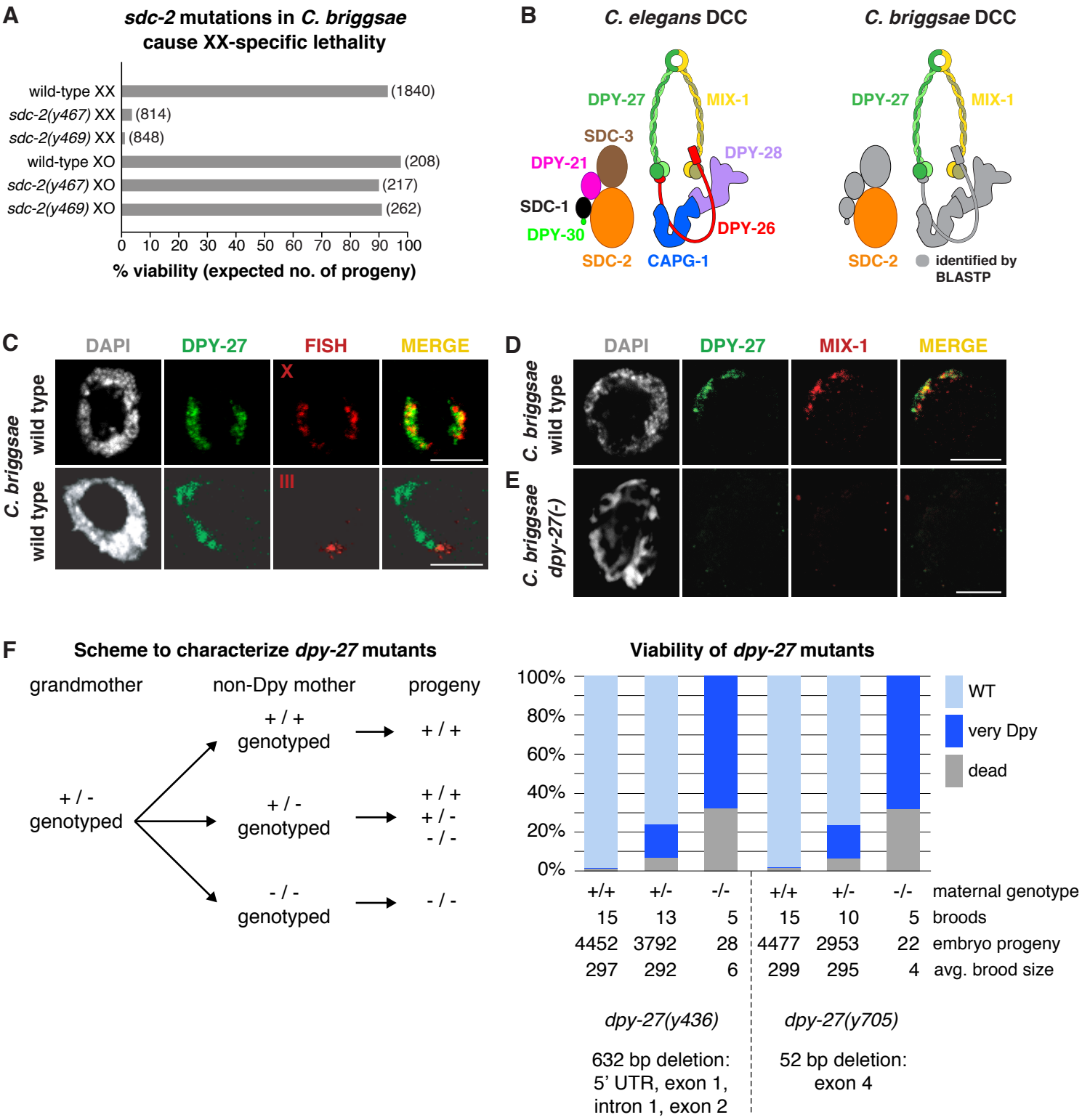

A
sdc-2 mutations in C. briggsae cause XX-specific lethality
wild-type XX (1840)
sdc-2(y467) XX (814)
sdc-2(y469) XX (848)
wild-type XO (208)
sdc-2(y467) XO (217)
sdc-2(y469) XO (262)
0 10 20 30 40 50 60 70 80 90 100
% viability (expected no. of progeny)
B
C. elegans DCC
C. briggsae DCC
DPY-27
MIX-1
SDC-3
DPY-28
DPY-21
SDC-1
DPY-30
DPY-26
SDC-2
CAPG-1
identified by BLASTP
C
DAPI
DPY-27
FISH
MERGE
C. briggsae wild type
X
III
D
MIX-1
E
C. briggsae dpy-27(-)
F
Scheme to characterize dpy-27 mutants
grandmother
non-Dpy mother
progeny
+ / - genotyped
+ / + genotyped
+ / - genotyped
- / - genotyped
+ / +
+ / +
+ / -
- / -
- / -
Viability of dpy-27 mutants
100%
80%
60%
40%
20%
0%
WT
very Dpy
dead
+/+ +/- -/- +/+ +/- -/- maternal genotype
15 13 5 15 10 5 broods
4452 3792 28 4477 2953 22 embryo progeny
297 292 6 299 295 4 avg. brood size
dpy-27(y436)
632 bp deletion: 5' UTR, exon 1, intron 1, exon 2
dpy-27(y705)
52 bp deletion: exon 4

**Figure 1.** Conservation of X-chromosome dosage compensation machinery between *C. briggsae* and *C. elegans*.

(A) *sdc-2* mutations cause XX-specific lethality in *C. briggsae.* Graph shows percent viability of wild-type and *Cbr sdc-2* mutant XX and XO adults. Viability of homozygous XX and hemizygous XO *Cbr sdc-2* mutants is expressed as the percentage of live adults for each karyotype relative to the number expected (shown in parentheses) in the progeny of a cross if all mutant animals were viable. Crosses and calculations are described in Materials and Methods. Sequence changes of *sdc-2* mutations derived from genome editing using zinc-finger nucleases are shown in ***Figure 1—figure supplement 2A***.

(B) Schematic of the *C. elegans* dosage compensation complex (left) and *C. briggsae* orthologs identified by BLASTP (right). The *C. elegans* DCC includes homologs of all core condensin subunits (MIX-1, DPY-27, DPY-26, DPY-28, and CAPG-1). *C. briggsae* DCC components identified and characterized in this study are shown in color; other orthologs are in grey.

(C) Condensin subunits bind X chromosomes and mediate dosage compensation in *C. briggsae.* Confocal images of *C. briggsae* hermaphrodite gut nuclei co-stained with the DNA dye DAPI (grey), antibodies to *Cbr* DPY-27 (green), and FISH probes to either 5% of X (red) or 1% of chromosome III (red) show that *Cbr* DPY-27 co-localizes with X but not III, consistent with a role in dosage compensation. Scale bars, 5 μm.

(D,E) Confocal images of *C. briggsae* gut nuclei from wild-type or *dpy-27* mutant XX adult hermaphrodites co-stained with DAPI (grey) and antibodies to *Cbr* DPY-27 (green) and *Cbr* MIX-1 (red) show that *Cbr* DPY-27 and *Cbr* MIX-1 co-localize on X in wild-type hermaphrodites (**D**). Association of *Cbr* MIX-1 with X is disrupted in *Cbr dpy-27* mutants, in accord with its participation in a protein complex with *Cbr* DPY-27 (E). The *Cbr dpy-27(y436)* allele was recovered by screening a *C. briggsae* deletion library. Scale bars, 5 μm.

(F) Viability of *dpy-27* mutant XX *C. briggsae* animals. The left panel shows the genetic scheme to characterize the effect of maternal genotype on viability of *dpy-27* null XX mutants. Comparison is made between homozygous null *dpy-27* progeny from heterozygous or homozyogus non-Dpy mutant mothers. The genotype of non-DPY mothers was established through PCR analysis. The right panel shows the percent viability of progeny from wild-type hermaphrodites and heterozygous or homozygous *dpy-27* mutant hermaphrodites. The maternal genotype, number of broods, total number of embryo progeny from all broods, and average brood size are provided for two null alleles of *dpy-27*. Molecular characterization of mutations is shown below the graph and in ***Figure 1—figure supplement 2B***. Almost all progeny of *dpy-27* null mutant mothers are dead; a homozygous *dpy-27* null strain cannot be propagated. More than 20% of progeny of *dpy-27*/+ heterozygous mutant mothers are very Dpy or dead, indicating that a wild-type DPY-27 maternal contribution has minimal effect on suppressing the deleterious effect of the homozygous null zygotic genotype. The complete XX lethality is consistent with a major role for condensin subunit DPY-27 in dosage compensation.

# Figure 1—figure supplement 1

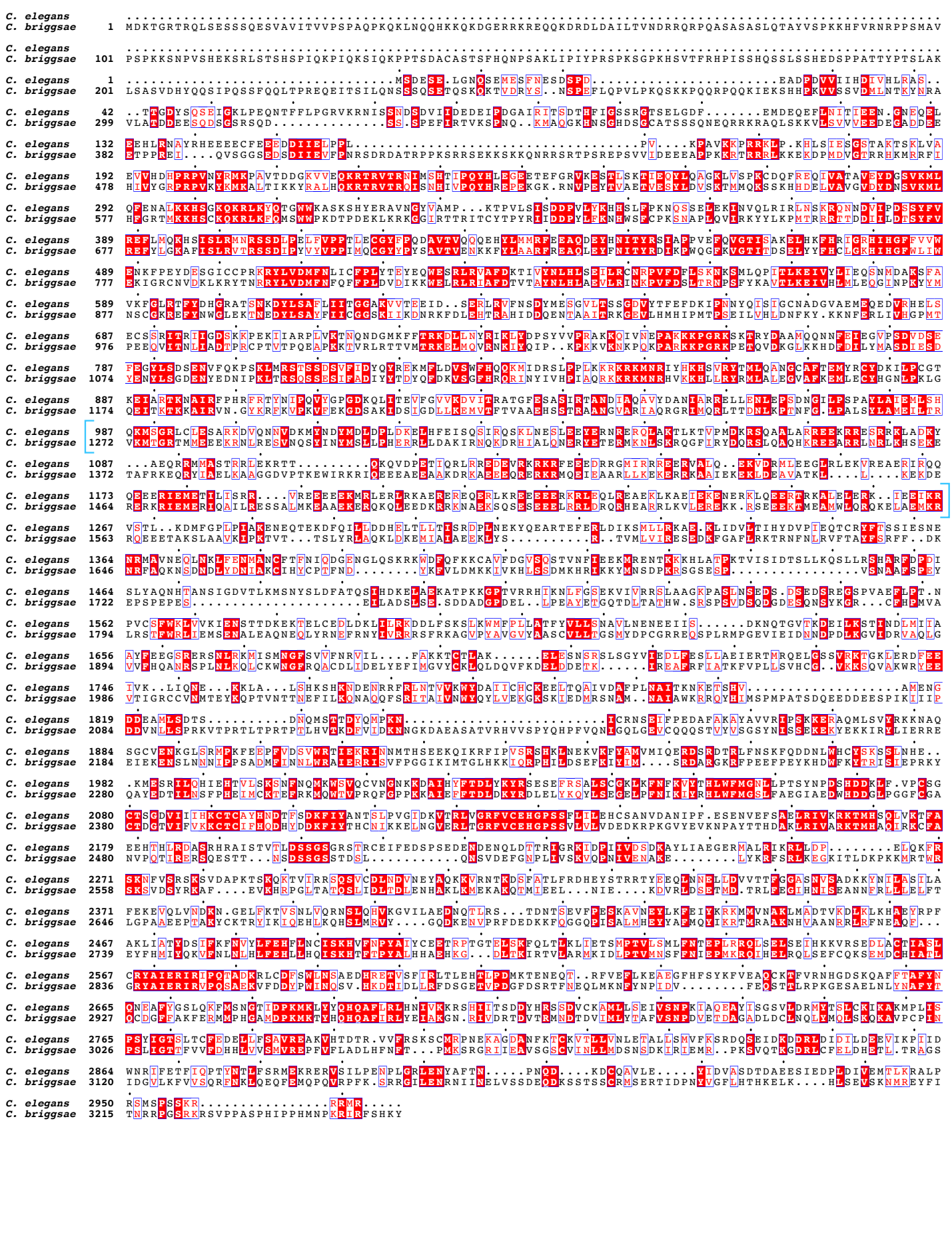


**Figure 1—figure supplement 1.** Protein sequence alignment comparing SDC-2 in *C. elegans* vs. *C. briggsae*. The proteins share 31% identity and 44% similarity. Red background indicates amino acid identity, and red characters demark similarity. The predicted coiled-coil regions for the two proteins are delineated by blue brackets.

**A**

```
                    CGATTTGGAACACACCCGAGCTCACATCGATGATCAAGAAAACACGGCAGCAATCACCAGGAAAGGAGAAG  (WT sdc-2)
                                                   CG
Cbr sdc-2(y467)     CGATTTGGAACACACCCGAGCTCACATCGAT:::::::::AACACGGCAGCAATCACCAGGAAAGGAGAAG  (Δ9, +2)

Cbr sdc-2(y469)     CGATTTGGAACACACC:::::::::::::::::::::::::::GGCAGCAATCACCAGGAAAGGAGAAG  (Δ29)
```

**B**

```
                    ATCGATTTGCTCTACAAGTTGAAGCTCCCGCGCGCATCGTTCTTCGCTCTGGAGTACGGTAAAACGGTAAAC  (WT dpy-27)
                                                                              DSB PAM
Cbr dpy-27(y705)    ATCGATTTGC:::::::::::::::::::::::::::::::::::::::::::::::::AACGGTAAAC  (Δ52)
```

**C**

```
                    GAAGTCGTCCAAACACCATGATGAACTTGTAGCTGTTGGTGTCGATTACGATAACTCGGTGAAGATG  (WT sdc-2)
                                                   CG
Cbr sdc-2(y453)     GAAGTCGTCCAAACACCATGATGAACTTGTA:::::::::::::::::::::::::::::::::::::C  (Δ39, +2)
                                      TATGGTGTATATGGTATA
Cbr sdc-2(y454)     GAAGTCGTCCAAACACCATGATGAACTTGTA:::::TGGTGTCGATTACGATAACTCGGTGAAGATGCTAC  (Δ5, +18)
```

**D**

```
                    CGATTTGGAACACACCCGAGCTCACATCGATGATCAAGAAAACACGGCAGCAATCACCAGGAAAGGA  (WT sdc-2)

Cbr sdc-2(y455)     CGATTTGGAACACACCCGAGCTCACATCGATGATCA::::::::CGGCAGCAATCACCAGGAAAGGAGAAG  (Δ8)
                                      AATCGTTGTCACATCG
Cbr sdc-2(y460)     CGATTTGGAACACACCCGAGCTCACATCGA:::::::::AAACACGGCAGCAATCACCAGGAAAGGAGAAG  (Δ9, +16)
```

**Figure 1—figure supplement 2.** DNA sequence changes mediated by genome editing.
(A) DNA sequences of mutant *Cbr sdc-2* alleles that were created by genome editing using zinc-finger nucleases, as described in Wood *et al.* 2011. Mutations include short insertions (green) and deletions (red colons) that generate in-frame deletions and frame-shift mutations. Inserted sequences (green) frequently share homology (underlined in green) with sequences flanking the break site, as is typical of NHEJ-mediated repair.

The deletions in both *sdc-2(y467)* and *sdc-2(y469)* create premature translation stop codons, thereby preventing formation of full-length SDC-2 proteins and causing complete loss of gene function. For *y467*, the wild-type sequence ends at codon 926-Asp. The deletion and insertion cause 18 incorrect amino acids to be translated, and a stop codon occurs in place of codon 945. For *y469*, the wild-type sequence ends at codon 921-Thr. The deletion causes 26 incorrect amino acids to be translated, and a stop codon occurs in place of codon 948.

(B) DNA sequence of mutant *Cbr dpy-27(y705)* allele created by genome editing using CRISPR/Cas9. The Cas9 target sequence was 5' CGCTCTGGAGTACGGTAAAA 3'. The PAM is the CGC (red) immediately 3' of the target sequence. The double strand break (DSB) site is indicated by a blue line. The mutation is a 52 bp deletion (red colons) in exon 4 that creates a premature translation stop codon and prevents formation of the full-length DPY-27 protein. The deletion starts at codon 689, and the in-frame stop codon is 2 codons past the 3' end of the deletion.

(C, D) DNA sequences of mutant *Cbr sdc-2* alleles that were obtained as suppressors of the XO-specific lethality caused by a *xol-1* mutation. Alleles *sdc-2(y453)*, *sdc-2(y454)*, *sdc-2(y455)*, and *sdc-2(y460)*, were created by genome editing using zinc-finger nucleases, as described in Wood *et al.* 2011.

The mutations in both *sdc-2(y453)* and *sdc-2(y454)* create premature translation stop codons, thereby preventing formation of full-length SDC-2 proteins and causing complete loss of gene function. For *sdc-2(y453)*, the wild-type sequence ends at codon 563-Val. The deletion and insertion cause 6 incorrect amino acids to be translated, and a stop codon occurs in place of codon 570 (554). For *sdc-2(y454)*, the wild-type sequence ends at codon 563-Val. The deletion and insertion cause 11 incorrect amino acids to be translated, and a stop codon occurs in place of codon 575.

The mutations in both *sdc-2(y455)* and *sdc-2(y460)* create premature translation stop codons, thereby preventing formation of full-length SDC-2 proteins. For *sdc-2(y455)*, the wild-type sequence ends at codon 927-His. The deletion causes 27 incorrect amino acids to be translated, and a stop codon occurs in place of codon 955. For *sdc-2(y460)*, the wild-type sequence ends at codon 925-Ile. The deletion and insertion cause 34 incorrect amino acids to be translated, and a stop codon occurs in place of codon 960.

**Figure 1—table supplement 1.** MALDI-TOF identification of *Cbr* MIX-1 peptides

| m/z Submitted | $MH^+$ Matched | Delta ppm | Peptide | Missed Cleavage | Database Sequence |
|---|---|---|---|---|---|
| 916.47 | 916.46 | 9.5 | 674-680 | 0 | (K)YHENVVR(L) |
| 1163.59 | 1163.58 | 3.3 | 375-384 | 1 | (K)LRGELEGMSR(G) |
| 1214.65 | 1214.66 | -3.6 | 631-641 | 0 | (R)VLIESQCLPGR(R) |
| 1224.63 | 1224.62 | 8.8 | 713-723 | 1 | (R)EVAYTDGVKSR(T) |
| 1263.74 | 1263.74 | -0.87 | 524-534 | 0 | (R)DVEGLVLHLIR(L) |
| 1285.69 | 1285.69 | -2.8 | 631-641 | 0 | (R)VLIESQCLPGR(R) |
| 1350.69 | 1350.70 | -8.9 | 656-666 | 0 | (R)YTIINDQSLQR(A) |
| 1881.97 | 1881.98 | -2.3 | 134-150 | 0 | (R)GVGLNVNNPHFLIMQGR(I) |
| 1886.89 | 1886.91 | -6.8 | 86-101 | 0 | (K)QSPFGMDHLDELVVQR(H) |
| 2064.01 | 2064.00 | 3.4 | 460-477 | 0 | (K)ITQQVQSLGYNADEDVQR(R) |
| 2377.18 | 2377.16 | 5.6 | 385-415 | 1 | (R)GTVTNDKGEHVSLETYIQETR(A) |

This table lists the mass-to-charge ratio (m/z) of measured peptides, the predicted masses ($MH^+$ Matched), and the deviation from predicted masses (Delta ppm). The ID of each measured peptide is described by the residue range within full-length MIX-1 (Peptide) and its corresponding amino acid sequence (Database Sequence). The number of uncut tryptic peptide bonds is listed for each peptide (Missed Cleavage).

Figure 2

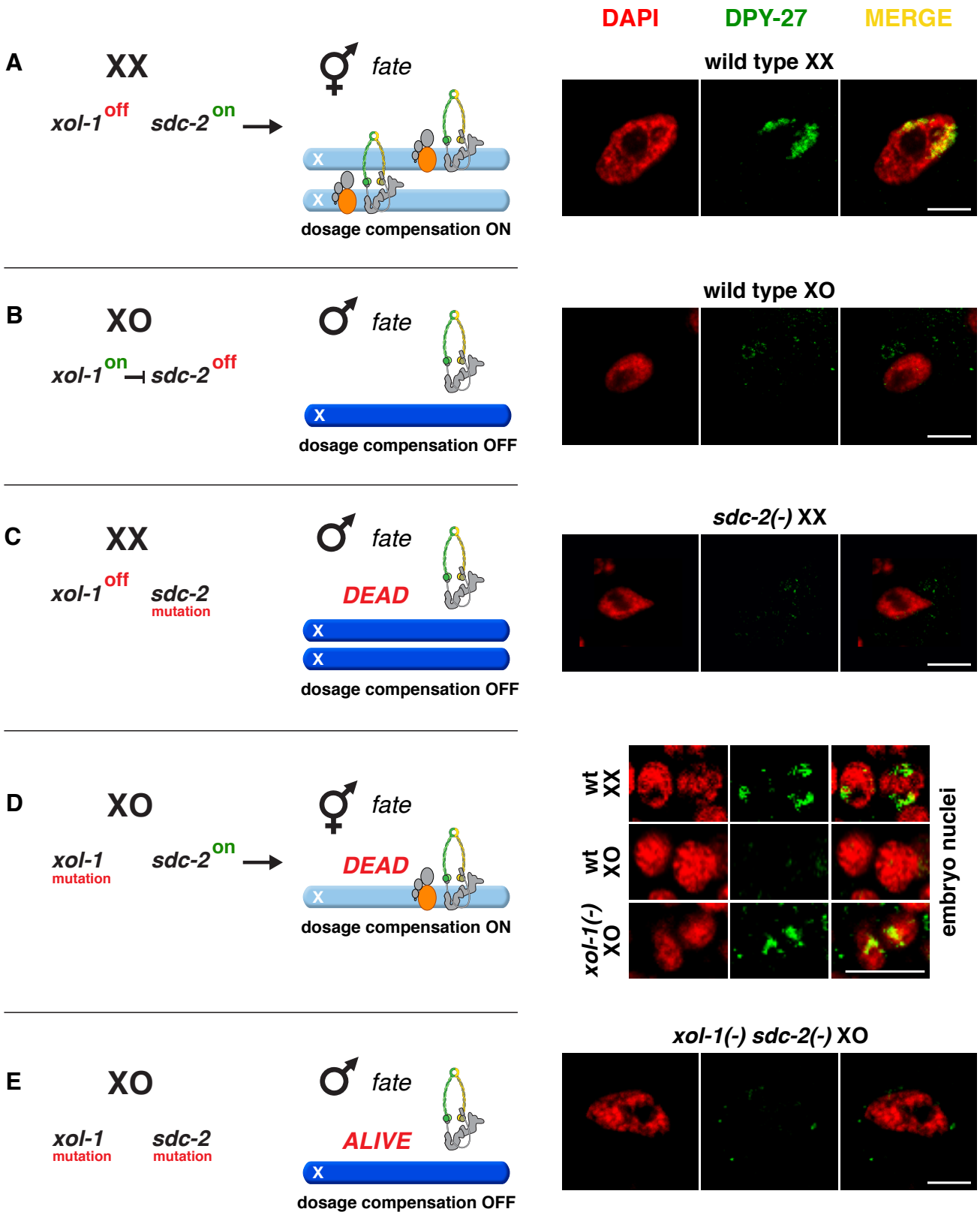

DAPI
DPY-27
MERGE
A
XX
xol-1 off
sdc-2 on
fate
X
X
dosage compensation ON
wild type XX
B
XO
xol-1 on
sdc-2 off
fate
X
dosage compensation OFF
wild type XO
C
XX
xol-1 off
sdc-2
mutation
fate
DEAD
X
X
dosage compensation OFF
sdc-2(-) XX
D
XO
xol-1
mutation
sdc-2 on
fate
DEAD
X
dosage compensation ON
wt XX
wt XO
xol-1(-) XO
embryo nuclei
E
XO
xol-1
mutation
sdc-2
mutation
fate
ALIVE
X
dosage compensation OFF
xol-1(-) sdc-2(-) XO

**Figure 2.** Conserved genetic hierarchy targets the *C. briggsae* DCC to X chromosomes of hermaphrodites.

(A-E) Schematic depiction of the genetic hierarchy controlling sex-specific DCC recruitment to *C. briggsae* X chromosomes (left) paired with representative immunofluorescence experiments exemplifying DCC localization (right). Scale bars, 5 μm. Gut nuclei (A, B, C, E) or embryos (D) were co-stained with DAPI (red) and antibodies to *Cbr* DPY-27 (green). In wild-type XX, but not XO gut nuclei (A, B), DPY-27 co-localizes with X chromosomes, consistent with a role for condensin subunit DPY-27 in dosage compensation (see also Figure 1C).

(C) SDC-2 is required for recruitment of DPY-27 to X chromosomes of hermaphrodites. Failure of the DCC to bind X chromosomes of *sdc-2* XX mutants underlies the XX-specific lethality. Shown is the gut nucleus of a rare XX *sdc-2* mutant escaper near death. *sdc-2* mutant XX escaper animals are masculinized.

(D) Lethality of *Cbr xol-1(y430)* XO animals corresponds to inappropriate binding of the DCC to the single X in embryos.

(E) Mutation of the DCC recruitment factor *Cbr sdc-2* in a *Cbr xol-1* XO mutant prevents DCC recruitment to X and suppresses the XO lethality. See Figure 3B for quantification.

**Figure 3**

**A**

*xol-1(y430) X*
XX ⚥

X

*mfIs27[Cel lip-1::gfp, Cel myo-2::gfp] X*
XO ♂

overnight

inject ZFN mRNA into XX animals with mating plugs

ZFN mRNA

*xol-1(y430) X*
XX

non-green XO progeny

**DEAD**

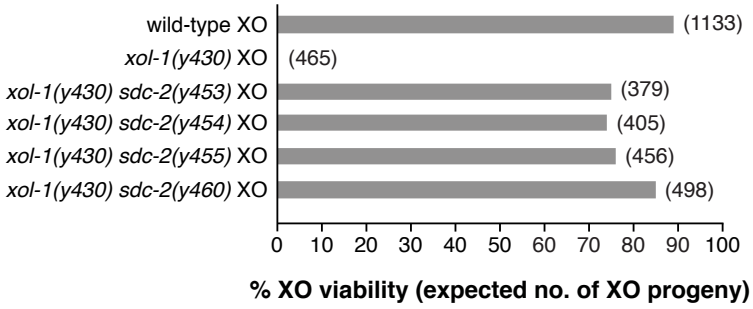
*xol-1(y430) X*
XO

**ALIVE**

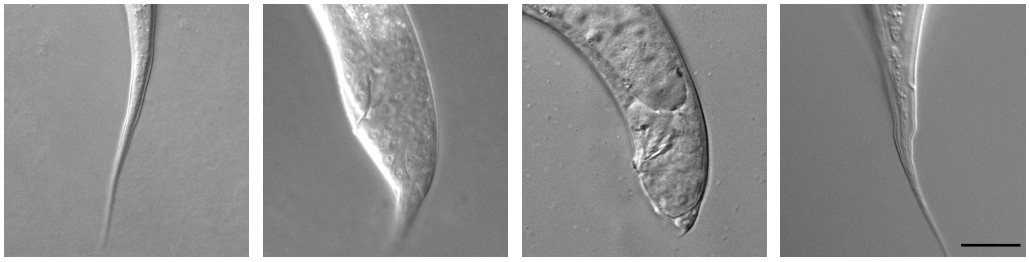
*xol-1(y430) sdc-2 X*
XO

**B**

***sdc-2* mutations in *C. briggsae* rescue XO lethality caused by loss of *xol-1***

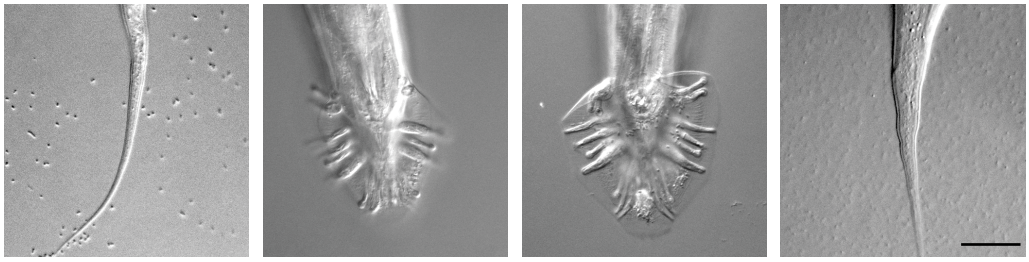



**C**



**wild-type XX L4** | **wild-type XO L4** | ***sdc-2* XX L4** | ***dpy-27* XX L4**

**D**



**wild-type XX adult** | **wild-type XO adult** | ***xol-1 sdc-2* XO adult** | ***dpy-27; xol-1* XO adult**

**Figure 3.** *sdc-2* controls dosage compensation and sex determination in *C. briggsae*.

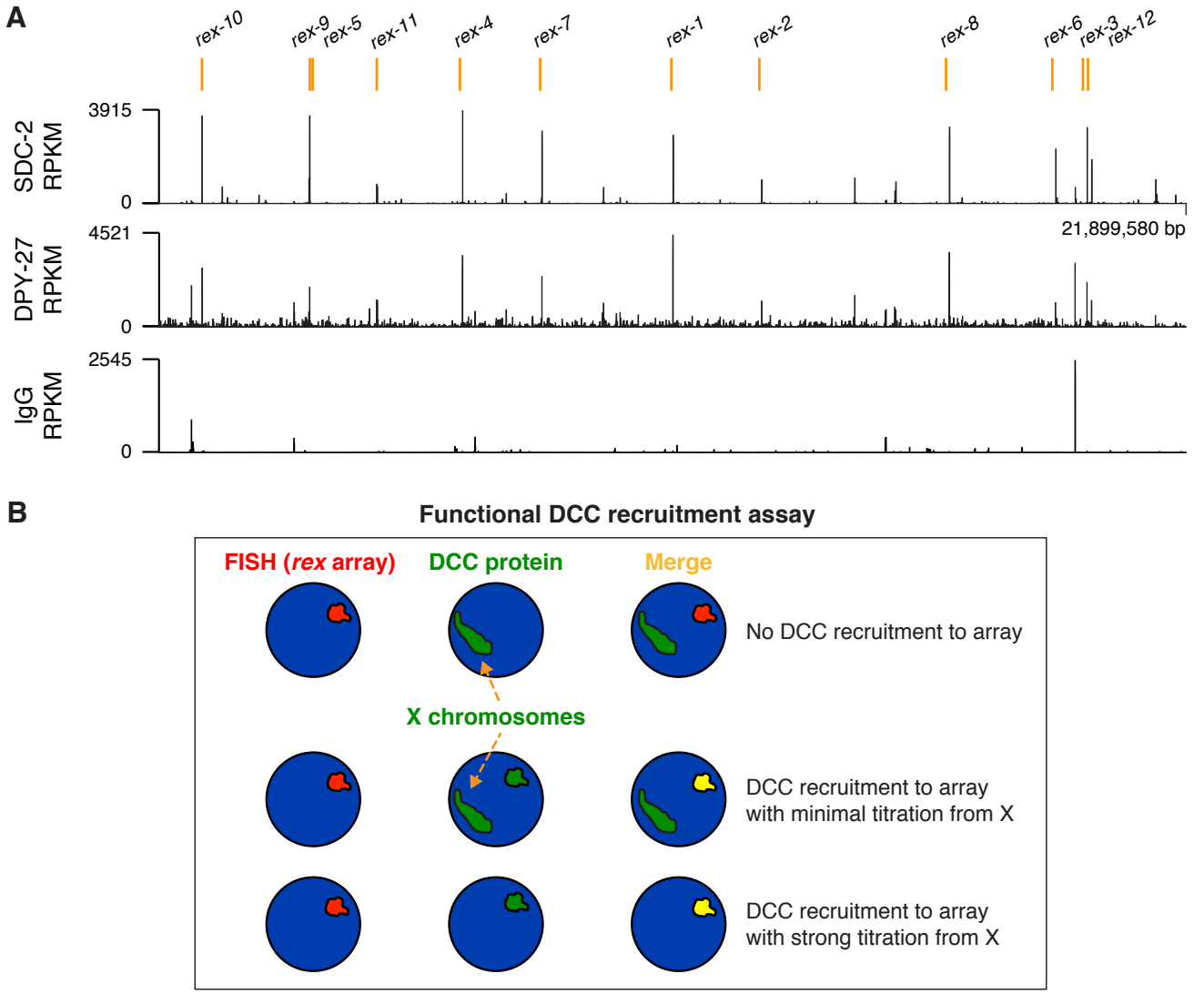
(A) Diagram of the screening strategy to recover *Cbr sdc-2* mutations as suppressors of the XO-specific lethality caused by a *xol-1* mutation. *Cbr xol-1* XX hermaphrodites were mated with males carrying a *gfp*-marked X chromosome to allow F1 XO males to be monitored for the parental origin of the X chromosome. Animals with mating plugs (indicating successful mating) were injected with mRNAs to *sdc-2* zinc-finger nucleases, and all F1 males were examined for GFP fluorescence. Non-green males necessarily inherited an X chromosome carrying a *Cbr-xol-1* mutation and, assuming conservation of the DCC regulatory hierarchy, would be inviable without a concomitant *Cbr sdc-2* mutation. GFP-positive males arose at low frequency from fertilization of nullo-X oocytes (caused by non-disjunction of the maternal X chromosome) with *gfp*-X-bearing sperm. These false positives were discarded from further study.

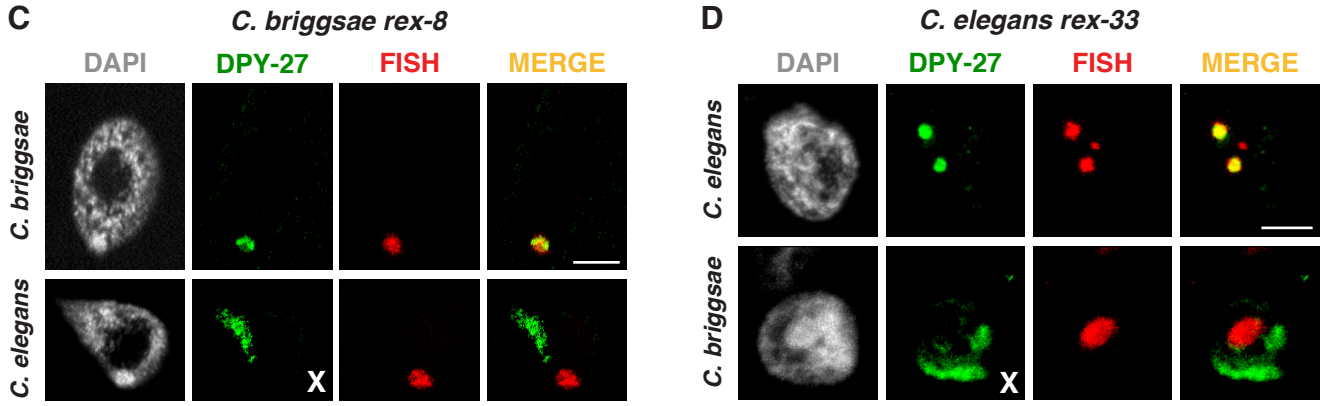
(B) *Cbr sdc-2* mutations rescue *Cbr xol-1(y430)* XO lethality. Graph shows percent viability of wild-type XO animals and mutant XO animals carrying combinations of *Cbr xol-1* and *Cbr sdc-2* mutations. The % XO viability is expressed as the percentage of live XO adults relative to the number expected (shown in parentheses) in the progeny of the cross. Formulae for viability calculations are given in the Materials and Methods. Sequence changes of *sdc-2* mutations are shown in ***Figure 1—figure supplement 2C and 2D***.

(C) *sdc-2* activates the program for C*br* hermaphrodite sexual development. DIC images show comparison of tail morphologies for *Cbr* L4 animals of different genotypes. *sdc-2* mutations, but not *dpy-27* mutations, cause masculinization of XX animals. Scale bar, 20 μm.

(D) DIC images show tail morphologies of wild-type or doubly mutant *Cbr* adults. An *sdc-2* mutation suppresses both the XO lethality and feminization caused by a *xol-1* mutation, consistent with a role for *sdc-2* in controlling both dosage compensation and sex determination. *xol-1 sdc-2* XO animals are viable, fertile males, indicating that the *sdc-2* mutation suppressed the lethality and feminization caused by *xol-1* mutations in XO animals. A *dpy-27* mutation suppresses the XO lethality but not feminization caused by a *xol-1* mutation, consistent with a role for *dpy-27* in dosage compensation but not sex determination. *dpy-27; xol-1* XO animals are fertile hermaphrodites. Scale bar, 20 μm.

**Figure 4**




**E**

| Site | Site Position on *Cbr* X | % Recruitment *in vivo* (No. of Nuclei) |
|---|---|---|
| *Cbr rex-1* | 10,780,533 | 92% (59) |
| *Cbr rex-3* | 19,468,721 | 88% (74) |
| *Cbr rex-4* | 6,358,591 | 85% (68) |
| *Cbr rex-7* | 8,026,460 | 97% (65) |
| flat 1 | 12,489,156 | 0% (181) |

**Figure 4.** Identification of *C. briggsae* DCC recruitment elements on X.

(A) ChIP-seq profiles of *Cbr* SDC-2 and *Cbr* DPY-27 binding to X chromosomes. ChIP-seq experiments were performed using an anti-FLAG antibody to immunoprecipitate SDC-2 from a strain encoding FLAG-tagged SDC-2. An anti-FLAG antibody was also used in ChIP-seq experiments to immunoprecipitate DPY-27 from a strain encoding FLAG-tagged DPY-27. The control IgG ChIP-seq profile on X is also shown. Peaks that correspond to recruitment elements on X (*rex* sites), as determined by the assay in (B), are indicated in blue above the ChIP-seq profiles. RPKM is the abbreviation for reads per kilobase per million reads mapped.

(B) Assay performed *in vivo* to determine whether DNAs from ChIP-seq peaks recruit the DCC when detached from X. XX embryos carrying extrachromosomal arrays with multiple copies of DNA from a ChIP-seq peak in (A) were stained with a DNA FISH probe to the array (red) and DPY-27 antibody (green). If the DNA from a peak failed to recruit the DCC, DPY-27 staining would identify X chromosomes but not the array. If DNA from a peak encoded a recruitment site (*rex* site), DPY-27 staining would co-localize with the array and the X chromosome. In the merged image, the array would appear yellow and the X chromosome would appear green. Often, an array carries enough copies of a *rex* site that it titrates most of the DCC from X, and only the array itself shows evidence of DCC binding, appearing yellow in the merged image. In that case, the X chromosome is not detectable by DPY-27 antibody staining. XX strains carrying *rex* arrays that titrate the DCC from X cannot be propagated due to the defect in dosage compensation cause by DCC titration.

(C) *C. briggsae rex* sites recruit the *C. briggsae* DCC but not the *C. elegans* DCC. Shown is a *C. briggsae* or *C. elegans* XX gut nucleus carrying an extrachromosomal array containing multiple copies of the *C. briggsae* DCC recruitment site *rex-8*. Nuclei were stained with appropriate species-specific *C. briggsae* or *C. elegans* antibodies to the DCC subunit DPY-27 (green), DAPI (grey), and an array FISH probe (red). In *C. briggsae*, DPY-27 bound to arrays in about 40% of the 52 scored nuclei carrying a *Cbr rex-8* array, and the DCC was titrated from X. In *C. elegans*, the DPY-27 bound to arrays in 0% of the 27 scored nuclei carrying a *Cbr rex-8* array, and DPY-27 binding to the *C. elegans* X was evident. Scale bar, 5 μm.

(D) *C. elegans rex* sites do not recruit the *C. briggsae* DCC. Shown is a *C. elegans* or *C. briggsae* XX gut nucleus carrying an extrachromosomal array containing multiple copies of the *C. elegans* recruitment site *rex-33* with 3 MEX motifs (ln[P] scores of -13.13, -15.33, -15.35). Nuclei were stained with *C. elegans* or *C. briggsae* antibodies to DCC subunit DPY-27 (green), DAPI (grey), and an array FISH probe (red). In *C. elegans*, DPY-27 bound to arrays in 100% of the 63 scored nuclei carrying a *Cel rex-33* array, and the DCC was titrated from X. In *C. briggsae*, DPY-27 bound to arrays in 0% of the 53 scored nuclei carrying a *Cel rex-33* array, but did bind to *Cbr* X chromosomes in the same nuclei (***Figure 4—table supplement 1***). Scale bar, 5 μm.

(E) Quantification of exemplary *Cbr* recruitment assays *in vivo* using extrachromosomal arrays containing multiple copies of DNA from *Cbr* DCC ChIP-seq peaks that define *rex* sites. Data are shown for DPY-27 recruitment to DNA from four strong *Cbr* ChIP-seq peaks and a control region of DNA lacking a DCC peak (flat 1 containing the gene *mom-1*). Shown are the locations of the sites on X, the total number of embryonic nuclei scored for DPY-27 recruitment to the array, and the percent of nuclei recruiting the DCC. Arrays carrying *rex* sites recruit the DCC but arrays carrying the control flat region fail to recruit the DCC. Results of DCC recruitment assays *in vivo* for all *rex* sites are presented in ***Figure 4—table supplement 1***.

**Figure 4—table supplement 1.** Results of DCC recruitment assays *in vivo*

**A.** *C. briggsae* DCC binds *C. briggsae* DCC recruitment sites

| ***Cbr rex* Site** | ***Cbr* Chr X Peak Position** | ***Cbr* SDC-2 RPKM** | ***Cbr* Array Assay *in vivo* % Recruitment (No. of Nuclei)** |
|---|---|---|---|
| *rex-1* | 10,780,533 | 2890 | 92% (59) |
| *rex-2* | 12,642,866 | 999 | 90% (101) |
| *rex-3* | 19,468,721 | 3219 | 88% (74) |
| *rex-4* | 6,358,591 | 3915 | 85% (68) |
| *rex-5* | 3,153,011 | 3562 | 98% (45) |
| *rex-6* | 18,811,390 | 2203 | 74% (68) |
| *rex-7* | 8,026,460 | 2964 | 97% (65) |
| *rex-8* | 16,578,214 | 3217 | 37% (52) |
| *rex-9* | 3,135,562 | 1029 | 85% (62) |
| *rex-10* | 895,450 | 3605 | 80% (55) |
| *rex-11* | 4,563,250 | 830 | 89% (54) |
| *rex-12* | 19,564,937 | 1786 | 79% (77) |
| flat 2 | 11,762,995 | 2890 | 6% (48) |
| flat 3 | 20,918,257 | 999 | 0% (144) |

Extrachromosomal arrays composed of DNA fragments (2 kb) that were PCR-amplified from *C. briggsae* X chromosome regions corresponding to Cbr SDC-2 ChIP-seq peaks were tested for their ability to recruit the *Cbr* DCC. Gut nuclei from *C. briggsae* transgenic lines were scored for the presence of the array using a FISH probe against the *myo-2::gfp* vector and the presence or absence of DCC binding to the array by immunofluorescence signal using *Cbr* DPY-27 antibodies. The % recruitment is the percentage of total scored array-bearing nuclei that showed DPY-27 to the array.

**B.** *C. briggsae* DCC does not bind *C. elegans* DCC recruitment sites

| ***Cel rex* Site** | ***Cel* Chr X Peak Position** | ***Cel* Array Assay *in vivo* % Recruitment (No. of Nuclei)** | ***Cbr* Array Assay *in vivo* % Recruitment (No. of Nuclei)** |
|---|---|---|---|
| *rex-4* | 11,522,205 | 100% (16) | 1% (116) |
| *rex-33* | 6,296,501 | 100% (63) | 0% (53) |

Identical DNA fragments encoding individual *C. elegans* DCC recruitment sites *(rex)* were injected into *C. elegans* and *C. briggsae* to create extrachromosomal arrays containing multiple copies of the *rex* site. Gut nuclei from *C. elegans* or *C. briggsae* transgenic lines were scored for the presence of the array using a FISH probe against the *myo-2::gfp* vector and for the presence or absence of DCC binding to the array by immunofluorescence signal from the species-matched DPY-27 antibody. % recruitment is the percentage of total scored array-bearing nuclei that showed DCC binding to the array.

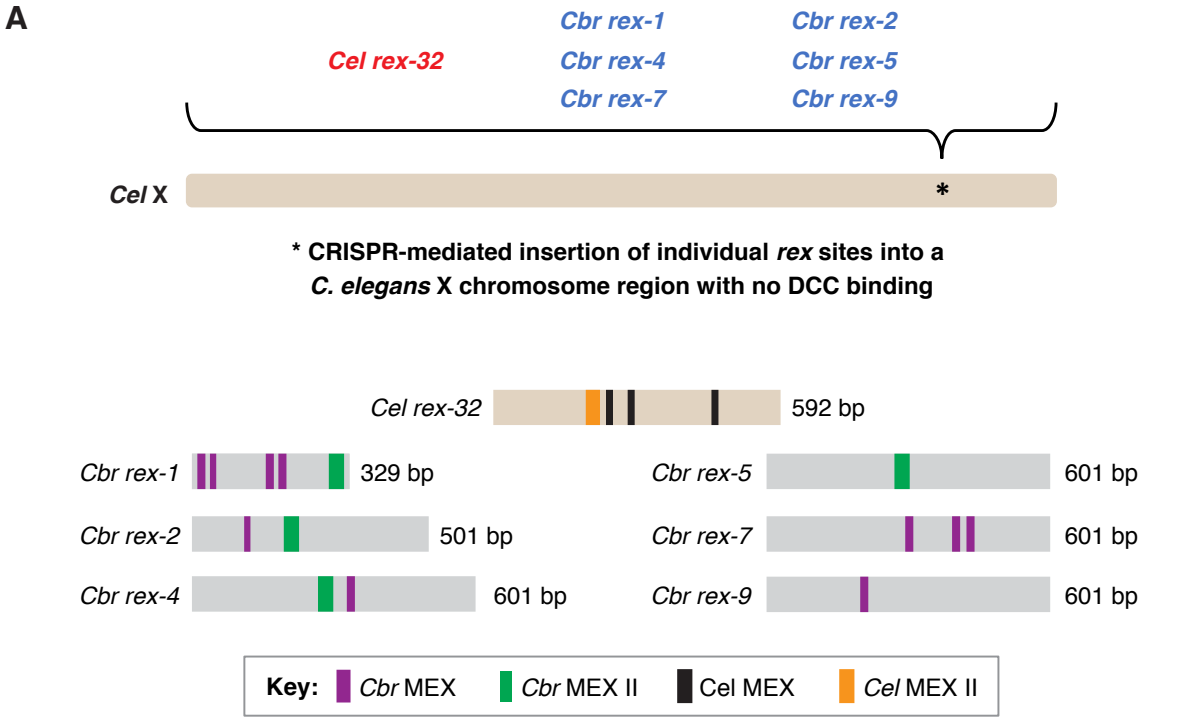
**Table 1.** Motifs within *rex* sites

| ***Cbr rex* Site** | **Chr X Peak Position** | **SDC-2 RPKM** | ***Cbr* MEX motif ln(P) < -12** | ***Cbr* MEX II ln(P) < -12** |
|---|---|---|---|---|
| *rex-1* | 10,780,533 | 2890 | **-15.57 (13 bp) -15.57 (106 bp) -14.63 (14 bp) -14.47 (93 bp)** | **-27.58** |
| *rex-2* | 12,642,866 | 999 | **-14.25 (73 bp)** | **-22.69** |
| *rex-3* | 19,468,721 | 3219 | **-12.36 (178 bp)** | **-20.04** |
| *rex-4* | 6,358,591 | 3915 | **-19.09 (33 bp)** | **-13.80** |
| *rex-5* | 3,153,011 | 3562 | **-18.98** | |
| *rex-6* | 18,811,390 | 2203 | **-15.43 (289 bp)** | **-13.35** |
| *rex-7* | 8,026,460 | 2964 | **-18.72 (85 bp) -12.26 (22 bp)** | **-12.58** |
| *rex-8* | 16,578,214 | 3217 | **-13.00 (60 bp) -14.31 (69 bp) -13.22 (23 bp)** | **-13.52** |
| *rex-9* | 3,135,562 | 1029 | **-12.8** | |
| *rex-10* | 895,450 | 3605 | **-12.60 (63 bp)** | **-14.68** |
| *rex-11* | 4,563,250 | 830 | | |
| *rex-12* | 19,564,937 | 1786 | | |

Listed are the *rex* sites analyzed in this study and their motifs. Motif cutoffs used include MEX with ln(P) < -12 and MEX II with ln(P) < - 12. Also listed are the coordinates (in bp) with the maximum SDC-2 ChIP-seq signal in each *rex* site and the maximum SDC-2 ChIP signal in reads per kilobase per million reads mapped (RPKM) within a 50 bp window.

**Figure 5**

**A**


Cel rex-32
Cbr rex-1
Cbr rex-4
Cbr rex-7
Cbr rex-2
Cbr rex-5
Cbr rex-9
Cel X
*
* CRISPR-mediated insertion of individual rex sites into a C. elegans X chromosome region with no DCC binding
Cel rex-32
592 bp
Cbr rex-1
329 bp
Cbr rex-2
501 bp
Cbr rex-4
601 bp
Cbr rex-5
601 bp
Cbr rex-7
601 bp
Cbr rex-9
601 bp
Key:
Cbr MEX
Cbr MEX II
Cel MEX
Cel MEX II

**B**

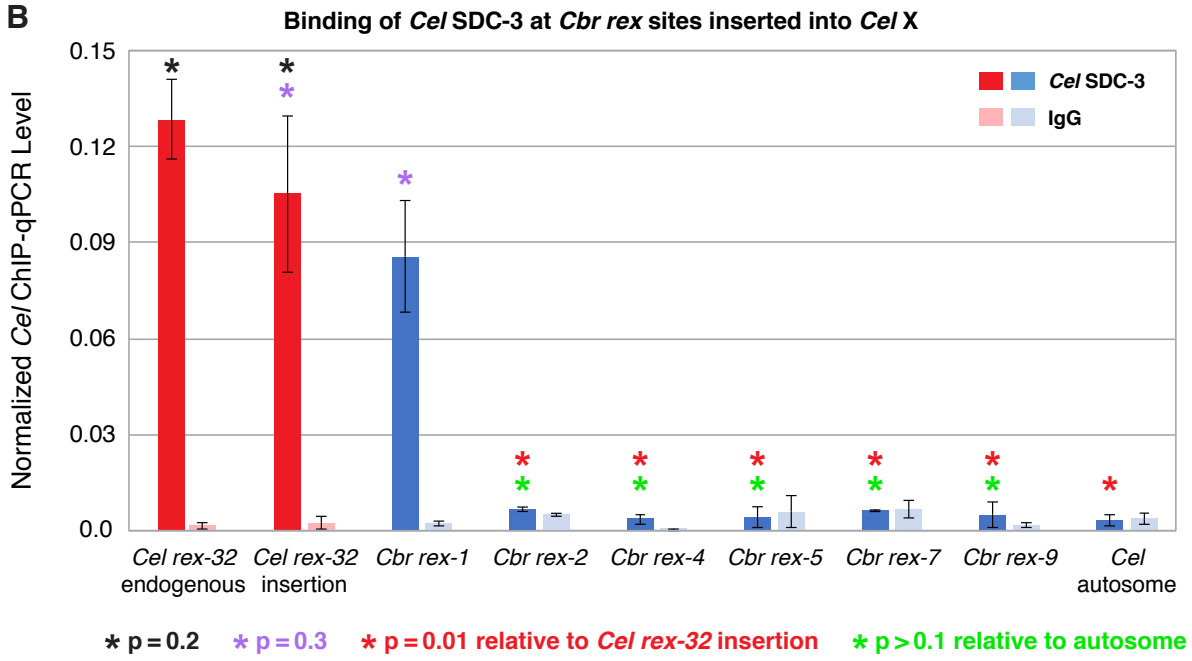

Binding of Cel SDC-3 at Cbr rex sites inserted into Cel X
Normalized Cel ChIP-qPCR Level
0.15
0.12
0.09
0.06
0.03
0.0
Cel SDC-3
IgG
Cel rex-32 endogenous
Cel rex-32 insertion
Cbr rex-1
Cbr rex-2
Cbr rex-4
Cbr rex-5
Cbr rex-7
Cbr rex-9
Cel autosome
* p = 0.2
* p = 0.3
* p = 0.01 relative to Cel rex-32 insertion
* p > 0.1 relative to autosome

**Figure 5.** *C. briggsae rex* sites integrated into the *C. elegans* X chromosome by genome editing failed to recruit the *C. elegans* DCC.

Binding of *C. elegans* DCC protein *Cel* SDC-3 and an IgG control were examined by ChIP-qPCR for *Cel rex-32* at its endogenous location on X, and for six *C. briggsae rex* sites (*Cbr rex-1*, *Cbr rex-2*, *Cbr rex-4*, *Cbr rex-5*, *Cbr rex-7*, and *Cbr rex-9*) plus the control *Cel rex-32* that were inserted by Cas9 genome editing into position 15,574,674 bp of the *C. elegans* X chromosome.

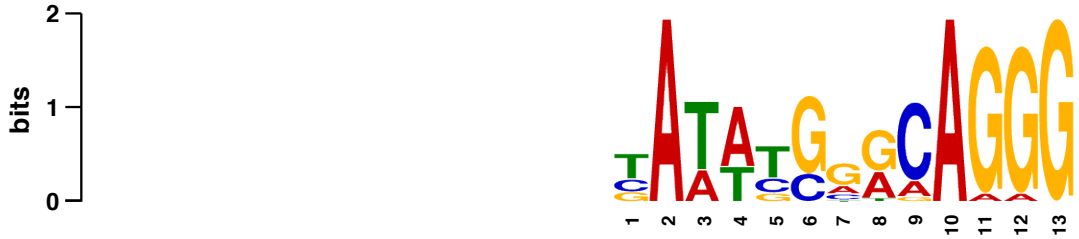
(A) Schematic shows the location of *Cbr rex* insertions in the *Cel* X chromosomes and shows the different combinations of *Cbr* MEX and MEX II motifs in the inserted *Cbr rex* sites.

(B) The graph of *Cel* SDC-3 ChIP-qPCR data shows that all *Cbr rex* sites except *rex-1* exhibited SDC-3 binding that was not significantly different from that of the autosomal negative control. *Cbr rex-1* contains a *Cel* Motif C variant within each *Cbr* MEX motif, thereby accounting for the exceptional SDC-3 binding. The Motif C variants within *Cbr rex-1* MEX include GGGCAGGGT (-11.68), GGGCAGGGG (-14.16), GCGCAGGGC (-12.06), and CGGCAGGGG (-10.72). A fifth Motif C variant lies between the -14.16 and -12.06 variants: TCCAAGGGG (-9.84).

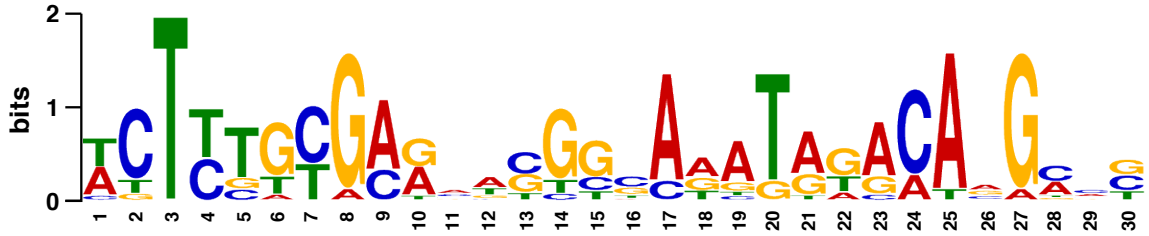
*Cel* SDC-3 levels for each replicate were normalized to the average levels at the five *Cel rex* sites: *Cel rex-8, Cel rex-16*, *Cel rex-32*, *Cel rex-48*, and *Cel rex-35*. Error bars represent the SD for three replicates of *Cel rex-32* and *Cbr rex-1* and two replicates for each of *Cbr rex-2*, *Cbr rex-4*, *Cbr rex-5*, *Cbr rex-7*, and *Cbr rex-9*. *Cel* SDC-3 binding to the endogenous *Cel rex-32* site and the inserted *rex-32* site were not significantly different ($p = 0.2$). *Cel* SDC-3 binding to all *Cbr rex* sites except *Cbr rex-1* was significantly lower than binding to the *Cel rex-32* insertion ($p = 0.01$, Student's *t* test). *Cel* SDC-3 binding at *Cel rex-32* versus *Cbr rex-1* is not significantly different ($p = 0.3$).

Figure 6

A *Cbr* MEX




*Cbr* MEX II




B *Cel* Motif C

*Cel* MEX

*Cel* MEX II

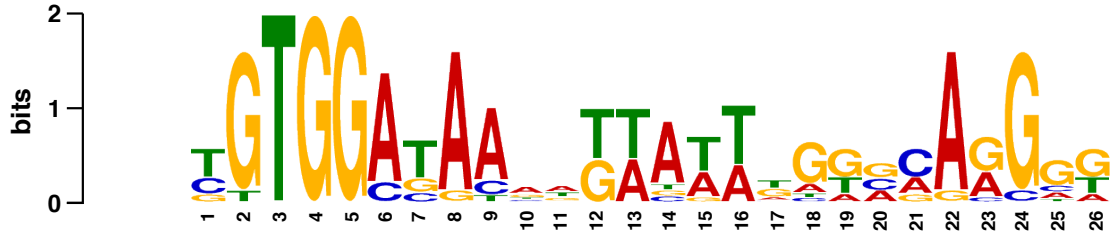


**Figure 6.** Comparison of *C. briggsae* and *C. elegans* DNA motifs on X that occur within respective *rex* sites and recruit respective DCC complexes. (A) Shown are the *C. briggsae* consensus motifs for the 13-bp MEX and 30-bp MEX II variants that recruit the DCC. Also shown are the *C. elegans* consensus motifs for the 12-bp MEX, 26-bp MEX II, and 9-bp Motif C variants that recruit the *Cel* DCC (B). The sequences were aligned relative to the conserved adenine in the 5'-CAGGG-3' common core of the motifs. Predominantly, the *Cel* MEX motif has a cytosine in the fourth position of the motif. Mutating it to a guanine (C4G) severely reduced DCC binding in assays conducted *in vivo* and *in vitro*. The consensus *Cbr* MEX motif has a guanine at the equivalent position relative to the CAGGG core. Hence, the *Cbr* MEX motif is predicted not to function as a DCC recruitment motif in *C. elegans.*

**Figure 6—figure supplement 1**

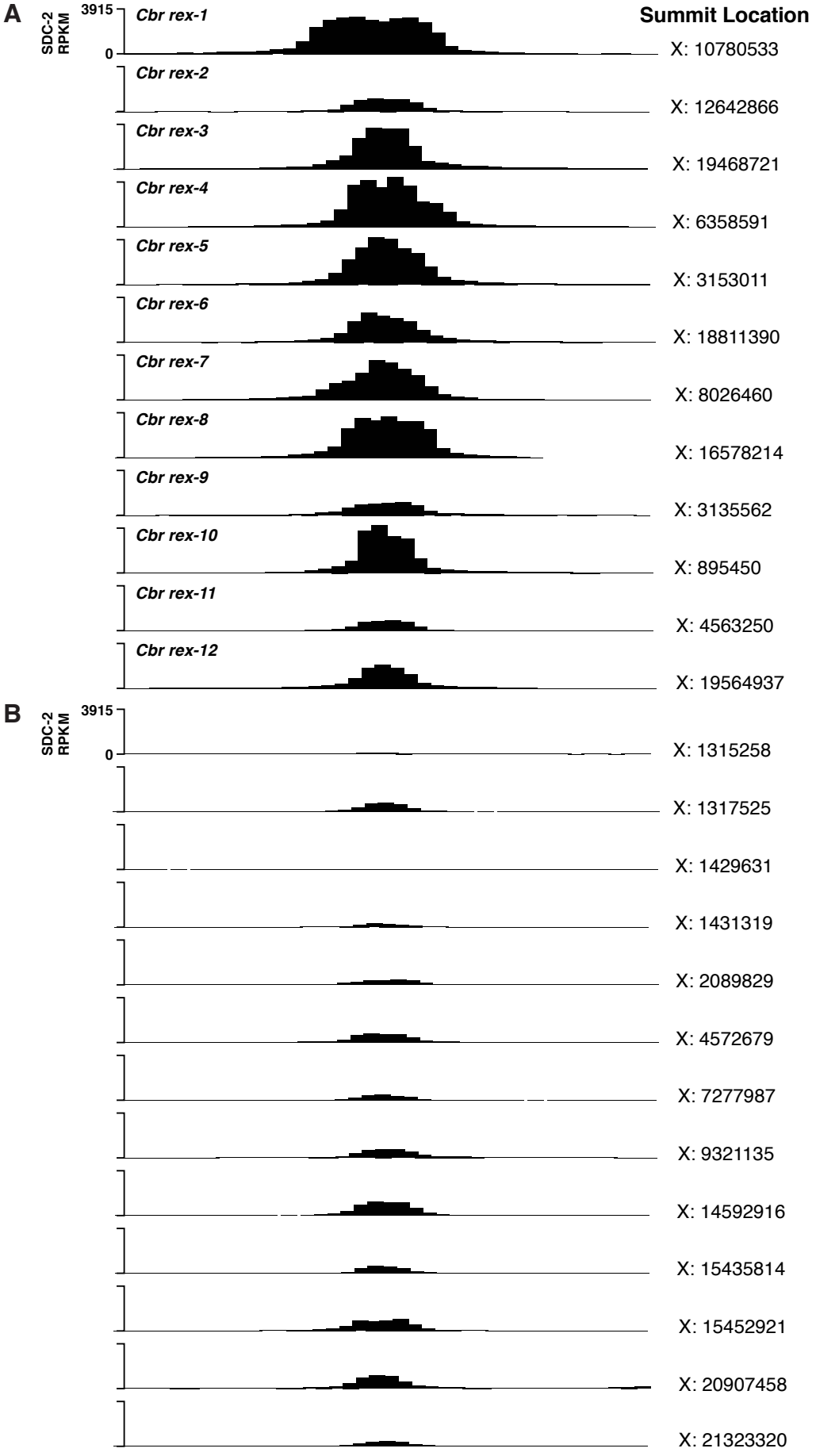


**Figure 6—figure supplement 1.** *C. briggsae* SDC-2 ChIP-seq peak profiles for *rex* sites and non-*rex* sites on X. (A) SDC-2 ChIP-seq profiles for all twelve *Cbr rex* sites. X coordinates for the peak summit locations are shown on the right, and the name of each *rex* site is shown on the left. The y-axis shows the SDC-2 signal in RPKM (reads per kilobase per million reads mapped). (B) SDC-2 Chip-seq peak profiles for the thirteen non-*rex* sites that were analyzed for motif candidates. No motif candidates that correlate with SDC-2 binding were found. The profiles show intermediate levels of SDC-2 binding that is equivalent to or lower than that at *rex-2*. Peak summit locations are shown on the right, and the y-axis shows the SDC-2 RPKM signal.

Figure 7

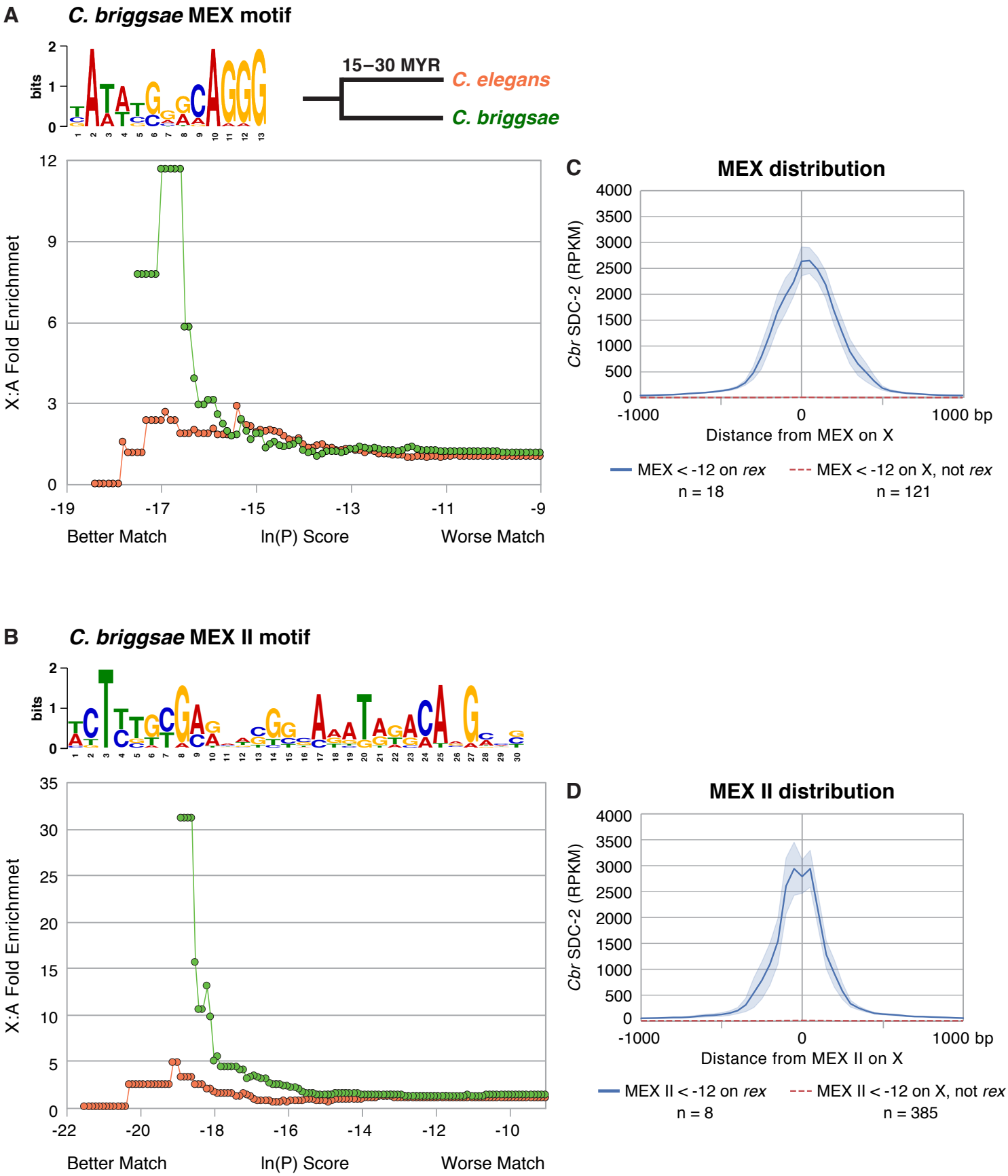

A
C. briggsae MEX motif
bits
15–30 MYR
C. elegans
C. briggsae
X:A Fold Enrichmnet
Better Match
ln(P) Score
Worse Match
C
MEX distribution
Cbr SDC-2 (RPKM)
Distance from MEX on X
1000 bp
MEX < -12 on rex
n = 18
MEX < -12 on X, not rex
n = 121
B
C. briggsae MEX II motif
bits
X:A Fold Enrichmnet
Better Match
ln(P) Score
Worse Match
D
MEX II distribution
Cbr SDC-2 (RPKM)
Distance from MEX II on X
1000 bp
MEX II < -12 on rex
n = 8
MEX II < -12 on X, not rex
n = 385

**Figure 7.** Enrichment of *Cbr* MEX and *Cbr* MEX II motifs on X chromosomes between *C. briggsae* and *C. elegans*. (A, B) Graphs show the enrichment (y axis) of *Cbr* MEX (A) or *Cbr* MEX II (B) variants (x axis) on X chromosomes compared to autosomes in the *C. briggsae* (green circles) and *C. elegans* (orange circles) genomes. For MEX, the ln(P) is the natural log of the probability that a 13-mer matches the MEX consensus motif matrix (shown above graphs) as calculated by the Patser program. For MEX II, the ln(P) is the natural log of the probability that a 30-mer matches the MEX II consensus motif matrix (shown above graphs) as calculated by Patser. The lower the score, the better the match. The maximum theoretical ln(P) value for MEX is -18.7 and for MEX II is -29.3. The best MEX score found on *Cbr* X is -18.7 and for MEX II is - 27.58. The graphs reflect cumulative scores. For example, the 12-fold X:A enrichment of MEX for *C. briggsae* at -17.58 reflects all motifs with ln(P) ≤ -17.58. The *C. elegans* X chromosome is not enriched for the *Cbr* MEX or MEX II consensus motifs that are enriched on *Cbr* X chromosomes and that are pivotal for *Cbr* DCC recruitment to *Cbr* X, as we show subsequently.

(C) The graph plots the mean (dark blue) and standard error (light blue) of *Cbr* SDC-2 ChIP-seq signal (RPKM) at various distances from MEX motifs (< -12) in *rex* sites versus the mean (dashed red) and standard error (light red) of SDC-2 signal at varying distances from MEX motifs (< -12) on X but not in *rex* sites. Abundant SDC-2 binding was found at MEX motifs in *rex* sites, but negligible SDC-2 binding was found at individual MEX motifs on X that were not in *rex* sites. n, total number of MEX motifs in each category.

(D) The graph plots the mean (dark blue) and standard error (light blue) of *Cbr* SDC-2 ChIP-seq signal (RPKM) at various distances from MEX II motifs (< -12) in *rex* sites versus the mean (dashed red line) and standard error (light red) of SDC-2 signal at varying distances from MEX II motifs (< -12) on X but not in *rex* sites. Abundant SDC-2 binding was found at MEX II motifs in *rex* sites, but negligible SDC-2 binding was found at individual MEX II motifs on X that were not in *rex* sites. n, total number of MEX II motifs in each category.

Figure 7—figure supplement 1

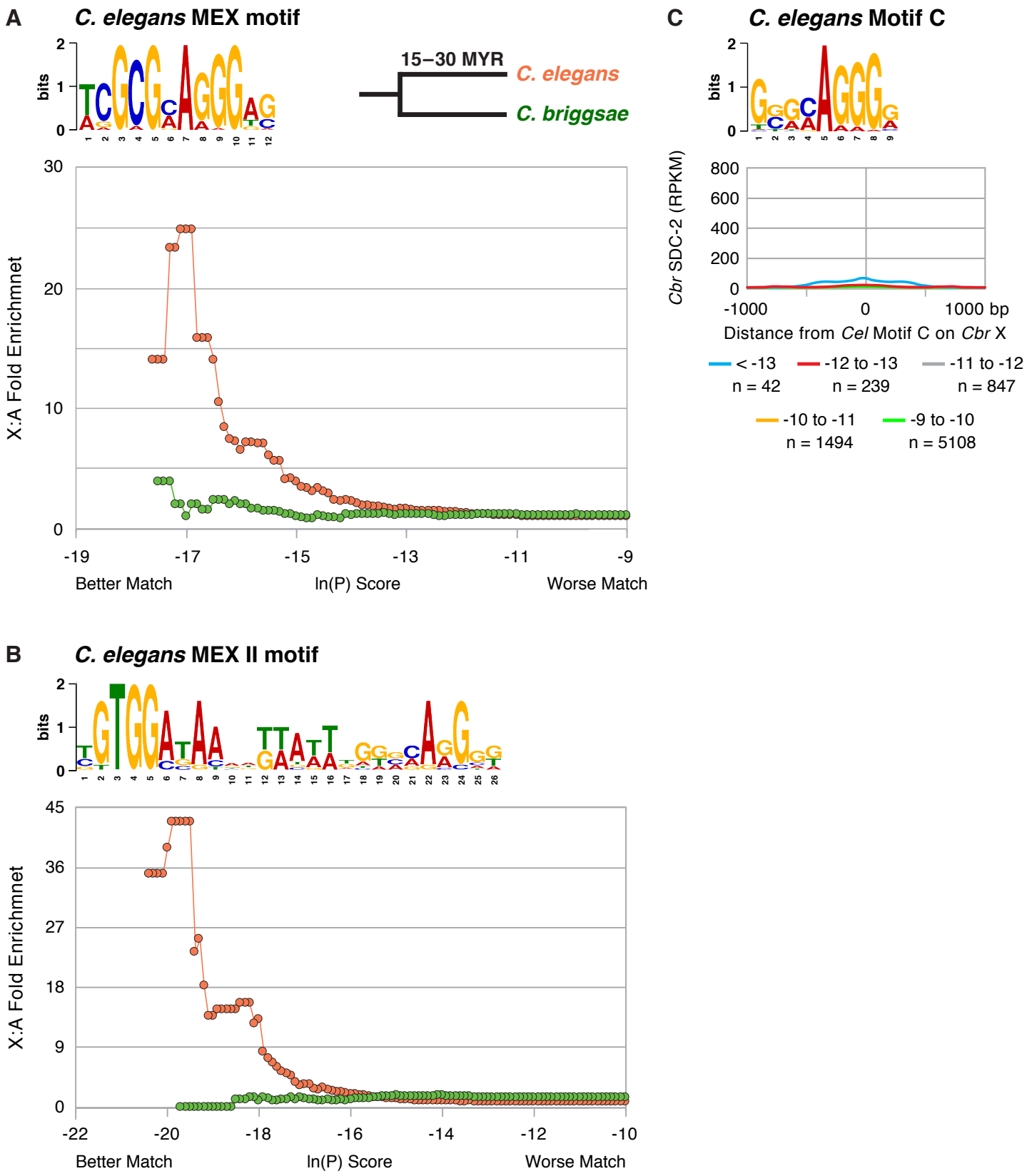
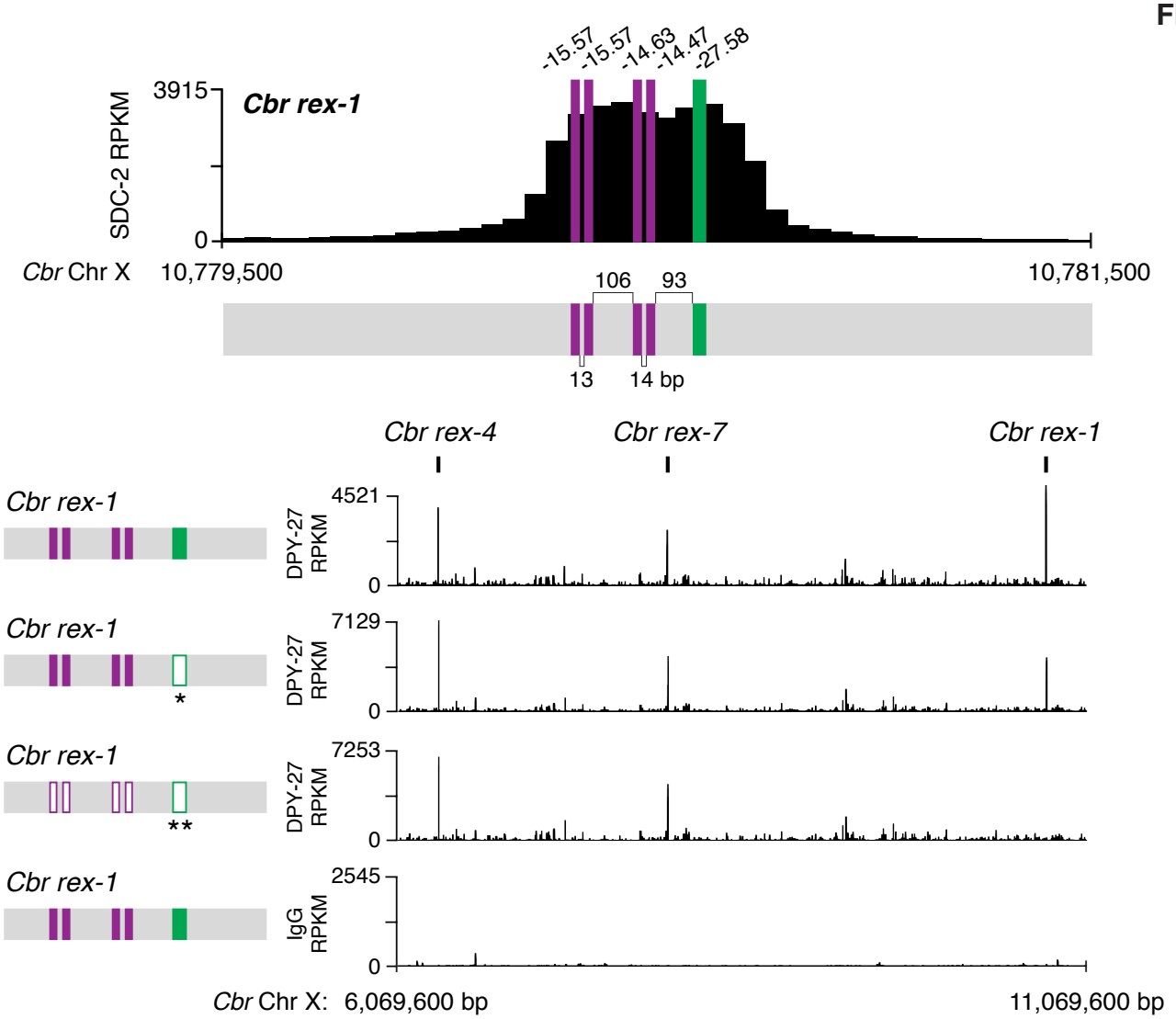


**Figure 7—figure supplement 1.** The *C. briggsae* X chromosome is not enriched for the *C. elegans* MEX (A) or MEX II (B) motifs that are highly enriched on *Cel* X chromosomes and pivotal for DCC binding to *Cel* X chromosomes *in vivo.* The descriptions of these graphs are the same as those presented in the legend to Figure 7. (C) Graph shows the *Cbr* SDC-2 RPKM signal from ChIP-seq experiments as a function of the distance from *Cel* Motif C variants of different matches (ln[P] score) to the consensus motif found on *Cbr* X chromosomes. *Cbr* SDC-2 binding is negligible at most *Cel* Motif C variants, indicating that *Cel* Motif C fails to participate in *Cbr* DCC recruitment to X chromosomes. The slight increase in SDC-2 signal at Motif C (< -13) variants is due to their location within MEX and MEX II motifs.

**Figure 8**



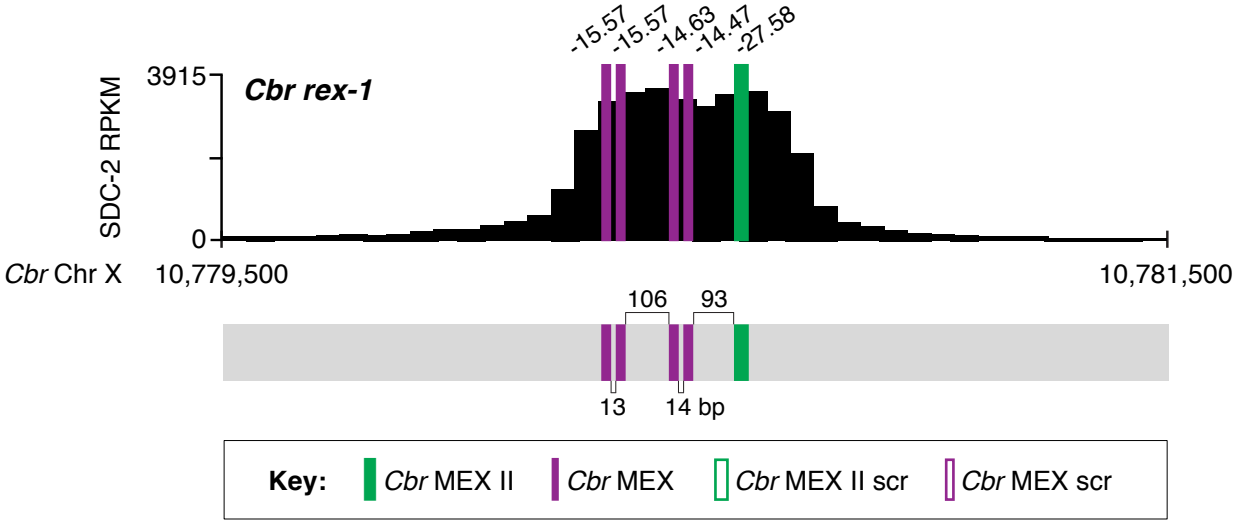

**Key:** *Cbr* MEX II | *Cbr* MEX | *Cbr* MEX II scr | *Cbr* MEX scr

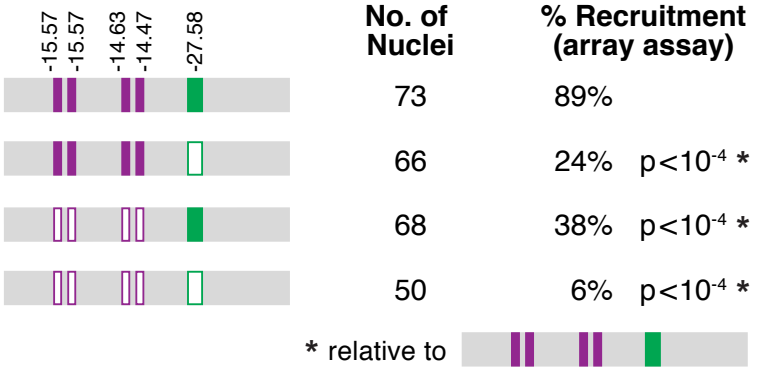
| Motif | Sequence |
|---|---|
| *Cbr rex-1* MEX (-15.57) | TAAAGGGGCAGGGT |
| *Cbr rex-1* MEX (-15.57) scr | AGAGTGGAGTGGCA |
| *Cbr rex-1* MEX (-15.57) | TAAAGGGGCAGGGG |
| *Cbr rex-1* MEX (-15.57) scr | TGAGGAGGAGCGGA |
| *Cbr rex-1* MEX (-14.63) | CATTTGCGCAGGGC |
| *Cbr rex-1* MEX (-14.63) scr | ACTGTGCGTCGCAG |
| *Cbr rex-1* MEX (-14.47) | GATTTCGGCAGGGG |
| *Cbr rex-1* MEX (-14.47) scr | ATGCGTGCGGTGGA |
| *Cbr rex-1* MEX II (-27.58) | AGGCTTGTCTATATGCCCTTATCGCAAAGA |
| *Cbr rex-1* MEX II (-27.58) scr* | AGTACTGCTATACGTCTCGTTAGACTCAGA |
| *Cbr rex-1* MEX II (-27.58) scr** | CGTACTAGCAGTAGTCTCGTAGATACTATC |

**Figure 8.** Combinatorial clustering of MEX and MEX II motifs in *Cbr rex-1* facilitates DCC binding to the endogenous *rex-1* site on X. Shown is an enlargement of the SDC-2 ChIP-seq peak profile for *Cbr rex-1* with its associated MEX and MEX II motifs and their ln(P) scores.

ChIP-seq analysis (graph) was performed using anti-FLAG antibody on an otherwise genetically wild-type *C. briggsae* strain encoding FLAG-tagged DPY-27 and on FLAG-tagged DPY-27 *C. briggsae* mutant variants carrying either a scrambled (scr) version of MEX II or a scrambled version of MEX II and all four MEX motifs. The DPY-27 and control IgG ChIP-seq profiles are also shown for *Cbr* sites *rex-7* and *rex-4* as an internal standard since DPY-27 binding is not disrupted at these sites. Sequences of the wild-type *Cbr rex-1* MEX motifs and their scrambled versions are shown below the graph. Underlined is the C*el* Motif C variant within each C*br* MEX motif. For analyzing MEX II, two different MEX II mutant variants were used, as indicated by asterisks. Numbers between motifs indicate the base pairs separating the motifs. ChIP-seq profiles reveal that mutating only MEX II reduces some DCC binding at *rex-1*, and mutating MEX II and all MEX motifs eliminates DCC binding. The motifs act cumulatively to recruit the DCC.

**Figure 8—figure supplement 1**

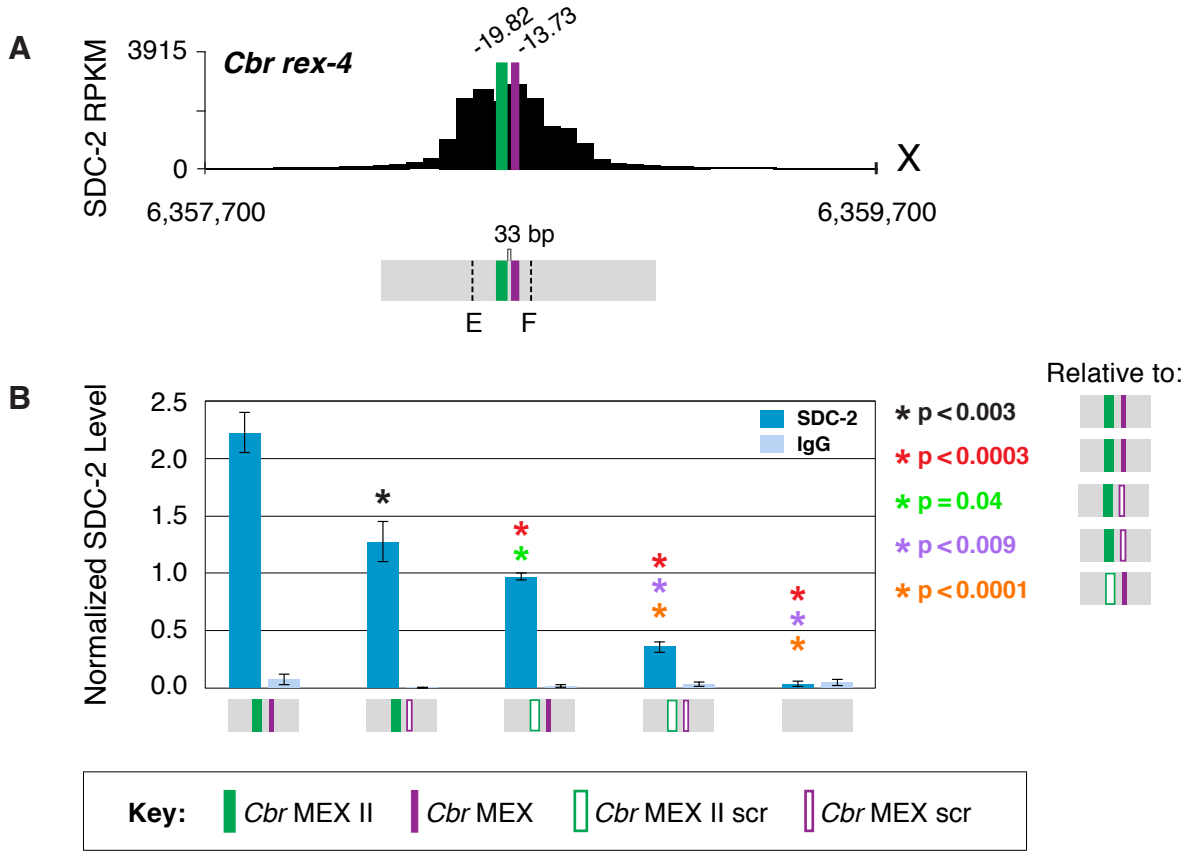



| Motif | Sequence |
|---|---|
| *Cbr rex-1* MEX (-15.57) | TAAAGGGGCAGGGT |
| *Cbr rex-1* MEX (-15.57) scr | AGAGTGGAGTGGCA |
| *Cbr rex-1* MEX (-15.57) | TAAAGGGGCAGGGG |
| *Cbr rex-1* MEX (-15.57) scr | TGAGGAGGAGCGGA |
| *Cbr rex-1* MEX (-14.63) | CATTTGCGCAGGGC |
| *Cbr rex-1* MEX (-14.63) scr | ACTGTGCGTCGCAG |
| *Cbr rex-1* MEX (-14.47) | GATTTCGGCAGGGG |
| *Cbr rex-1* MEX (-14.47) scr | ATGCGTGCGGTGGA |
| *Cbr rex-1* MEX II (-27.58) | AGGCTTGTCTATATGCCCTTATCGCAAAGA |
| *Cbr rex-1* MEX II (-27.58) scr | CGTACTAGCAGTAGTCTCGTAGATACTATC |

**Figure 8—figure supplement 1.** Clustering of MEX and MEX II motifs in *Cbr rex-1* confers DCC binding *in vivo.* Shown is an enlargement of the SDC-2 ChIP-seq peak profile for *Cbr rex-1* with its associated MEX and MEX II motifs and their ln(P) scores. Numbers between motifs indicate the base pairs separating the motifs. *Cbr* DPY-27 binding to wild-type and mutant versions of *rex-1* was assayed using extrachromosomal arrays carrying multiple copies of wild-type *rex-1* or mutant *rex-1* variants with either a scrambled MEX II sequence, four scrambled MEX sequences, or a scrambled MEX II sequence and four scrambled MEX sequences. Shown is the total number of array-bearing nuclei that were assayed and the percentage of those nuclei exhibiting DPY-27 binding. The assays show that mutating only MEX II or only the four MEX motifs reduces DPY-27 binding, while mutating both MEX and MEX II motifs virtually eliminates DPY-27 binding. These results indicate that both MEX and MEX II motifs are important for DCC binding at *rex* sites *in vivo*. The p values were determined using the Student's *t* test and are relative to DPY-27 recruitment to arrays carrying wild-type *rex-1* sequences. Sequences of the wild-type *Cbr rex-1* motifs and their scrambled versions are shown below the graph.

**Figure 9**

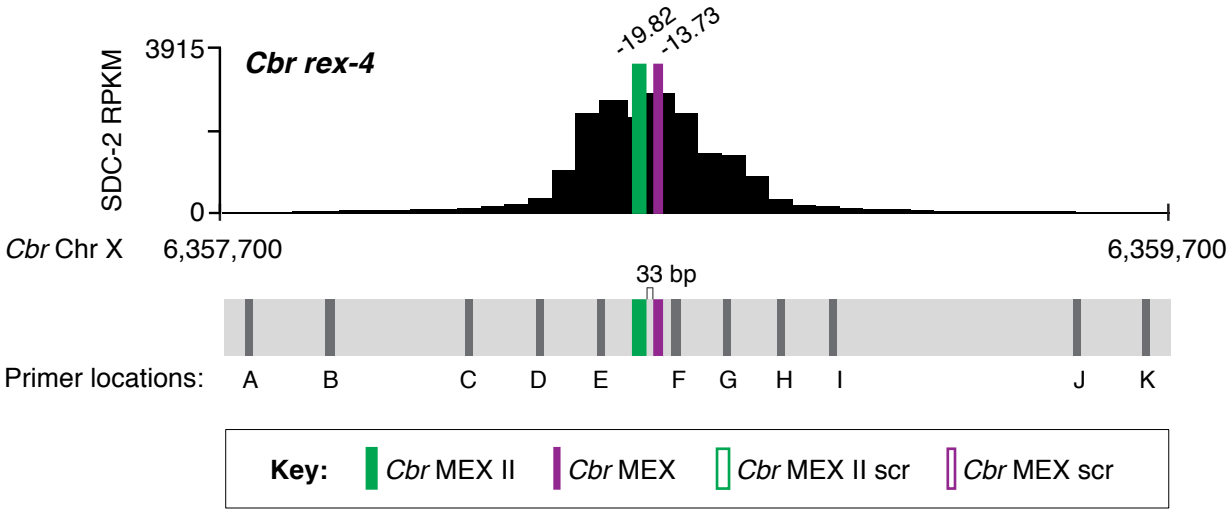


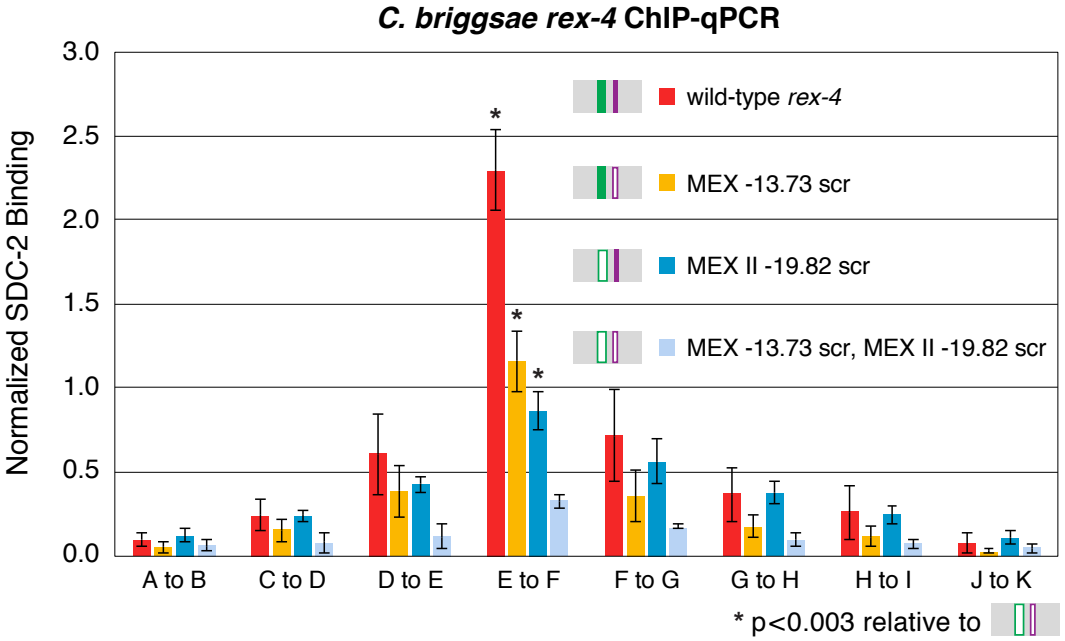

C

| Motif | Sequence |
|---|---|
| *Cbr rex-4* MEX (-13.73) | TATATGAACAGGG |
| *Cbr rex-4* MEX (-13.73) scr | GATAAGTGATGAC |
| *Cbr rex-4* MEX II (-19.82) | TCTCTGCGCGTATGGCCAATTGGCATGGGC |
| *Cbr rex-4* MEX II (-19.82) scr | TGCTGCTGCAGTCACTGTGCTCGCTCTGAC |

**Figure 9.** MEX and MEX II motifs are critical for DCC binding to *Cbr rex-4 in vivo.*
Shown is an enlargement of the SDC-2 ChIP-seq profile for *rex-4*, a schematic of the MEX (purple) and MEX II (green) motifs in *rex-4*, and the location of primers (E and F, dashed lines) to evaluate DCC binding *in vivo* using ChIP-qPCR. Motifs are separated by 33 bp. The graph shows ChIP qPCR levels for SDC-2 (dark blue) and control IgG (light blue) at endogenous wild-type *rex-4*, at endogenous *rex-4* with different combinations of motif mutations created by genome editing, and at a negative control site on X of 107 bp that lacks DCC binding centered at (7,000,213 bp). Strains carrying wild-type and mutant motifs encoded FLAG-tagged SDC-2. SDC-2 levels for each replicate were normalized to the average level of five endogenous non-edited *rex* sites (*Cbr rex-1*, *Cbr rex-2*, *Cbr rex-5*, and *Cbr rex-9*). Error bars represent the standard deviation (SD) of three replicates. Asterisks of the same color specify data compared using the Student's *t* test. If more than two motif combinations are compared, the schematic to the right of the p-value indicates the motif combination to which the other combinations were compared. DNA sequences of wild-type and mutant motifs (scr) are shown below the graph. Both MEX and MEX II motifs are critical for DCC binding at *rex-4*. Mutating each motif independently causes an equivalent reduction in DCC binding, and mutating both motifs is necessary to eliminate DCC binding. ChIP-qPCR analysis of SDC-2 binding at intervals across the entire peak are presented in ***Figure 9—figure supplement 1***.

**Figure 9—figure supplement 1**

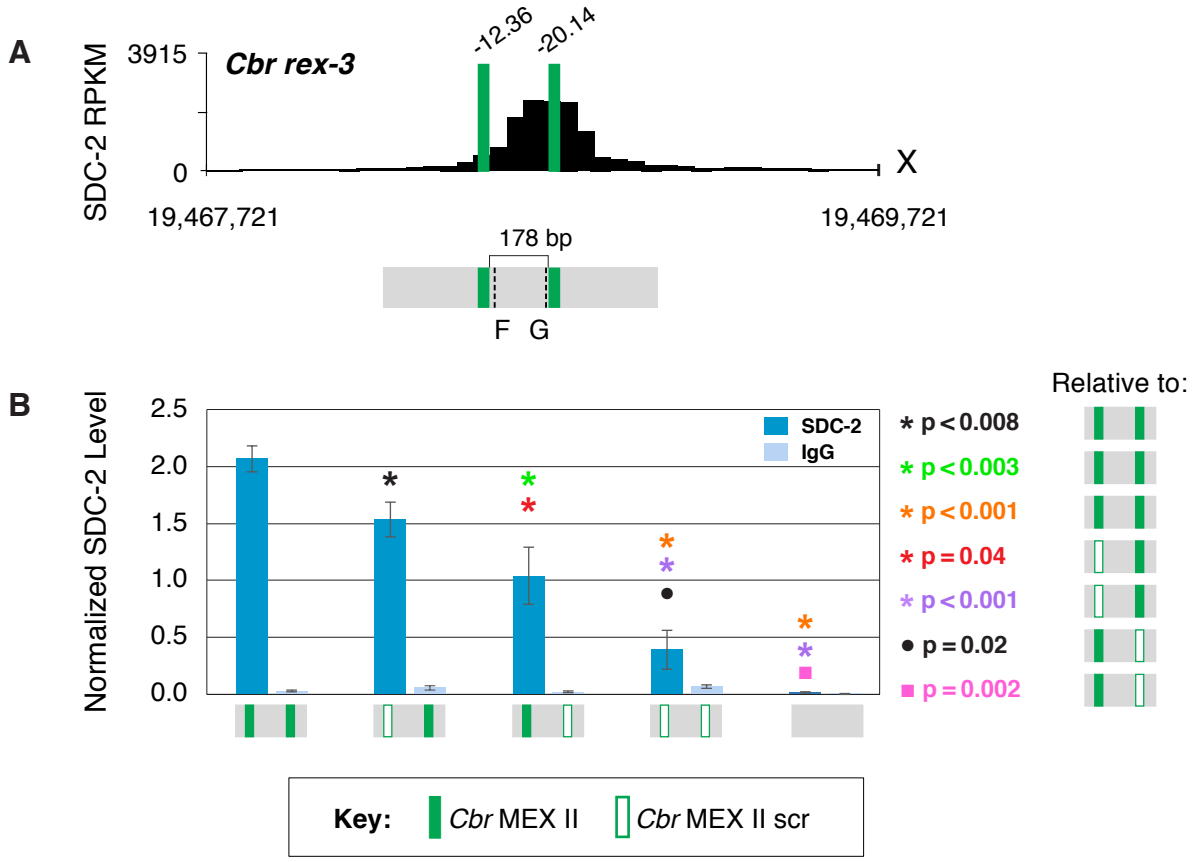



| Motif | Sequence |
|---|---|
| *Cbr rex-4* MEX (-13.73) | TATATGAACAGGG |
| *Cbr rex-4* MEX (-13.73) scr | GATAAGTGATGAC |
| *Cbr rex-4* MEX II (-19.82) | TCTCTGCGCGTATGGCCAATTGGCATGGGC |
| *Cbr rex-4* MEX II (-19.82) scr | TGCTGCTGCAGTCACTGTGCTCGCTCTGAC |

**Figure 9—figure supplement 1.** MEX and MEX II motifs are critical for SDC-2 binding to *Cbr rex-4 in vivo.* This figure extends the analysis of SDC-2 binding at *rex-4* in wild-type and *rex-4* mutant strains presented in Figure 9 by including SDC-2 ChIP-qPCR analysis at intervals extending all along the entire SDC-2 peak. The schematic showing motifs in *rex-4* includes the locations of primers (grey) used for the PCR analysis presented in the graph below it. Error bars represent the standard deviation (SD) of three replicates. The asterisks highlight the SDC-2 binding values that are significantly different in the E-F interval of the *rex-4* site that is mutant for both MEX and MEX II versus wild-type rex-4 or *rex-4* with either MEX or MEX II scrambled. Statistics were determined using the Student's *t* test. All other aspects of the figure resemble those explained in the legend to Figure 9.

**Figure 10**

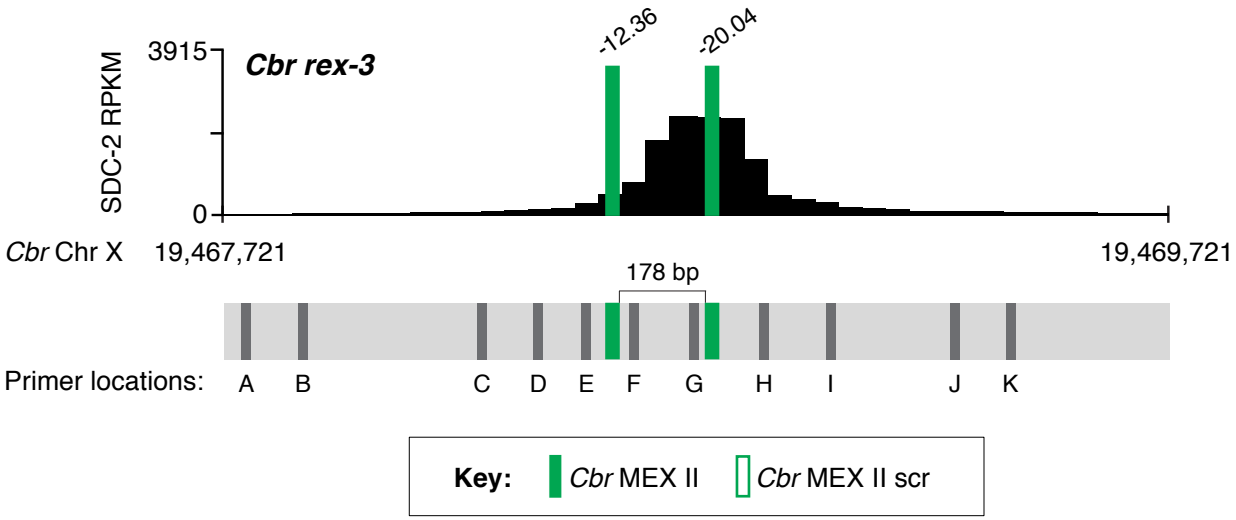


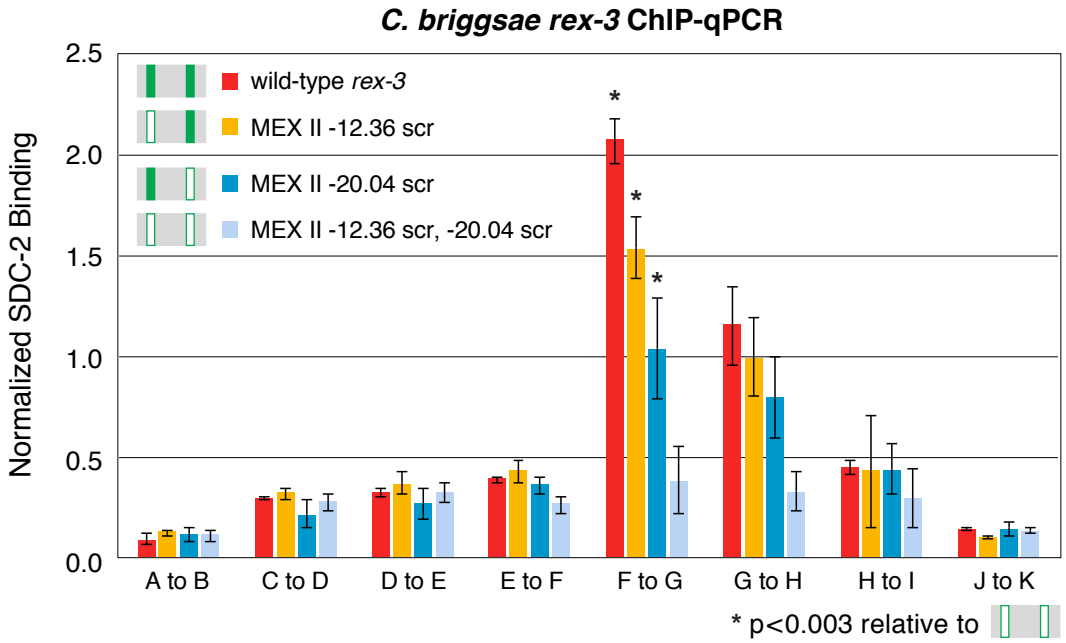

C

| Motif | Sequence |
|---|---|
| *Cbr rex-3* MEX II (-12.36) | ACTCTGTGAAAATAGGAAATCTACAAAATT |
| *Cbr rex-3* MEX II (-12.36) scr | TACAGATGTATGATTATCACAGTATTACTA |
| *Cbr rex-3* MEX II (-20.04) | ACTTCGCGCGTTGTGGAAGTAGGCAGGAGG |
| *Cbr rex-3* MEX II (-20.04) scr | GAGCTCGCTGTGCTGACGAGAGCAGCTCAT |

**Figure 10.** Both MEX II motifs are critical for DCC binding to *Cbr rex-3 in vivo*. Shown is an enlargement of SDC-2 ChIP-seq profile for *Cbr rex-3* with its associated MEX II motifs (green) and their ln(P) scores. Motifs are separated by 178 bp. Locations of primers (F and G, dashed lines) to evaluate DCC binding *in vivo* using ChIP-qPCR are shown. The graph shows ChIP qPCR levels for SDC-2 (dark blue) and control IgG (light blue) at endogenous wild-type *rex-3*, at endogenous *rex-3* with different combinations of motif mutations created by genome editing, and at a negative control site on X that lacks DCC binding. Strains carrying wild-type and mutant motifs encoded FLAG-tagged SDC-2. SDC-2 levels for each replicate were normalized to the average level of five endogenous non-edited *rex* sites (*Cbr rex-1*, *Cbr rex-2*, *Cbr rex-5*, and *Cbr rex-9*). Error bars represent the standard deviation (SD) of three replicates. Symbols of the same color specify data compared using the Student's *t* test. If more than two motif combinations are compared, the schematic to the right of the p value indicates the motif combination to which the other combinations were compared. DNA sequences of wild-type and mutant motifs (scr) are shown below the graph. Both MEX II motifs are critical for DCC binding at *rex-3*. Mutating each motif independently causes an equivalent reduction in DCC binding, and mutating both motifs is necessary to eliminate DCC binding. ChIP-qPCR analysis of SDC-2 binding at intervals across the entire peak are presented in ***Figure 10—figure supplement 1***.

**Figure 10—figure supplement 1**

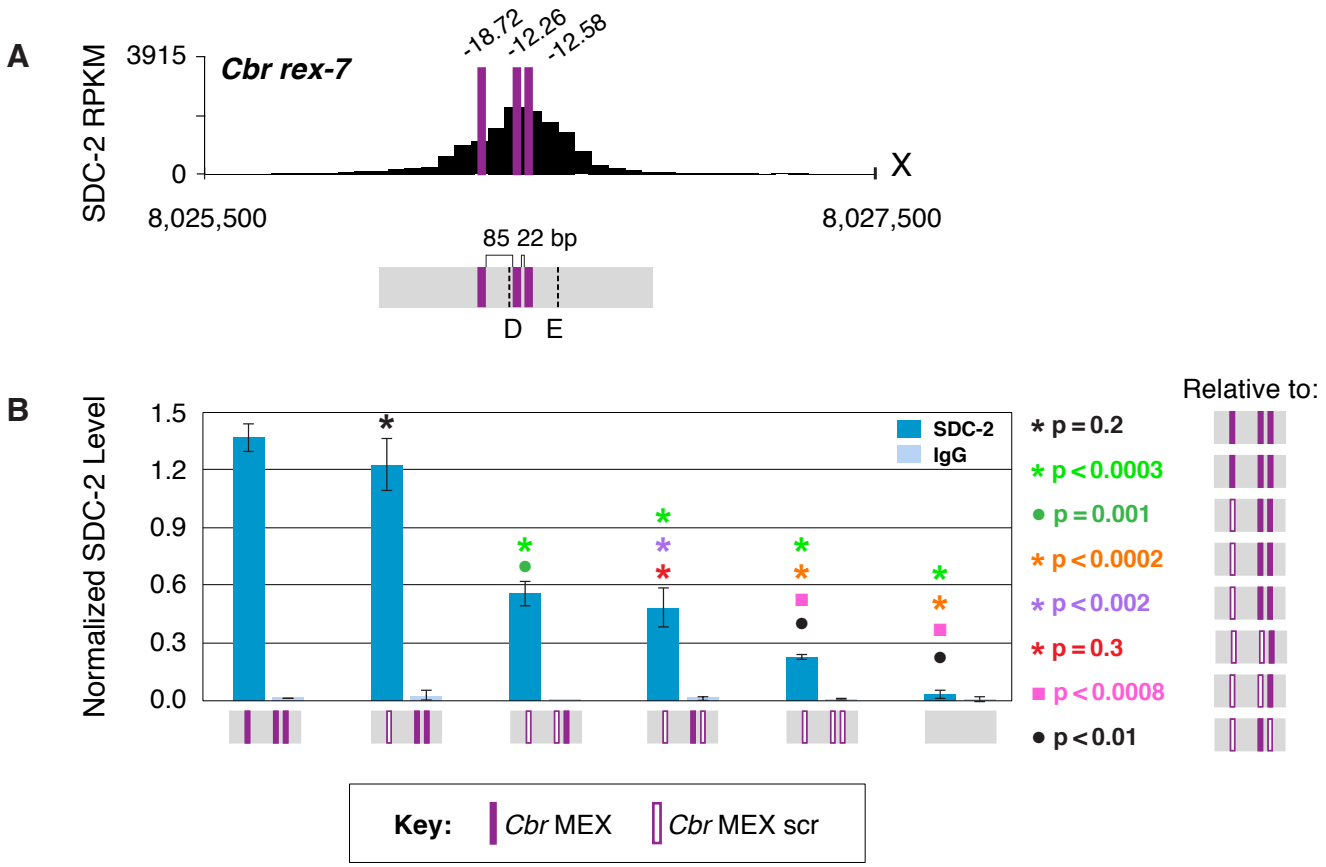



| Motif | Sequence |
|---|---|
| *Cbr rex-3* MEX II (-12.36) | ACTCTGTGAAAATAGGAAATCTACAAAATT |
| *Cbr rex-3* MEX II (-12.36) scr | TACAGATGTATGATTATCACAGTATTACTA |
| *Cbr rex-3* MEX II (-20.04) | ACTTCGCGCGTTGTGGAAGTAGGCAGGAGG |
| *Cbr rex-3* MEX II (-20.04) scr | GAGCTCGCTGTGCTGACGAGAGCAGCTCAT |

**Figure 10—figure supplement 1.** Both MEX II motifs are critical for DCC binding to *Cbr rex-3 in vivo*. This figure extends the analysis of SDC-2 binding at *rex-3* in wild-type and *rex-3* mutant strains in Figure 10 by including SDC-2 ChIP-qPCR analysis at intervals extending all along the SDC-2 entire peak. The schematic of motifs in *rex-3* includes the locations of primers (grey) used for the PCR analysis presented in the graph below it. Error bars represent the standard deviation (SD) of three replicates. The asterisks highlight the SDC-2 binding values that are significantly different in the F-G interval of the *rex-3* site that is mutant for both MEX II motifs versus wild-type *rex-3* or *rex-3* with one scrambled MEX II motif. Statistics were determined using the Student's *t* test. All other aspects of the figure resemble those explained in the legend to Figure 10.

**Figure 11**

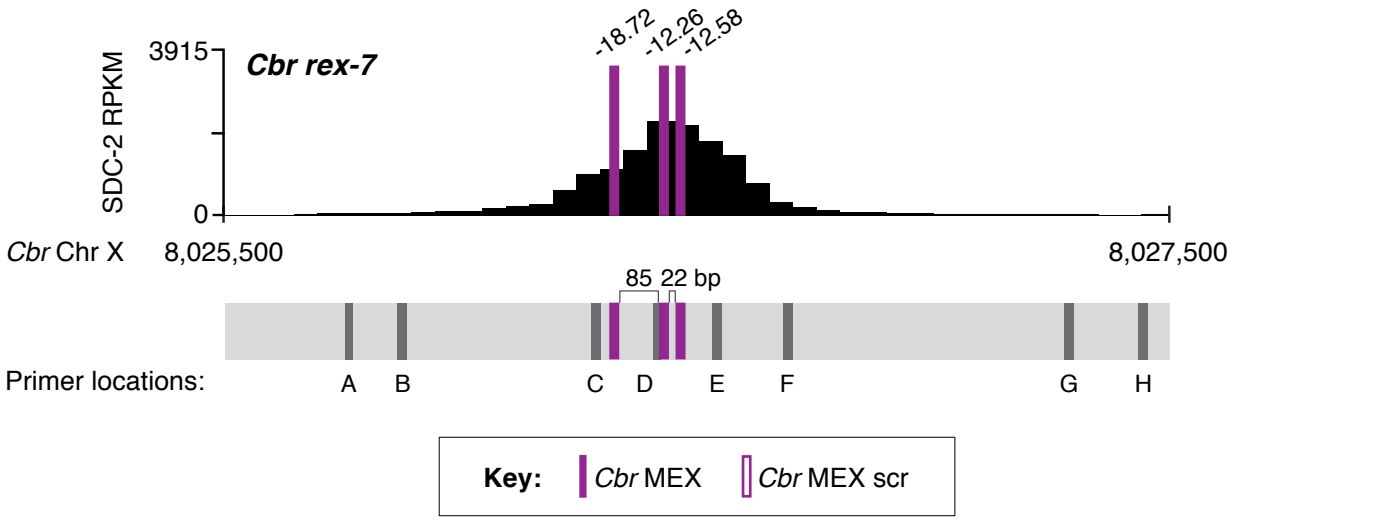


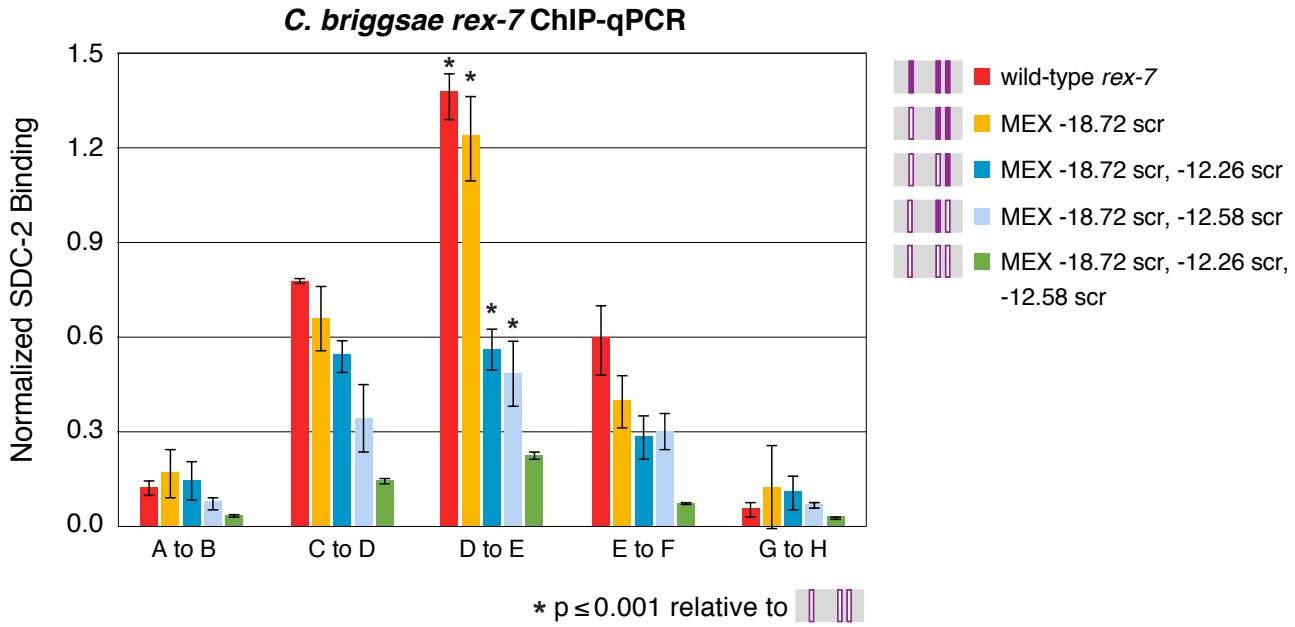

C

| Motif | Sequence |
|---|---|
| *Cbr rex-7* MEX (-18.72) | TATATGGGCAGGG |
| *Cbr rex-7* MEX (-18.72) scr | GTGAGCAGTAGTG |
| *Cbr rex-7* MEX (-12.26) | GAATGCGGCAGGG |
| *Cbr rex-7* MEX (-12.26) scr | CACGACGACGCGA |
| *Cbr rex-7* MEX (-12.58) | TAATTGGGAAGGG |
| *Cbr rex-7* MEX (-12.58) scr | AGTCAGTGTCAGA |

**Figure 11.** Multiple MEX motifs in *Cbr rex-7* contribute to DCC binding *in vivo*. Shown is an enlargement of SDC-2 ChIP-seq profile for *Cbr rex-7* with its associated MEX motifs (purple) and their ln(P) scores. Motifs are separated by 85 bp and 22 bp. Locations of primers (D and E, dashed lines) to evaluate DCC binding *in vivo* using ChIP-qPCR are shown. The graph shows ChIP qPCR levels for SDC-2 (dark blue) and control IgG (light blue) at endogenous wild-type *rex-7*, at endogenous *rex-7* with different combinations of motif mutations created by genome editing, and at a negative control site on X that lacks DCC binding. Strains carrying wild-type and mutant motifs encoded FLAG-tagged SDC-2. SDC-2 levels for each replicate were normalized to the average level of five endogenous non-edited *rex* sites (*Cbr rex-1*, *Cbr rex-2*, *Cbr rex-5*, and *Cbr rex-9*). Error bars represent the standard deviation (SD) of three replicates. Symbols of the same color specify data compared using the Student's *t* test. If more than two motif combinations are compared, the schematic to the right of the p value indicates the motif combination to which the other combinations were compared. Sequences of wild-type and mutant motifs (scr) are shown below the graph. Multiple MEX motifs contribute to DCC binding at *rex-7*. Mutating the first MEX motif has an insignificant effect on DCC binding, but mutating the first MEX motif and either of the other two motifs reduces binding equivalently. Mutating all three MEX motifs eliminates DCC binding. ChIP-qPCR analysis of SDC-2 binding at intervals across the entire peak are presented in ***Figure 11—figure supplement 1***.

**Figure 11—figure supplement 1**




| Motif | Sequence |
|---|---|
| *Cbr rex-7* MEX (-18.72) | TATATGGGCAGGG |
| *Cbr rex-7* MEX (-18.72) scr | GTGAGCAGTAGTG |
| *Cbr rex-7* MEX (-12.26) | GAATGCGGCAGGG |
| *Cbr rex-7* MEX (-12.26) scr | CACGACGACGCGA |
| *Cbr rex-7* MEX (-12.58) | TAATTGGGAAGGG |
| *Cbr rex-7* MEX (-12.58) scr | AGTCAGTGTCAGA |

**Figure 11—figure supplement 1.** Multiple MEX motifs in *Cbr rex-7* contribute to DCC binding *in vivo*. This figure extends the analysis of SDC-2 binding at *rex-7* in wild-type and *rex-7* mutant strains in Figure 11 by including SDC-2 ChIP-qPCR analysis at intervals extending all along the SDC-2 entire peak. The schematic of motifs in *rex-7* includes the locations of primers (grey) used for the PCR analysis presented in the graph below it. Error bars represent the standard deviation (SD) of three replicates. The asterisks highlight the SDC-2 binding values that are significantly different in the D-E interval of the *rex-7* site that is mutant for all three MEX motifs versus wild-type *rex-7* or *rex-7* with different combinations of scrambled MEX motifs. Statistics were determined using the Student's *t* test. All other aspects of the figure resemble those explained in the legend to Figure 11.

**Figure 12**

**A** ***Cel rex-39***

*Cel* MEX II TGTGGAGAACATTATTTGGGCACGTG CGTGGCTAACGTAATTGGTGCAAGCT

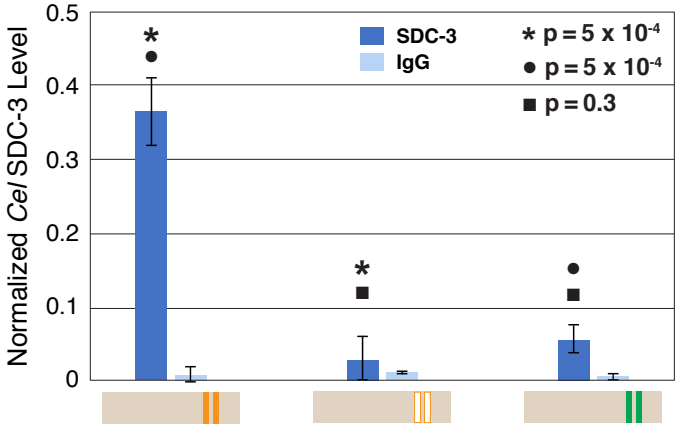
*Cbr* MEX II ACTTCGCGCGTTGTGGAAGTAGGCAGGAGG ACTTCGCGCGTTGTGGAAGTAGGCAGGAGG

*Cel* scrambled MEX II AAAAAGATGAACGCGCTCCTTTTTTC TTGCATATCTCGCGCAAACAATGAAA

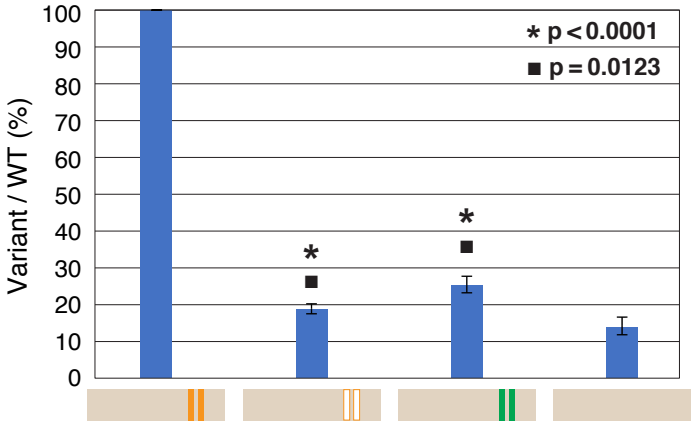
*Cel* MEX II spacer TGTAGTTACTTGTTTCT

*Cbr* MEX II spacer AGTAGTTACTTGTTACT

**B** ***Cel* SDC-3 on *rex-39 in vivo***




**C** ***Cel* SDC-2 on *rex-39 in vitro***

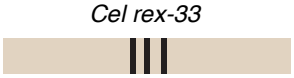


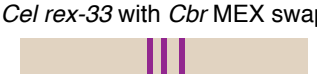

**D** ***Cel rex-33***

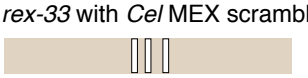
*Cel* MEX TTTAAGCGAAGGGAC ATTTGGCGCAGGGGG GTATCGCGCAGGGAA

*Cbr* MEX TATATGGGCAGGGAC GAATGCGGCAGGGGG TAATTGGGAAGGGAA

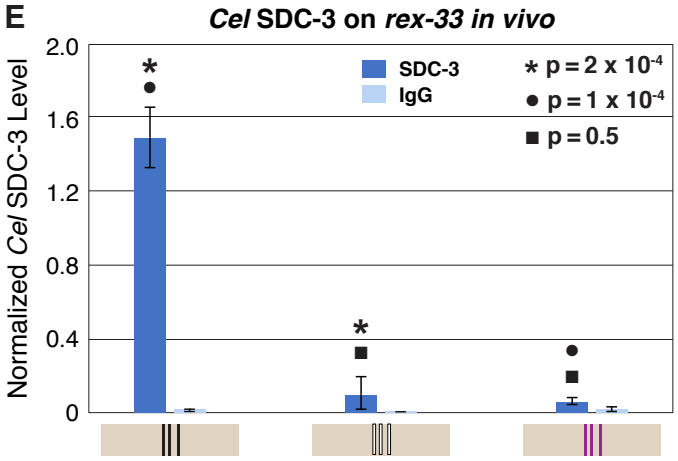
*Cel* scrambled MEX TTTAGTCGACGTGAC ATTGCGCTCGAGCGC GTATGACGCGACGCT




**E** ***Cel* SDC-3 on *rex-33 in vivo***




**F** ***Cel* SDC-2 on *rex-33 in vitro***

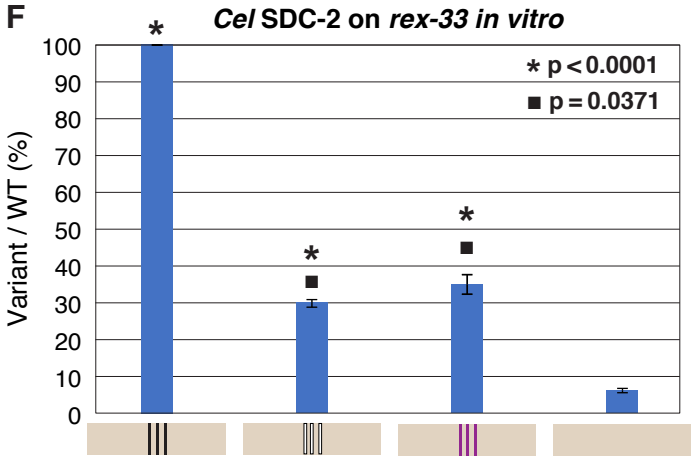


**Figure 12.** Functional divergence of X motifs demonstrated by *C. elegans* DCC binding studies *in vivo* and *in vitro* to *Cel rex* sites engineered to replace *Cel* motifs with *Cbr* MEX and MEX II motifs.

(A) Comparison of DNA sequences for the two MEX II motifs in wild-type *Cel rex-39* (*Cel* ln[P] of -21.23 and -20.74) with the *Cbr* MEX II motifs (*Cbr* ln[P] of -20.04 and *Cel* ln[P] > -9 for both) that replaced them. DNA sequences of the spacer region between wild-type *Cel* MEX II motifs and inserted *Cbr* MEX II motifs are shown as are sequences of the scrambled *Cel* MEX II motifs used as negative controls. Schematics show keys for *rex* sites analyzed for *Cel* SDC-3 binding *in vivo* and *Cel* SDC-2 binding *in vitro*: wild-type *Cel rex-39* (orange, MEX II motifs), *Cel rex-39* with *Cbr* MEX II motifs (green), *Cel rex-39* with scrambled *Cel* MEX II motifs (orange outline).

(B) Graph shows ChIP qPCR levels for *Cel* SDC-3 (dark blue) and control IgG (light blue) at wild-type *Cel rex-39* and mutant *rex-39* with *Cbr* MEX II motifs *in vivo*. *Cel* SDC-3 binds *in vivo* to endogenous *Cel rex-39* sites with wild-type MEX II motifs but not to mutant *Cel rex-39* sites with either scrambled *Cel* MEX II motifs or *Cbr* MEX II motif replacements. SDC-3 levels for each replicate were normalized to the average SDC-3 level at 7 control *rex* sites (*Cel rex-8*, *Cel rex-14*, *Cel rex-16*, *Cel rex-32*, *Cel rex-35*, *Cel rex-36*, and *Cel rex-48*). Error bars represent the standard deviation (SD) of three replicates. Statistical comparisons were calculated using the Student's *t* test.

(C) Graph of *in vitro* assay assessing *Cel* SDC-2 binding to a wild-type *Cel rex-39* DNA template and a mutant *rex-39* template with *Cbr* MEX II motifs. *Cel* SDC-2 binds to the *Cel rex-39* template with wild-type MEX II motifs but not to mutant *rex-39* templates with either scrambled *Cel* MEX II motifs or *Cbr* MEX II motif replacements. *Cel* SDC-2 does not bind to the control template (beige) made of DNA from a site on the *Cel* X that lacks SDC-2 binding *in vivo*. SDC-2 levels detected for the mutant variants of *rex-39* templates are shown as the percentage (%) of SDC-2 binding to the wild-type *rex-39* template. The plot represents the average of three independent experiments, with error bars indicating SD. Statistical comparisons were calculated using the Student's *t* test.

(D) Comparison of DNA sequences for the three MEX motifs in wild-type *Cel rex-33* and the *Cbr* MEX motifs that replaced them. Also shown are sequences for the scrambled *Cel* MEX motifs used as negative controls. Schematics show keys for *rex* sites analyzed for *Cel* SDC-3 binding *in vivo* and *Cel* SDC-2 binding *in vitro*: wild-type *Cel rex-33* (black, MEX motifs), *Cel rex-33* with *Cbr* MEX motifs (purple), *Cel rex-39* with scrambled *Cel* MEX motifs (black outline).

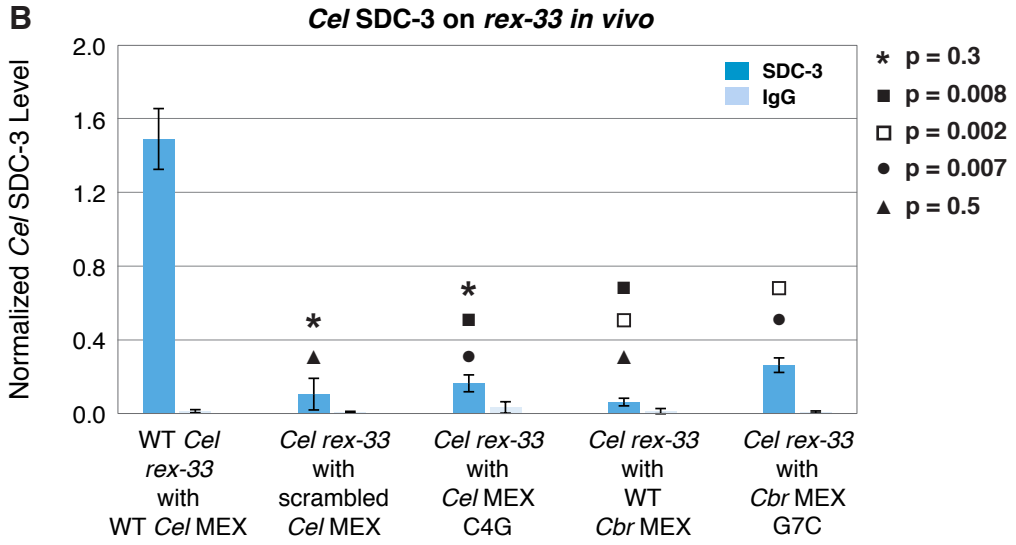
(E) Graph shows ChIP qPCR levels for *Cel* SDC-3 (dark blue) and control IgG (light blue) at wild-type *Cel rex-33* and mutant *rex-33* with *Cbr* MEX motifs *in vivo*. *Cel* SDC-3 binds to endogenous *Cel rex-33* sites with wild-type MEX motifs but not to mutant *Cel rex-33* sites with either scrambled *Cel* MEX motifs or *Cbr* MEX motif replacements. Details of the experiment and graph are the same as in (B).

(F) Graph of *in vitro* assay assessing *Cel* SDC-2 binding to a wild-type *Cel rex-33* DNA template and a mutant *rex-33* template with *Cbr* MEX motifs. *Cel* SDC-2 binds to the *Cel rex-33* template with wild-type MEX motifs but not to mutant *Cel rex-33* templates with either scrambled *Cel* MEX motifs or *Cbr* MEX motif replacements. *Cel* SDC-2 does not bind to the control template (beige). SDC-2 levels detected for the mutant variant *rex-33* templates are shown as the percentage (%) of SDC-2 binding to the wild-type *rex-33* template. The plot represents the average of three independent experiments, with error bars indicating SD. Statistical comparisons were calculated using the Student's *t* test.

**Figure 13**

**A** ***Cel rex-33***

| | | | |
|---|---|---|---|
| WT *Cel rex-33* with WT *Cel* MEX | TTTAAGCGAAGGGAC | ATTTGGCGCAGGGGG | GTATCGCGCAGGGAA |
| *Cel rex-33* with WT *Cbr* MEX | TATATGGGCAGGGAC | GAATGCGGCAGGGGG | TAATTGGGAAGGGAA |
| *Cel rex-33* with *Cel* MEX C4G | TTTAAGGGAAGGGAC | ATTTGGGGCAGGGGG | GTATCGGGCAGGGAA |
| *Cel rex-33* with *Cbr* MEX G7C | TATATGCGCAGGGAC | GAATGCCGCAGGGGG | TAATTGCGAAGGGAA |
| *Cel rex-33* with scrambled *Cel* MEX | TTTAGTCGACGTGAC | ATTGCGCTCGAGCGC | GTATGACGCGACGCT |



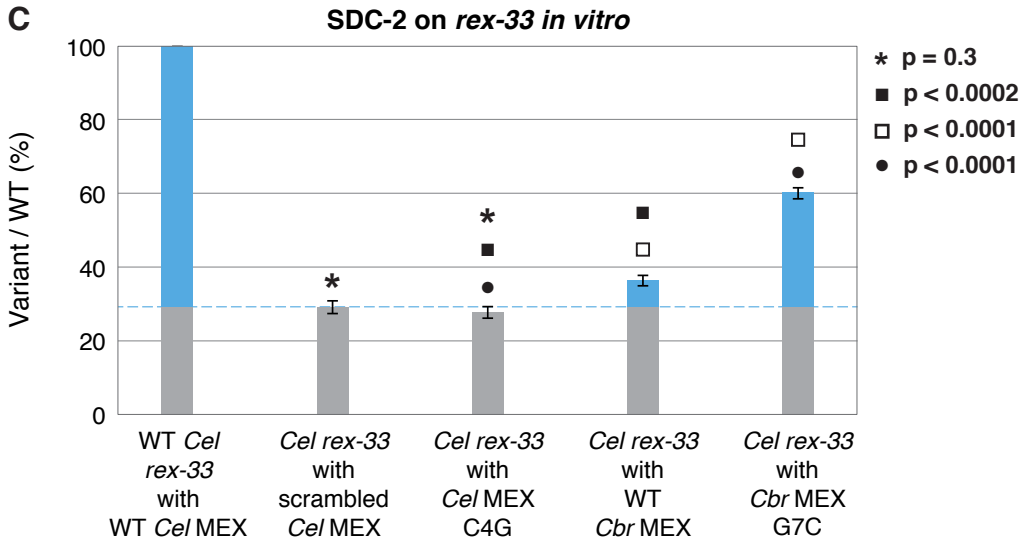



**Figure 13.** A nucleotide position in the consensus *Cbr* MEX motif can act as a critical determinant for whether *Cel* DCC binds *in vivo* and *in vitro*.

(A) Shown are DNA sequences of three wild-type or mutant *Cel* or *Cbr* MEX motifs within *Cel rex-33* assayed for *Cel* SDC-3 binding *in vivo* (B) and *Cel* SDC-2 binding *in vitro* (C). The ln(P) scores for the wild-type *Cel* MEX motifs in *rex-33* are -13.13, -15.33, and -15.35. The *Cel* ln(P) scores for the 3 substituted *Cbr* MEX motifs are all greater than -9. The three *Cbr* ln[P] scores for those substituted *Cbr* MEX motifs are -18.72, -12.26, and -12.58. The *Cel* ln(P) scores for the 3 *Cel* MEX motifs with the C4G change are -9.58, -11.20, and -11.26. The *Cel* ln(P) scores for the 3 *Cbr* MEX motifs with the G7C change are -12.20, -11.16, and -10.84. The *Cel* ln(P) scores for the *Cel rex-33* scrambled MEX motifs are all > -9.

(B) Graph shows normalized ChIP qPCR levels for C*el* SDC-3 (dark blue) and control IgG (light blue) *in vivo* at endogenous *Cel rex-33* with wild-type or mutant *Cel* MEX motifs and wild-type or mutant *Cbr* MEX motifs. Replacing the critical cytosine (red asterisk) in each of the three MEX motifs of endogenous *Cel rex-33* with a guanine (C4G) eliminates *Cel* SDC-3 binding, as does scrambling the three C*el* MEX motifs. Substituting three *Cbr* MEX motifs for *Cel* MEX motifs also severely reduces *Cel* DCC binding. Each *Cbr* MEX motif has a guanine instead of a cytosine in the critical location. Replacing the guanine with a cytosine (G7C) in each of the *Cbr* MEX motifs increased C*el* SDC-3 binding 4.2-fold, resulting in a *Cel* SDC-3 binding level representing 18% of that at wild-type *rex-33*. SDC-3 levels for each replicate were normalized to the average SDC-3 level at 7 control *rex* sites (*Cel rex-8*, *Cel rex-14*, *Cel rex-16*, *Cel rex-32*, *Cel rex-35*, *Cel rex-36*, and *Cel rex-48*). Error bars represent the standard deviation (SD) of three replicates. Statistical comparisons were calculated using the Student's *t* test.

(C) Graph of the *in vitro Cel* SDC-2 binding assay shows that replacing the critical cytosine (red asterisk) in each of the three MEX motifs of *Cel rex-33* with a guanine (C4G) eliminates *Cel* SDC-2 binding, as does scrambling the three MEX motifs. Substituting three *Cbr* MEX motifs for *Cel* MEX motifs severely reduces *Cel* DCC binding. Each *Cbr* MEX motif has a guanine instead of a cytosine in the critical location. Replacing the guanine with a cytosine (G7C) in each of the *Cbr* MEX motifs increases specific *Cel* SDC-2 binding 4.3-fold and restores it to 44% of that at the wild-type *rex-33* DNA template. SDC-2 levels detected for the mutant variants of *rex-33* templates are shown as the percentage (%) of SDC-2 binding to the wild-type *rex-33* template. The plot represents the average of three independent experiments, with error bars indicating SD. Statistical comparisons were calculated using the Student's *t* test.

**Supplementary File 1. List of alleles and strains used in this study**

| Strain | Species | Genotype | Description |
|---|---|---|---|
| TY5004 | *Cbr* | *mix-1(y435)* / + II | 460 bp deletion (Δ 5' UTR into 2nd intron, out of frame), null allele |
| TY5005 | *Cbr* | *dpy-27(y436)* / + III | 632 bp deletion (Δ 5' UTR, exon 1, intron 1, and part of exon 2), null |
| TY5006 | *Cbr* | *xol-1(y430)* X | 589 bp deletion (Δ promoter into 2nd exon, out of frame), null |
| TY5153 | *Cbr* | *dpy-27(y436)* / + III; *xol-1(y430)* X | |
| TY5230 | *Cbr* | *xol-1(y430) sdc-2(y453)* / + *xol-1(y430)* X | *sdc-2* null deletion allele created by genome editing |
| TY5231 | *Cbr* | *xol-1(y430) sdc-2(y454)* / + *xol-1(y430)* X | *sdc-2* null deletion allele created by genome editing |
| TY5232 | *Cbr* | *xol-1(y430) sdc-2(y455)* / + *xol-1(y430)* X | *sdc-2* null deletion allele created by genome editing |
| TY5237 | *Cbr* | *xol-1(y430) sdc-2(y460)* / + *xol-1(y430)* X | *sdc-2* null deletion allele created by genome editing |
| TY5363 | *Cbr* | *sdc-2(y467)* / + X | *sdc-2* null deletion allele created by genome editing |
| TY5365 | *Cbr* | *sdc-2(y469)* / + X | *sdc-2* null deletion allele created by genome editing |
| TY5753 | *Cel* | *dpy-27(y679)* III | 3xFlag-tagged *Cel dpy-27* |
| TY5773 | *Cbr* | *dpy-27(y705)* / + III | 52 bp deletion in exon 4 (Δ starts at codon 689, out of frame), null |
| TY5774 | *Cbr* | *dpy-27(y706)* III | 3xFlag-tagged *Cbr dpy-27* |
| TY5775 | *Cbr* | *sdc-2(y716)* X | 3xFlag-tagged *Cbr sdc-2* |
| TY5836 | *Cbr* | *dpy-27(y706)* III; *rex-1(y747)* X | *Cbr* MEX II (-27.58) scrambled at *Cbr rex-1* |
| TY5837 | *Cbr* | *dpy-27(706)* III; *rex-1(y749)* X | 5 motifs (4 *Cbr* MEX and 1 *Cbr* MEX II) scrambled in *Cbr rex-1* |
| TY5847 | *Cel* | *dpy-27(y679)* III; *yIs185 (Cbr rex-1)* X | *Cbr rex-1* insertion at *Cel* site 2 |
| TY5852 | *Cel* | *dpy-27(y679)* III; *yIs176 (Cbr rex-2)* X | *Cbr rex-2* insertion at *Cel* site 2 |
| TY5854 | *Cel* | *dpy-27(y679)* III; *yIs187 (Cel rex-32)* X | *Cel rex-32* insertion at *Cel* site 2 |
| TY5862 | *Cel* | *dpy-27(y679)* III; *yIs193 (Cbr rex-9)* X | *Cbr rex-9* insertion at *Cel* site 2 |
| TY5863 | *Cel* | *dpy-27(y679)* III; *yIs194 (Cbr rex-7)* X | *Cbr rex-7* insertion at *Cel* site 2 |
| TY5865 | *Cel* | *dpy-27(y679)* III; *yIs195 (Cbr rex-5)* X | *Cbr rex-5* insertion at *Cel* site 2 |
| TY5942 | *Cel* | *dpy-27(y679)* III; *yIs204 (Cbr rex-4)* X | *Cbr rex-4* insertion at *Cel* site 2 |
| TY5945 | *Cbr* | *sdc-2(y716) rex-4(y799)* X | *Cbr rex-4* MEX (-13.8) scrambled |
| TY5975 | *Cbr* | *sdc-2(y716) rex-4(y824)* X | *Cbr rex-4* MEX (-13.8) scrambled and MEX II (-19.09) scrambled |
| TY5976 | *Cbr* | *sdc-2(y716)* rex-4(y825) X | *Cbr rex-4* MEX II (-19.09) scrambled |

| | | | |
|---|---|---|---|
| TY6075 | *Cbr* | *sdc-2(y716) rex-3(y849)* X | *Cbr rex-3* MEX II (-12.36) scrambled |
| TY6076 | *Cbr* | *sdc-2(y716) rex-3(y850)* X | *Cbr rex-3* MEX II (-20.04) scrambled |
| TY6121 | *Cbr* | *sdc-2(y716) rex-3(y868)* X | *Cbr rex-3* MEX II (-12.36) scrambled and MEX II (-20.04) scrambled |
| TY5946 | *Cbr* | *sdc-2(y716) rex-7(y800)* X | *Cbr rex-7* MEX (-18.72) scrambled |
| TY6072 | *Cbr* | *sdc-2(y716) rex-7(y846)* X | *Cbr rex-7* MEX (-18.72) scrambled and MEX (-12.26) scrambled |
| TY6085 | *Cbr* | *sdc-2(y716) rex-7(y855)* X | *Cbr rex-7* MEX (-18.72) scrambled and MEX (-12.58) scrambled |
| TY6086 | *Cbr* | *sdc-2(y716) rex-7(y857)* X | *Cbr rex-7* MEX (-18.72) scrambled, MEX (-12.26) scrambled, and MEX (-12.58) scrambled |
| TY4573 | *Cel* | *sdc-2(y74)* X; *yEx992* | Expression of *3xflag::sdc-2* from an extrachromosomal array for the *in vitro* assay |
| TY6122 | *Cel* | *rex-33(y869)* X | 3 *Cel* MEX motifs replaced by *Cbr* MEX motifs in *Cel rex-33* |
| TY6123 | *Cel* | *rex-33(y870)* X | 3 *Cel* MEX motifs scrambled |
| TY6106 | *Cel* | *rex-39(y861)* X | 2 *Cel* MEX II motifs replaced by *Cbr* MEX II motifs in *Cel rex-39* |
| TY5759 | *Cel* | *rex-39(y686)* X | 2 *Cel* MEX II motifs scrambled |
| AF16 | *Cbr* | wild-type *C. briggsae* | |
| JU935 | *Cbr* | *mfIs27(Ce-lip-1::gfp, Ce-myo-2::gfp)* X | Used to determine parental origin of *Cbr* X chromosome |

**Supplementary File 2. List of primers**

| Target | Figure | Location | Primer name | Sequence | Function |
|---|---|---|---|---|---|
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | A | QY171 | GATGATGAACTAAATCGTAAGCTTCC | qPCR for DCC binding |
| | | B | QY172r | CAGGGAAGATTAACTTGAAACTTCAG | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | C | QY173 | GCCTCAGGTCTTACGGTAGAAG | qPCR for DCC binding |
| | | D | QY174r | CTCAGAGACTTTTTGTACATTGTATTTG | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | D | QY175 | CAAATACAATGTACAAAAAGTCTCTGAG | qPCR for DCC binding |
| | | E | QY176r | CTAGCTTGCACATCAAGAAGAC | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | E | QY157 | GTCTTCTTGATGTGCAAGCTAG | qPCR for DCC binding |
| | | F | QY177r | CACGTTTCTATTAAACATTTCCTC | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | F | QY010 | GAGGAAATGTTTAATAGAAACGTG | qPCR for DCC binding |
| | | G | QY011r | CTTTGCATATGTCCTTTCACG | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | G | QY178 | CGTGAAAGGACATATGCAAAG | qPCR for DCC binding |
| | | H | QY156r | GCTATTCGACAAACACTCCACAC | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | H | QY179 | GTGTGGAGTGTTTGTCGAATAGC | qPCR for DCC binding |
| | | I | QY180r | CCCAATATGTTCCGTTTCTTACTG | qPCR for DCC binding |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | J | QY181 | GGTTACCTAACGGAAATCCTGTG | qPCR for DCC binding |
| | | K | QY182r | GATCGTAAATGCACACATGCATTC | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | A | QY121 | GGACTGTGCTCTGGCG | qPCR for DCC binding |
| | | B | QY122r | GGGACCATGGTTACTTTTCTTG | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | C | QY123 | GAAAAGTTAACGCTCCGCTC | qPCR for DCC binding |
| | | D | QY124r | GTCAGTTGACCTTACTCATTCAG | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | D | QY125 | CTGAATGAGTAAGGTCAACTGAC | qPCR for DCC binding |
| | | E | QY126r | GAATGCCATACGATGTCTGAC | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | E | QY012 | GTCAGACATCGTATGGCATTC | qPCR for DCC binding |
| | | F | QY013r | GTAGCAGGCCACTAGTTTCC | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | F | QY127 | GGAAACTAGTGGCCTGCTAC | qPCR for DCC binding |
| | | G | QY128r | CTATCTTCGCAGAAAGTCTGAC | qPCR for DCC binding |

| | | | | | |
|---|---|---|---|---|---|
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | G | QY129 | GTCAGACTTTCTGCGAAGATAG | qPCR for DCC binding |
| | | H | QY130r | CCATAGAACATAGTTCCTGGTTC | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | H | QY131 | GAACCAGGAACTATGTTCTATGG | qPCR for DCC binding |
| | | I | QY132r | GAAAACATTGCGAAGACTCAAC | qPCR for DCC binding |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | J | QY133 | CTTTGGAAAGTCAGTTCCTC | qPCR for DCC binding |
| | | K | QY134r | CATGAATAGTATGTGCAGTGATG | qPCR for DCC binding |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | A | QY135 | GATGTTGCTCTATTCAAAATGCG | qPCR for DCC binding |
| | | B | QY136r | CATAGATGCGGGATTTTTTGTG | qPCR for DCC binding |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | C | QY018 | CATTGCAATAAACTGGTGGG | qPCR for DCC binding |
| | | D | QY019r | GCAGGGGATTAAGACAACATT | qPCR for DCC binding (MEX -12.26 wt) |
| | | D | QY190r | ACGCGAGATTAAGACAACATT | qPCR for DCC binding (MEX -12.26 scr) |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | D | QY137 | AATGTTGTCTTAATCCCCTGC | qPCR for DCC binding (MEX -12.26 wt) |
| | | D | QY191 | AATGTTGTCTTAATCTCGCGT | qPCR for DCC binding (MEX -12.26 scr) |
| | | E | QY138r | GACTTGTAGAATCCTTTTTATCGC | qPCR for DCC binding |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | E | QY139 | GCGATAAAAAGGATTCTACAAGTC | qPCR for DCC binding |
| | | F | QY140r | TAACACGTCTCCTATCACTC | qPCR for DCC binding |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | G | QY141 | GGTTTTATGGCCGTGGTG | qPCR for DCC binding |
| | | H | QY142r | GCTATTCGAACGTCGAACAG | qPCR for DCC binding |
| *Cbr rex-1* | Figure 9, Figure 9—Figure supplement 1, Figure 10, Figure 10—Figure supplement 1, Figure 11, Figure 11—Figure supplement 1 | | QY006 | CCCTTCCACTCTAGTCTAATCG | qPCR for DCC binding normalization |
| | | | QY007r | GGTGTGTTTGATGATGTAGGC | qPCR for DCC binding normalization |
| *Cbr rex-2* | Figure 9, Figure 9—Figure supplement 1, Figure 10, Figure 10—Figure supplement 1, Figure 11, Figure 11—Figure supplement 1 | | QY040 | CAAATTTGATCGAGTCAACCTC | qPCR for DCC binding normalization |
| | | | QY041r | GAAAAGGAGAGTTATCACTCAATG | qPCR for DCC binding normalization |

| | | | | |
|---|---|---|---|---|
| *Cbr rex-5* | Figure 9, Figure 9—Figure supplement 1, Figure 10, Figure 10—Figure supplement 1, Figure 11, Figure 11—Figure supplement 1 | QY014 | CGAAGAAAGCATATGAAAGC | qPCR for DCC binding normalization |
| | | QY015r | CTCTAAAATAATTGTCCTCCGTC | qPCR for DCC binding normalization |
| *Cbr rex-9* | Figure 9, Figure 9—Figure supplement 1, Figure 10, Figure 10—Figure supplement 1, Figure 11, Figure 11—Figure supplement 1 | QY022 | GATACGAACAGGGTGCAAGG | qPCR for DCC binding normalization |
| | | QY023r | TCACATACTCGTTTCGTCCG | qPCR for DCC binding normalization |
| *Cbr* X negative control | Figure 9, Figure 9—Figure supplement 1, Figure 10, Figure 10—Figure supplement 1, Figure 11, Figure 11—Figure supplement 1 | QY099 | CCAGAAATAGCTATTCTAAGAGG | qPCR for negative control |
| | | QY100r | GTTTTTGAGTTCCCTGGCAC | qPCR for negative control |
| *Cel rex-8* | Figure 5, Figure 12 | rex8-F | TTTATCCACCAACATGCATAAG | qPCR for DCC binding normalization |
| | | rex8-R | CAGTGGATAACTACACAAGGG | qPCR for DCC binding normalization |
| *Cel rex-14* | Figure 12 | rex14-F | ACCTCCTTTCACAACACTCTTT | qPCR for DCC binding normalization |
| | | rex14-R | TCGAACCCAACTCGTTTATCTC | qPCR for DCC binding normalization |
| *Cel rex-16* | Figure 5, Figure 12 | rex16-F | GTACAAACGCAGGGAAGAGA | qPCR for DCC binding normalization |
| | | rex16-R | GACGCTACCACACCTTCAATA | qPCR for DCC binding normalization |
| *Cel rex-32* | Figure 5, Figure 12 | rex32-F | CACTCCCCAGCTAATTTGGA | qPCR for DCC binding normalization |
| | | rex32-R | TTCCCTTGTTGCGGAGATAG | qPCR for DCC binding normalization |
| *Cel rex-33* | Figure 12 | QY212 | GTGTGTTGCTGCCAAAGCCTG | *Cel rex-33* mutagenesis genotyping, sequencing |
| | | QY243 | GCAAGCACAGACACTCAAAC | qPCR for DCC binding |
| | | QY213r | GGGCCCGTGGTTAATTTATTCG | *Cel rex-33* mutagenesis genotyping; sequencing; qPCR for DCC binding |
| *Cel rex-35* | Figure 5, Figure 12 | rex35-F | CCATATGTTGCCCAATGTTCC | qPCR for DCC binding normalization |
| | | rex35-R | CGCAGGGAACATCAAATTAGTC | qPCR for DCC binding normalization |
| *Cel rex-36* | Figure 12 | rex36-F | CCCTCTTCAGGCGATAAATG | qPCR for DCC binding normalization |
| | | rex36-R | CGTTCATGCGAATGTCTCTC | qPCR for DCC binding normalization |

| | | | | |
|---|---|---|---|---|
| *Cel rex-39* | Figure 12 | QY210 | CGATACATTTGTTTTTTATTAAATATCTACATTTCTCG | *Cel rex-39* mutagenesis genotyping; sequencing |
| | | QY211r | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAATG | *Cel rex-39* mutagenesis genotyping; sequencing |
| | | QY215 | AATGCACTCATGCACATGTTTC | qPCR for DCC binding |
| | | QY216r | CACAACAAGACCGAATAAATATAACAC | qPCR for DCC binding |
| *Cel rex-48* | Figure 5, Figure 12 | rex-48-F | CTGCGCGATAGGCAATAGT | qPCR for DCC binding normalization |
| | | rex-48-R | GCACAATTCCAAGTCATCCATAC | qPCR for DCC binding normalization |
| *Cel* site 2 | Figure 5 | ER589 | CAGCGTAGTTGCTGACACTTAATGGTTC | qPCR for DCC binding normalization |
| | | ER590 | CTTTTAAGCAGTCGTCATGTACGTGTTCG | qPCR for DCC binding normalization |
| *Cel* Chr I control | Figure 5, Figure 12 | autosome-F | ACCCACGACATTGCTCTTGT | qPCR for DCC binding normalization |
| | | autosome-R | AGTTTTGGGGCAGCTCTCTC | qPCR for DCC binding normalization |
| *Cel* X site 2 | Figure 5 | ER573 | CGTGCCAGTTGTTGACTTATG | *Cel* X site 2 insertion genotyping; sequencing |
| | | ER574 | CATGTTTTTGGCGCTGGTGAGTAGG | *Cel* X site 2 insertion genotyping; sequencing |
| *Cbr ben-1* | | BF-2041 | GCTCGCTTTCTTTCCAAAAACGAGCAGAAGCCCCAATCGGTCG | Cas9 co-conversion marker for *Cbr* |
| | | BF-2042 | CGTGCGCAGCTTGTGATTCATGCTCCGCCCACTTTTCCG | Cas9 co-conversion marker for *Cbr* |
| *Cbr dpy-27* | Figure 1F | CBDPY27.OL | GACGACAGAGTGGCTCTGCCGACAAGAGC | *Cbr* deletion library screening |
| | | CBDPY27.IL | GCCAACTTGCCGAATTTGAGC | *Cbr* deletion library screening |
| | | CBDPY27.PL | GGAGCTGTTGGAAGACTCGAGTGGTTGG | *Cbr* deletion library screening |
| | | CBDPY27.OR | CTTACAATGTCTTCAATCTGTTGGAAAAG | *Cbr* deletion library screening |
| | | CBDPY27.IR | GGCCATTTTGATCGTCGTTGTGG | *Cbr* deletion library screening |
| | | CBDPY27.PR | CCAGACGTCAATCTCAGCGATGAC | *Cbr* deletion library screening |
| *Cbr xol-1* | Figure 2 | CBXOL1.OL | GCCTAGTTTCACGTATTTCTCTAC | *Cbr* deletion library screening |
| | | CBXOL1.IL | GTAAGGCCAACCGGATTAGC | *Cbr* deletion library screening |
| | | CBXOL1.PL | CGCTTCAAGGAGACGCCGAGC | *Cbr* deletion library screening |
| | | CBXOL1.OR | CCCCGTGAAAAGAGTCTGCC | *Cbr* deletion library screening |
| | | CBXOL1.IR | CGGCACTTCTGGGTTTAGACG | *Cbr* deletion library screening |
| | | CBXOL1.PR | CGCATGTTCCTATGCAAACTTTGGC | *Cbr* deletion library screening |

| | | | |
|---|---|---|---|
| *Cel dpy-10* | BF-1853 | CGAACGTTCTCGCTGACAACGAACTATTCGCGTCAG | Cas9 co-conversion marker for *Cel* |
| | BF-1854 | GCATGTTTGATTTGGAGTAGTTCCTGGCATTCC | Cas9 co-conversion marker for *Cel* |

**Supplementary File 3. Chromosome-specific BACs used to generate FISH probes**

| BAC | Chromosome | Start | End |
|---|---|---|---|
| RPCI94_19F11 | III | 35266 | 135375 |
| RPCI94_21C16 | III | 53199 | 124610 |
| RPCI94_27L20 | III | 241669 | 341069 |
| RPCI94_27P10 | X | 4191 | 96616 |
| RPCI94_03E18 | X | 217136 | 321081 |
| RPCI94_01B13 | X | 237728 | 343819 |
| RPCI94_20J22 | X | 720980 | 839475 |
| RPCI94_28F15 | X | 1198052 | 1322029 |
| RPCI94_19L23 | X | 3179427 | 3291372 |
| RPCI94_19O24 | X | 14257820 | 14399879 |
| RPCI94_26I06 | X | 15980691 | 16042072 |
| RPCI94_28L18 | X | 16743266 | 16871287 |
| RPCI94_22H01 | X | 19865754 | 19995983 |

**Supplementary File 4. List of target-specific sequences for guide RNAs used in CRISPR / Cas9 genome editing experiments**

| Target | Figure | Target sequence (5' to 3') | Coordinates | Guide name |
|---|---|---|---|---|
| *Cbr dpy-27* | Figure 1F | CGCTCTGGAGTACGGTAAAA | III: 2729405..2729386 | *cbr-dpy-27* |
| *Cbr ben-1* | Figure 1, Figure 9, Figure 10, Figure 11 | CAACCTGATGGAACCTACAA | III: 8377724..8377705 | crispr_bf39 |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | GCGCGTATGGCCAATTGGCA | X: 6358568..6358587 | crispr_bf80 |
| *Cbr rex-4* | Figure 9, Figure 9—Figure supplement 1 | TTTTATATGAACAGGGTGCG | X: 6358623..6358642 | crispr_bf77 |
| *Cbr rex-7* | Figure 9, Figure 9—Figure supplement 1 | TCTGAGATTTTATATGGGCA | X: 8026313..8026332 | crispr_bf79 |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | CGAAGAGAAGAATGCGGCAG | X: 8026442..8026423 | crQY011 |
| *Cbr rex-7* | Figure 11, Figure 11—Figure supplement 1 | AATTTAAGTAATTGGGAAGG | X: 8026448..8026467 | crQY015 |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | GTAGCTAACTCTGTGAAAAT | X: 19468573..19468554 | crQY014 |
| *Cbr rex-3* | Figure 10, Figure 10—Figure supplement 1 | GCGTTGTGGAAGTAGGCAGG | X: 19468752..19468771 | crQY013 |
| *Cel dpy-10* | Figure 5, Figure 8 | GCTACCATAGGCACCACGAG | II: 6711193..6711212 | crispr_bf32 |
| *Cel rex-33* | Figure 12 | TGCCCTACTAAATAAGCGAA | X: 6296617..6296598 | crQY016 |
| *Cel rex-39* | Figure 12 | ACATGTGGAGAACATTATTT | X: 14813548..14813529 | crQY017 |
| *Cel* site 2 | Figure 5 | TTATGTAGTCTCTTTCAGTG | X: 15574657..15574676 | CS568 |

**Supplementary File 5. DNA sequences of repair templates used in CRISPR / Cas9 genome editing experiments**

| Target | Description | Figure | Sequence of repair template (5' to 3') | Related guide | Repair ID |
|---|---|---|---|---|---|
| *Cbr ben-1* | *Cbr* co-injection marker | | GGGAAGTGATTTCCGACGAGCACGGAATTCAACCTGATGGAACCTACATAT GGTGGAGAGAGTGACTTGCAGCTCGAGCGCATCAATGTCTACTACAACG | crispr_bf39 | BF-2036 |
| *Cbr rex-4* | *Cbr rex-4* MEX (-13.8) scrambled | Figure 9, Figure 9—Figure supplement 1 | GGCCAATTGGCATGGGCTGCCTGCTAACCTTTCCCTGCCTACGCATATTTGATAAGTGATGACTGCGCGGACAAAAGAGGGAAACTAGTGGCCTGCTACCCGAGAAAGAGAGA | crispr_bf77 | BF-2470 |
| *Cbr rex-4* | *Cbr rex-4* MEX II (-19.09) scrambled | Figure 9, Figure 9—Figure supplement 1 | GCATTCCTCAACCCGCAAAGAGAAGTCAATCGCGCAGATATTGTAATTGTTGCTGCTGCAGTCACTGTGCTCGCTCTGACTGCCTGCTAACCTTTCCCTGCCTACGCATATTTTATATGAACAGGGTGCG | crispr_bf80 | QY070t |
| *Cbr rex-4* | *Cbr rex-4* MEX II (-19.09) scrambled and MEX (-13.8) scrambled | Figure 9, Figure 9—Figure supplement 1 | GCATTCCTCAACCCGCAAAGAGAAGTCAATCGCGCAGATATTGTAATTGTTGCTGCTGCAGTCACTGTGCTCGCTCTGACTGCCTGCTAACCTTTCCCTGCCTACGCATATTTGATAAGTGATGACTGCG | crispr_bf80 | QY071t |
| *Cbr rex-7* | *Cbr rex-7* MEX (-18.72) scrambled | Figure 11, Figure 11—Figure supplement 1 | CGTCCTGTTCATTGCAATAAACTGGTGGGAGTTTTTCCAATCTGAGATTTGTGAGCAGTAGTGACAACAGTGATTTAATTTTATGTTCATGAAGTTTTCAGGTTTTTTGCAT | crispr_bf79 | BF-2472 |
| *Cbr rex-7* | *Cbr rex-7* MEX (-12.26) scrambled, and MEX (-12.58) scrambled | Figure 11, Figure 11—Figure supplement 1 | TCAGGTTTTTTGCATAAAATGCACAATATTCTGAGAAATGTTGTCTTAATCTCGCGTCGTCGTGTTCTCTTCGCACGCAATTTAAGAGTCAGTGTCAGAGGAGAAGACAAATTTGAGGGACCTCTTCTCTTATTTTTTTTTCGCAAAGT | crQY011, crQY015 | QY087t |
| *Cbr rex-3* | *Cbr rex-3* MEX II (-20.04) scrambled | Figure 10, Figure 10—Figure supplement 1 | GCAAAGTCAACATGTGTATTTTCCGTGAAAGGACATATGCAAAGGGGTGTCTGTCGAGCTCGCTGTGCTGACGAGAGCAGCTCATAGAGCGTAAATGGGCATTGCCCTCCGCGCAGATACGCGCGTTAAGCCATACCACACATATAACC | crQY013 | QY154t |
| *Cbr rex-3* | *Cbr rex-3* MEX II (-12.36) scrambled | Figure 10, Figure 10—Figure supplement 1 | TTGATCACACAGTCTTCTTGATGTGCAAGCTAGCTATTTCGAGTAGTTGGAAAATCAAAATTCTATAGTAATACTGTGATAATCATACATCTGTATAGCTACGATGATTTTGAGGAAATGTTTAATAGAAACGTGAAAAAAAGAAATAT | crQY014 | QY155t |

| | | | | | |
|---|---|---|---|---|---|
| *Cel dpy-10* | *Cel* co-injection marker | | CACTTGAACTTCAATACGGCAAGATGAGAATGACTGGAAACCGTACCGCAT<br>GCGGTGCCTATGGTAGCGGAGCTTCACATGGCTTCAGACCAACAGCCTAT | crispr_bf32 | BF-1813 |
| *Cel rex-33* | 3 *Cel* MEX motifs replaced by *Cbr* MEX motifs in *Cel rex-33* | Figure 12 | CGTTCAAACAGTCTTTCCTGCAAGCACAGACACTCAAACGTGAGTAATTAT<br>TATATGGGCAGGGACACCCAATCGATTGCCCATTTACGAATGCGGCAGGGG<br>GTCACCATAGATAGTAATTGGGAAGGGAAGATTTACCGCCTTTCGCTTATT<br>TAGTAGGGCACGCAAATTAGTATGCTT | crQY016 | QY199t |
| *Cel rex-33* | 3 *Cel* MEX motifs scrambled | Figure 12 | CGTTCAAACAGTCTTTCCTGCAAGCACAGACACTCAAACGTGAGTAATTAT<br>TTTAGTCGACGTGACACCCAATCGATTGCCCATTTACATTGCGCTCGAGCG<br>CTCACCATAGATAGGTATGACGCGACGCTGATTTACCGCCTTTCGCTTATT<br>TAGTAGGGCACGCAAATTAGTATGCTT | crQY016 | QY214t |
| *Cel rex-39* | 2 *Cel* MEX II motifs replaced by *Cbr* MEX II motifs in *Cel rex-39* | Figure 12 | TTTCATAACACAACAAGACCGAATAAATATAACACTTATTCCTCCTGCCTA<br>CTTCCACAACGCGCGAAGTAACAAGTAACTACTTCGCGCGTTGTGGAAGTA<br>GGCAGGAGGATTTTCAAGAAACATGTGCATGAGTGCATTTCAAAATTC | crQY017 | QY208t |
| *Cel* X site 2 | *Cel* X site 2 insertion of *Cbr rex-2* | Figure 5 | GAACATGTCGAACACGTACATGACGACTGCTTAAAAGTTGAAAATTTCCCA<br>TATCCGTTTCTCATTTTATGTAGTCTCTTTCAGTAATCGTATTCAGCACGT<br>TCGAAAGTACCTTTGTACAAATTTTGAGCTAATTCTCAACGCCCCTCTGAA<br>AACACTTCCCTTGTGAGTTTGAACGGTTTCAGTACAACCATATGGTCAGGG<br>GAACTAAAAAACTAGAAATTCATTACTCGAACATACTGTAGTTATCCCACC<br>ATCGCAAATTTGATCGAGTCAACCTCTGCGAAAACGCAAAATAGAAAGGAC<br>CACCACACACAAAACGCCCACGTAAACACTGCCCCCTTCCGAGATAAAACA<br>TTGAGTGATAACTCTCCTTTTCCGTTTTTCTGAGCGTTTCGCATTTTGGCA<br>CGGATCAGTTTCTAATCCACAACTTTAAAAAAATCAAAAATTTTCTTCGAA<br>ATTCGAAAGAAAATAAGGAGATTTTTTGACAAGTGAAAAATGAACTCATTC<br>AGTAAGAACGCATATTGTTTCTCAATATTTCTTTTCTATCGTGAAAACGCT<br>TCAACAATCGTTACAAAACTTCATTGGCTACTAACGATTTTGCAATTTTAC<br>ATTATACTTTGTTTGTGAGTTTCAGGAAACTTGTGAATTCGTTCAAACCCT<br>TCAGAATCAG | CS568 | ER567 |
| *Cel* X site 2 | *Cel* X site 2 insertion of *Cel rex-32* | Figure 5 | GTCGAACACGTACATGACGACTGCTTAAAAGTTGAAAATTTCCCATATCCG<br>TTTCTCATTTTATGTAGTCTCTTTCAGGAGACAAAAAGGAAAAAATGAAAT<br>TGGACAACCCCATAAATTTCATGAATTTTTAAAACTTCTTGCAGGAATATA<br>AATTACTCTAATTAAAGTTTTTTTGTTTGAAAATTTTTTGATAGGCCCAAA<br>TACATACTTATCTCTAAAAAAATTACTTTTGAATTCGTTCATTCAATGTGT<br>AAAATCTAAACAAAAGTGACCCCCCTTGTCCAAATATTTTATCCACGAGAC<br>TAGGTACACCTCCCTTCGCGATAAATAATTGGTACATCATTTTATCCACAG | CS568 | ER577 |

| | | | | | |
|---|---|---|---|---|---|
| | | | GGCTACTTCCTCCCTGCGCGATAAATTTAAAATTTTATAACTCTTTGGGTA<br>ACTAATTTTGGCAATATTTTCTTATTTTTTCACATCAACAAAAATATTTCA<br>AGTCTAAAAGTTGCAATAAATGCAGTTCAGAGGAAAATTGAGTCGTGTGCG<br>AATAACGAGGAAAAAGACAGACCCATACTTCCTCCCTGCGCGATACGATCT<br>CTATCGACTTTTCTGGTTTTATTGTTTGGCAATGTTTATTAAATTACTCCA<br>AAGATCAAGAAAACAATTATGAAAATTCGAGGAGGTGTAAAATAGAAAATG<br>TTGCAGTTGTGGCTACTAACGATTTTGCAATTTTACATTATACTTTGTTTG<br>TGAGTTTCAGGAAACTTGTGAATTCGTTCAAACCC | | |
| *Cel* X site 2 | *Cel* X site 2 insertion of *Cbr rex-9* | Figure 5 | GTCGAACACGTACATGACGACTGCTTAAAAGTTGAAAATTTCCCATATCCG<br>TTTCTCATTTTATGTAGTCTCTTTCAGGCAAATTTGAGCTTTCACCTTGAT<br>CTCAATTGTACATAATATTTCATGACATTTTTTGTATAAATGTGTTTTCTC<br>ATCACTCGATTTTCTTCTGCCAAAAAATAGAGCACTCCATTCCAAAAATAG<br>TATGTCTACGACATTCTCCGCTAATTGTAGTCTTCTGAACACCTCCTTTCG<br>CATGTAAGACGCTGATGGATATAAGATACGAACAGGGTGCAAGGACCCGCG<br>CACGTGCCTTGAATAGACGCTGTTAAAAAGGGCAGACGGCCAGATGGAAGA<br>CGTTTCGGAGACAGCGAGGCGGACGAAACGAGTATGTGAGGCCCATTACAA<br>CGTCTAATCCATTGGAGGAGAGAGGTTTGCAAAGGGGTGACTGGGGCGACC<br>AGCATTTTTATGTTTGATGGATGTGACCGGGAAAATGACGGGTTGTCATCA<br>GTGCAAGGTGACACAAAAAAACCGACAGTAGAATGGTAGTTTTTTTTTGCA<br>ATTTAAACGTTATCCATCATATTACGGTAGTGGAGGAGTAGTGACACCGCT<br>AAATGCATCTGATAAGTTTTATCAGGGTAGTCAAGATGATTTTTGCAACAA<br>ATTTTGAACTGTACTTTGTGGCTACTAACGATTTTGCAATTTTACATTATA<br>CTTTGTTTGTGAGTTTCAGGAAACTTGTGAATTCG | CS568 | ER581 |
| *Cel X site 2* | *Cel* X site 2 insertion of *Cbr rex-7* | Figure 5 | GTCGAACACGTACATGACGACTGCTTAAAAGTTGAAAATTTCCCATATCCG<br>TTTCTCATTTTATGTAGTCTCTTTCAGGTCAGTTGATCACTTTCTGATAAT<br>TCTTACATGAACCAAAACAAATTTGCTTGAAATTGAAAGTTGAACTGCTCA<br>TTTCATTAGAGTGTTTACGCTATTTTTCGAAAATTTGTTTAAAAGTATACT<br>TGAAACGTAGCAAAAAGCTTACTGAAAGAAAACGGAAATGTTTCATTTTAA<br>TTGATTTTGTCCGGTGTTTTTGTTAAATCAGCTTTATCGTTTCCTGAAATT<br>TTCGAGATTTGCCCGACCGTCCTGTTCATTGCAATAAACTGGTGGGAGTTT<br>TTCCAATCTGAGATTTTATATGGGCAGGGACAACAGTGATTTAATTTTATG<br>TTCATGAAGTTTTCAGGTTTTTTGCATAAAATGCACAATATTCTGAGAAAT<br>GTTGTCTTAATCCCCTGCCGCATTCTTCTCTTCGCACGCAATTTAAGTAAT<br>TGGGAAGGGGGAGAAGACAAATTTGAGGGACCTCTTCTCTTATTTTTTTTT<br>CGCAAAGTAGCAGCAATGAATTTTGCGATAAAAAGGATTCTACAAGTCGTT<br>TTCTATTTTCCTTCCTTTTCCAGTGGAAACTCTTCAAACAGACCAAATATT<br>TTGCACTCTGTTGTCATGTGGCTACTAACGATTTTGCAATTTTACATTATA<br>CTTTGTTTGTGAGTTTCAGGAAACTTGTGAATTCGTTCAAACCC | CS568 | ER586 |

| | | | | | |
|---|---|---|---|---|---|
| *Cel X site 2* | *Cel* X site 2 insertion of *Cbr rex-4* | Figure 5 | GTCGAACACGTACATGACGACTGCTTAAAAGTTGAAAATTTCCCATATCCG<br>TTTCTCATTTTATGTAGTCTCTTTCAGGTATGAAATTTGAAGTGTTTGAAT<br>ATATACTCTCTGCACTCCGACTATTTTACAGTGCGACTGGCAAATCTCACT<br>GAATGAGTAAGGTCAACTGACAATCAACAAAGATATAATTCCAAAGTTTTC<br>TCATTTCTTGAATATAAGTATTTGATGTTTGAAAAAATCGGCATTTCTTGC<br>AAATGTACTGAATGAATTGTAATCTTGGTCAGACATCGTATGGCATTCCTC<br>AACCCGCAAAGAGAAGTCAATCGCGCAGATATTGTAATTGTTCTCTGCGCG<br>TATGGCCAATTGGCATGGGCTGCCTGCTAACCTTTCCCTGCCTACGCATAT<br>TTTATATGAACAGGGTGCGCGGACAAAAGAGGGAAACTAGTGGCCTGCTAC<br>CCGAGAAAGAGAGAAATGCAACGTTTAATAAACCGATGACGAGCAGGCAAC<br>GTGCCCCTTCCGATTTGAATGGCGTTTCAAAATCAGAGTCAGACTTTCTGC<br>GAAGATAGTTTTTGTAAGCGCTTCGAGGGCAATGGAATCCTAAAATGATCA<br>ATTTTAATGAAATCGGGATGTAATGTAGGTAGAAATCTAGATCTACGTAGA<br>ACCAGGAACTATGTTCTGTGGCTACTAACGATTTTGCAATTTTACATTATA<br>CTTTGTTTGTGAGTTTCAGGAAACTTGTGAATTCG | CS568 | ER587 |
| *Cel X site 2* | *Cel* X site 2 insertion of *Cbr rex-5* | Figure 5 | GTCGAACACGTACATGACGACTGCTTAAAAGTTGAAAATTTCCCATATCCG<br>TTTCTCATTTTATGTAGTCTCTTTCAGGTAGTTTTTCGAAAGAACCGCCCA<br>AATTTTGAACTAGACTTGAAAATATTTTCGCGCGTTTAAAAACTTCATTTA<br>CTCAGAGTTACATCTTTCAACTTTACGAAGAAAGCATATGAAAGCGTTTTC<br>ACGTCTCGTTTCTCTAAATCTCCATTTCCATTTTTGATCCTTTCTCGTGCG<br>GATTTTCTCAACGCCCAAAGAAGTGAAAAATATTTCTTTGAAAAAGAAAAA<br>TAGACATTGACGGAGGACAATTATTTTAGAGAAAAACCAACTAACTCTACG<br>AAAAGGTTATATAGGCAAGCATATCAAAAATCAGATTTACATCAAAATTGC<br>AGAACAAAAGCAGAGAAAATCTGGTTCAACGGGATGCCAAGAATTTTCTCT<br>ATGTGGGTTTTCAGTTTCTTAAAAGGGATCAACTTCAAATATTATAGAAAC<br>CATTTGAGCTTTTAACTTTTCCATAATTGGTCATTGAAAGTTTGCTTGATT<br>ATTAAAAATCAAAAACCAAATGACTTTCTTGTACGGCTTTCATTCCGTCAG<br>AACCCTAATGAAAATATAGAATTTATAAAAGTACATTTATTAGGTTGGAAG<br>AAAAGTAACTGTCCACTGTGGCTACTAACGATTTTGCAATTTTACATTATA<br>CTTTGTTTGTGAGTTTCAGGAAACTTGTGAATTCG | CS568 | ER588 |

**Supplementary File 6. DNA templates used for *in vitro* DCC binding assays**

| Figure | ID | Description | Forward Primer | Forward Primer Sequence | Reverse Primer | Reverse Primer Sequence | Probe Sequence |
|---|---|---|---|---|---|---|---|
| Figure 12C | pKB460 | WT *rex-39* (reverse complement) | kb416 | CGATACATTTGTTTTTATTAAATATCTACATTTTCTCG | kb417r | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAAATG | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAAATGAAAATGTTAACCTTAGCTTAAACTGGCAAGACCTCAATTCCTGTTCAGCATGAAAAACTCTTCGCATTATGAACTTTCGTTTCCGAATGTTATCCTGGTATGCTTCCACTCGAGTGTTACACTTTTCCATGTTCTATTATAATACAATTGAATTTTATTACTTCATGTCAAAAAACTCACAACTAAATCAAACTCTTCAACCAGCTGTCTCAATTTCTCAACTCGCCGTTTTTGAAAGAAATCCTGCAATAAATCCGTAAAATTTTGCCTAATAAGTCGAATTTCAGCAGCAAGGTTGTCGTCAAATGATGCTTTGCAGGCTGAAAATAAGCTTTTAGAAATAGTGGCAGGTTCATTACTTTTTTTTCATAACACAACAAGACCGAATAAATATAACACTTATTCAGCTTGCACCAATTACGTTAGCCACGAGAAACAAGTAACTACATGTGGAGAACATTATTTGGGCACGTGTATTTTCAAGAAACATGTGCATGAGTGCATTTCAAAATTCAAATGTGGATTTTCGAGAAAATGTAGATATTTAATAAAAACAAATGTATCG |
| Figure 12C | pKB468 | *rex-39;* both MEX II mutated | kb416 | CGATACATTTGTTTTTATTAAATATCTACATTTTCTCG | kb417r | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAAATG | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAAATGAAAATGTTAACCTTAGCTTAAACTGGCAAGACCTCAATTCCTGTTCAGCATGAAAAACTCTTCGCATTATGAACTTTCGTTTCCGAATGTTATCCTGGTATGCTTCCACTCGAGTGTTACACTTTTCCATGTTCTATTATAATACAATTGAATTTTATTACTTCATGTCAAAAAACTCACAACTAAATCAAACTCTTCAACCAGCTGTCTCAATTTCTCAACTCGCCGTTTTTGAAAGAAATCCTGCAATAAATCCGTAAAATTTTGCCTAATAAGTCGAATTTCAGCAGCAAGGTTGTCGTCAAATGATGCTTTGCAGGCTGAAAATAAGCTTTTAGAAATAGTGGCAGGTTCATTACTTTTTTTTCATAACACAACAAGACCGAATAAATATAACACTTATTCTTTCATTGTTTGCGCGAGATATGCAAAGAAACAAGTAACTACAAAAAAGATGAACGCGCTCCTTTTTTCTATTTTCAAGAAACATGTGCATGAGTGCATTTCAAAATTCAAATGTGGATTTTCGAGAAAATGTAGATATTTAATAAAAACAAATGTATCG |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Figure 12C | pKB1015 | *rex-39;* both MEX II mutated to *Cbr* MEX II | kb416 | CGATACATTTGTTTTTATTAAATATCTACATTTTCTCG | kb417r | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAAATG | TTTCTGAAAAAATTGAAAGAATCTTGCTTAAAAATGAAAATGTTAACCTTAGCTTAAACTGGCAAGACCTCAATTCCTGTTCAGCATGAAAAACTCTTCGCATTATGAACTTTCGTTTCCGAATGTTATCCTGGTATGCTTCCACTCGAGTGTTACACTTTTCCATGTTCTATTATAATACAATTGAATTTTATTACTTCATGTCAAAAAACTCACAACTAAATCAAACTCTTCAACCAGCTGTCTCAATTTCTCAACTCGCCGTTTTTGAAAGAAATCCTGCAATAAATCCGTAAAATTTTGCCTAATAAGTCGAATTTCAGCAGCAAGGTTGTCGTCAAATGATGCTTTGCAGGCTGAAAATAAGCTTTTAGAAATAGTGGCAGGTTCATTACTTTTTTTTCATAACACAACAAGACCGAATAAATATAACACTTATTCCTCCTGCCTACTTCCACAACGCGCGAAGTAACAAGTAACTACTTCGCGCGTTGTGGAAGTAGGCAGGAGGATTTTCAAGAAACATGTGCATGAGTGCATTTCAAAATTCAAATGTGGATTTTCGAGAAAATGTAGATATTTAATAAAAACAAATGTATCG |
| Figure 12F, Figure 13C | pKB1023 | WT *rex-33* (reverse complement) | kb221 | AGATTTTGCGGTCAAATTGAGG | kb222r | GTCAGAAAGGGTATACGTTCCAAG | CTTGGAACGTATACCCTTTCTGACAAAAAACTATTTATAACGTCATGTTTTGATCATAACGTTTCATAACGTTCAAAATTATTGTTAGCTCAATTTATTTTGGACAACTCTTGCAAATTCTTAAATTTATAAACCATGCGATATTTGCAAACTCCAAATGTGTCCAAAAAAGGGCCCGTGGTTAATTTATTCGTGTTTTATCGTGTTTCTAACAGTTAAGCATACTAATTTGCGTGCCCTACTAAATAAGCGAAAGGCGGTAAATCTTCCCTGCGCGATACCTATCTATGGTGACCCCCTGCGCCAAATGTAAATGGGCAATCGATTGGGTGTCCCTTCGCTTAAAATAATTACTCACGTTTGAGTGTCTGTGCTTGCAGGAAAGACTGTTTGAACGGACAGGGCAGGCTTTGGCAGCAACACACATTGAGTATCGGGCAGAAAGTCGTTTTCGGAAAACTGAAAAAAAAATCTTAAACTATACAAGTGGAGTAGAGACATTTTGCAGTAACTTTTGAAATATATAAAATTACATGTATTGTTTCCATATAACATAGATAAAAGTTACTTTAAATTCAACCTCAATTTGACCGCAAAATCT |
| Figure 12F, Figure 13C | pKB1022 | *rex-33* with all 3 MEX motifs scrambled | kb221 | AGATTTTGCGGTCAAATTGAGG | kb222r | GTCAGAAAGGGTATACGTTCCAAG | CTTGGAACGTATACCCTTTCTGACAAAAAACTATTTATAACGTCATGTTTTGATCATAACGTTTCATAACGTTCAAAATTATTGTTAGCTCAATTTATTTTGGACAACTCTTGCAAATTCTTAAATTTATAAACCATGCGATATTTGCAAACTCCAAATGTGTCCAAAAAAGGGCCCGTGGTTAATTTATTCGTGTTTTATCGTGTTTCTAACAGTTAAGCATACTAATTTGCG |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | TGCCCTACTAAATAAGCGAAAGGCGGTAAATCAGCGTCG<br>CGTCATACCTATCTATGGTGAGCGCTCGAGCGCAATGTA<br>AATGGGCAATCGATTGGGTGTCACGTCGACTAAAATAAT<br>TACTCACGTTTGAGTGTCTGTGCTTGCAGGAAAGACTGT<br>TTGAACGGACAGGGCAGGCTTTGGCAGCAACACACATTG<br>AGTATCGGGCAGAAAGTCGTTTTCGGAAAACTGAAAAAA<br>AAATCTTAAACTATACAAGTGGAGTAGAGACATTTTGCA<br>GTAACTTTTGAAATATATAAAATTACATGTATTGTTTCC<br>ATATAACATAGATAAAAGTTACTTTAAATTCAACCTCAA<br>TTTGACCGCAAAATCT |
| Figure 12F, Figure 13C | pKB1021 | *rex-33* with all 3 MEX motifs mutated to *Cbr* MEX motifs | kb221 | AGATTTTGC<br>GGTCAAATT<br>GAGG | kb222r | GTCAGAAA<br>GGGTATAC<br>GTTCCAAG | CTTGGAACGTATACCCTTTCTGACAAAAAACTATTTATA<br>ACGTCATGTTTTGATCATAACGTTTCATAACGTTCAAAA<br>TTATTGTTAGCTCAATTTATTTTGGACAACTCTTGCAAA<br>TTCTTAAATTTATAAACCATGCGATATTTGCAAACTCCA<br>AATGTGTCCAAAAAAGGGCCCGTGGTTAATTTATTCGTG<br>TTTTATCGTGTTTCTAACAGTTAAGCATACTAATTTGCG<br>TGCCCTACTAAATAAGCGAAAGGCGGTAAATCTTCCCTT<br>CCCAATTACTATCTATGGTGACCCCCTGCCGCATTCGTA<br>AATGGGCAATCGATTGGGTGTCCCTGCCCATATAATAAT<br>TACTCACGTTTGAGTGTCTGTGCTTGCAGGAAAGACTGT<br>TTGAACGGACAGGGCAGGCTTTGGCAGCAACACACATTG<br>AGTATCGGGCAGAAAGTCGTTTTCGGAAAACTGAAAAAA<br>AAATCTTAAACTATACAAGTGGAGTAGAGACATTTTGCA<br>GTAACTTTTGAAATATATAAAATTACATGTATTGTTTCC<br>ATATAACATAGATAAAAGTTACTTTAAATTCAACCTCAA<br>TTTGACCGCAAAATCT |
| Figure 13C | pKB1026 | *rex-33* with all 3 MEX motifs mutated to *Cbr* MEX motifs with G7C change | kb221 | AGATTTTGC<br>GGTCAAATT<br>GAGG | kb222r | GTCAGAAA<br>GGGTATAC<br>GTTCCAAG | CTTGGAACGTATACCCTTTCTGACAAAAAACTATTTATA<br>ACGTCATGTTTTGATCATAACGTTTCATAACGTTCAAAA<br>TTATTGTTAGCTCAATTTATTTTGGACAACTCTTGCAAA<br>TTCTTAAATTTATAAACCATGCGATATTTGCAAACTCCA<br>AATGTGTCCAAAAAAGGGCCCGTGGTTAATTTATTCGTG<br>TTTTATCGTGTTTCTAACAGTTAAGCATACTAATTTGCG<br>TGCCCTACTAAATAAGCGAAAGGCGGTAAATCTTCCCTT<br>CGCAATTACTATCTATGGTGACCCCCTGCGGCATTCGTA<br>AATGGGCAATCGATTGGGTGTCCCTGCGCATATAATAAT<br>TACTCACGTTTGAGTGTCTGTGCTTGCAGGAAAGACTGT<br>TTGAACGGACAGGGCAGGCTTTGGCAGCAACACACATTG<br>AGTATCGGGCAGAAAGTCGTTTTCGGAAAACTGAAAAAA<br>AAATCTTAAACTATACAAGTGGAGTAGAGACATTTTGCA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | GTAACTTTTGAAATATATAAAATTACATGTATTGTTTCC<br>ATATAACATAGATAAAAGTTACTTTAAATTCAACCTCAA<br>TTTGACCGCAAAATCT |
| Figure 13C | pKB1028 | *rex-33* with C4G mutated in all 3 MEX motifs | kb221 | AGATTTTGC<br>GGTCAAATT<br>GAGG | kb222r | GTCAGAAA<br>GGGTATAC<br>GTTCCAAG | CTTGGAACGTATACCCTTTCTGACAAAAAACTATTTATA<br>ACGTCATGTTTTGATCATAACGTTTCATAACGTTCAAAA<br>TTATTGTTAGCTCAATTTATTTTGGACAACTCTTGCAAA<br>TTCTTAAATTTATAAACCATGCGATATTTGCAAACTCCA<br>AATGTGTCCAAAAAAGGGCCCGTGGTTAATTTATTCGTG<br>TTTTATCGTGTTTCTAACAGTTAAGCATACTAATTTGCG<br>TGCCCTACTAAATAAGCGAAAGGCGGTAAATCTTCCCTG<br>CCCGATACCTATCTATGGTGACCCCCTGCCCCAAATGTA<br>AATGGGCAATCGATTGGGTGTCCCTTCCCTTAAAATAAT<br>TACTCACGTTTGAGTGTCTGTGCTTGCAGGAAAGACTGT<br>TTGAACGGACAGGGCAGGCTTTGGCAGCAACACACATTG<br>AGTATCGGGCAGAAAGTCGTTTTCGGAAAACTGAAAAAA<br>AAATCTTAAACTATACAAGTGGAGTAGAGACATTTTGCA<br>GTAACTTTTGAAATATATAAAATTACATGTATTGTTTCC<br>ATATAACATAGATAAAAGTTACTTTAAATTCAACCTCAA<br>TTTGACCGCAAAATCT |
| Figure 12C, Figure 12F | pKB212 | np 1 | kb204 | ATTTGTATC<br>AAATCAAAG<br>AGCAGG | kb183r | GCGGTAAC<br>TGCTAGTT<br>TTCAGG | ATTTGTATCAAATCAAAGAGCAGGACACGGTTGTTGCTT<br>CAATCTACTATGTGCTAACGTTTATTTTCGAAATGACAG<br>CATTCTTTGTTATTAACAAAATGAATATCCCTTTCTTTT<br>CGGTAATTTTATGGTAGTTTACCGAGTAGAGACATTCAA<br>ATTTTAGGAGGACAATTTTCTGGATCACGTGGGAATGGA<br>TAATGATAACCAGGTAATGGTCACATTGTGTTGATGTAA<br>AAGAACAAGAAAATACAAAAAAGAAAGCAAGATCTTTTA<br>ATCAAAGTCCAAGATTTGTTTCTTCATGAAATCTGTGGA<br>AGTTGTTGGTTAAAGTACAACCCAGACCACGAGGGACTT<br>GAGTTATCTCGTCATTTAATTTTGTTTGATTTTCCGGTA<br>GTTATGTATGTAAACATCAGAATATTCCATTTGTCTGTA<br>GCTCATAATGATGCTGATAATAAATTTGTTATGCACTAA<br>TGACGAAAGCTAATGATTATTTTATCGTCTATTATTTTT<br>CGCATCTTTCAACTTCCTGGTATCTTGTTTTCTAAAATT<br>ATATTTTCATATTTTCTCGTTGCTGCCAAAAGTCCTGAA<br>AACTAGCAGTTACCGC |