# Supporting Information

## Beh et al. 10.1073/pnas.1118678109

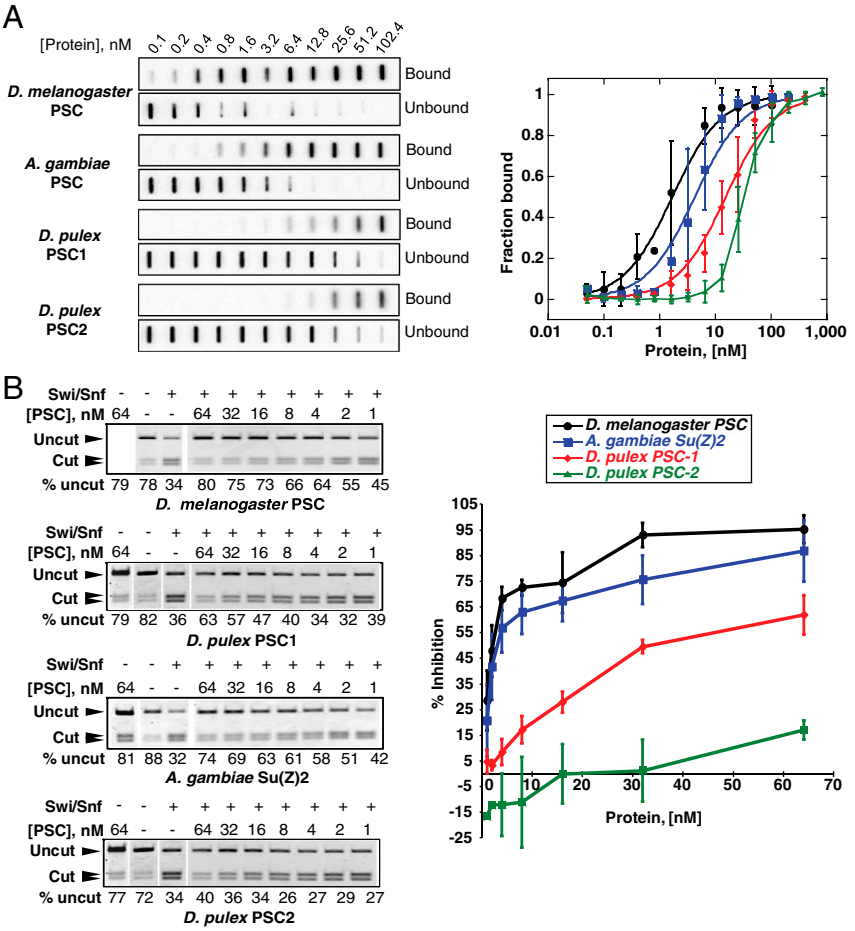



**Fig. S1.** Metazoan PSC-CTRs exhibit measurable variation in biochemical activities. (*A*) Left panel: representative data from double-filter binding assays that measure the $K_d$ of PSC-CTRs for free DNA. PSC-CTR protein was titrated into binding reactions containing 20pM 157-bp $^{32}$P-labeled DNA. Protein-DNA complexes are captured on the top (nitrocellulose) filter, while unbound DNA is captured on the bottom (charged nylon) filter. (*Right*) data from four different PSC-CTRs are graphed. Error bars denote standard deviation. (*B*) (*Left*) representative data from Restriction Enzyme Accessibility (REA) assays that measure the ability of PSC-CTRs to inhibit chromatin remodeling. The extent of Swi/Snf inhibition (labeled as "% inhibition") was calculated from: [(percent uncut with Swi/Snf and PSC-CTR) − (percent uncut with Swi/Snf)]/[(percent uncut without Swi/Snf) − (percent uncut with Swi/Snf)]) × 100%. (*Right*) representative data from four different PSC-CTRs are graphed. Error bars denote standard deviation.

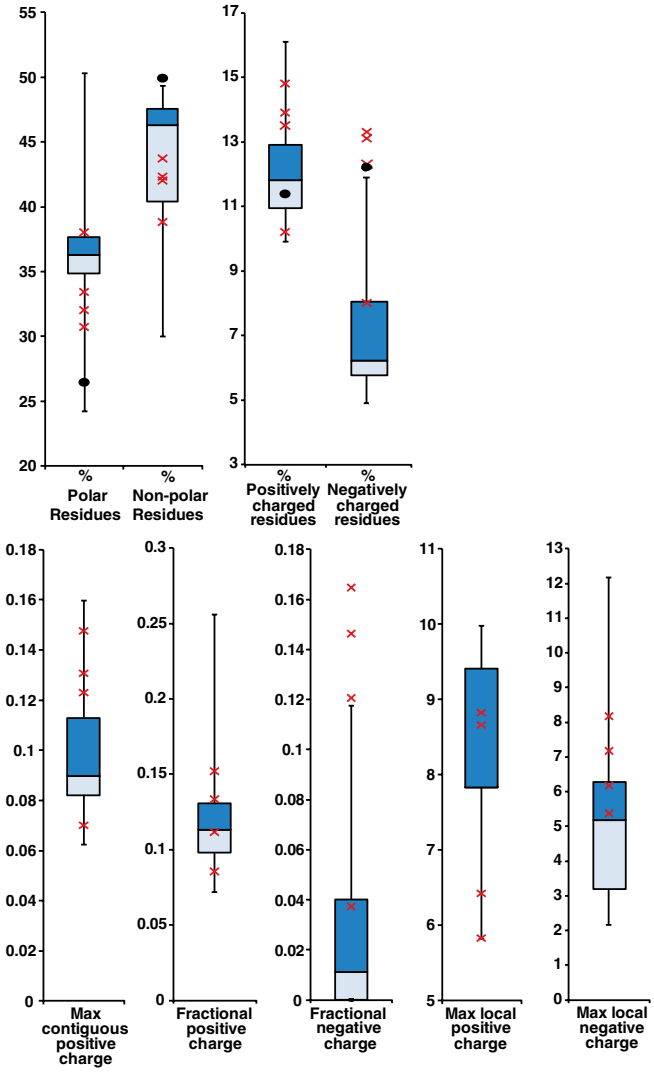



**Fig. S2.** Comparative analysis of PSC-CTRs. Overall amino acid composition and several charge properties cannot distinguish PSC-CTR activity. Repressive PSC-CTRs are depicted with box and whisker plots, while nonrepressive PSC-CTRs are represented by red crosses. The ends of the whiskers respectively represent the maximum and minimum data points; upper and lower bounds of the box represent the upper and lower quartiles; horizontal line through the box is the median. The average of the UniProt Knowledgebase is represented as a filled circle. Nonrepressive PSC-CTRs are *Daphnia pulex* PSC1 and PSC2, which were experimentally tested in this study, and *Mus musculus* BMI1 and *Xenopus laevis* PCGF2, which were previously tested (1). The amino acids aspartic acid and glutamic acid are classified as negatively charged; lysine and arginine are classified as positively charged; cysteine, serine, threonine, histidine, glutamine, asparagine, and tyrosine are classified as polar; alanine, glycine, isoleucine, leucine, methionine, phenylalanine, proline, tryptophan, and valine are classified as nonpolar. The "% residues" was calculated as the percentage of respective residues in each protein. "Max contiguous positive charge" represents the length of the longest positively charged stretch in the protein, normalized by protein length. "Fractional positive (or negative) charge" represents the fraction of 25-amino acid windows with charge greater than +3.5 (or below −3.5) in each PSC-CTR. "Max local positive (or negative) charge" represents the maximum (or minimum) charge attained amongst all 25-amino acid windows in each PSC-CTR.

1 Grau DJ, et al. (2011) Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev* 25:2210–2221.

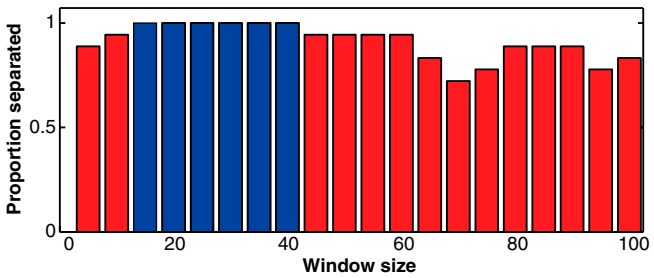



**Fig. S3.** Maximum contiguous negative charge distinguishes PSC-CTR activity over a wide range of window sizes. The proportion of nonrepressive PSC-CTRs separated from repressive PSC-CTRs by maximum contiguous negative charge is plotted against the window size used for this calculation. Complete separation between repressive and nonrepressive PSC-CTRs is represented by blue bars. This is observed from window sizes of 15–40 amino acids. Incomplete separation between repressive and nonrepressive PSC-CTRs is represented by red bars.

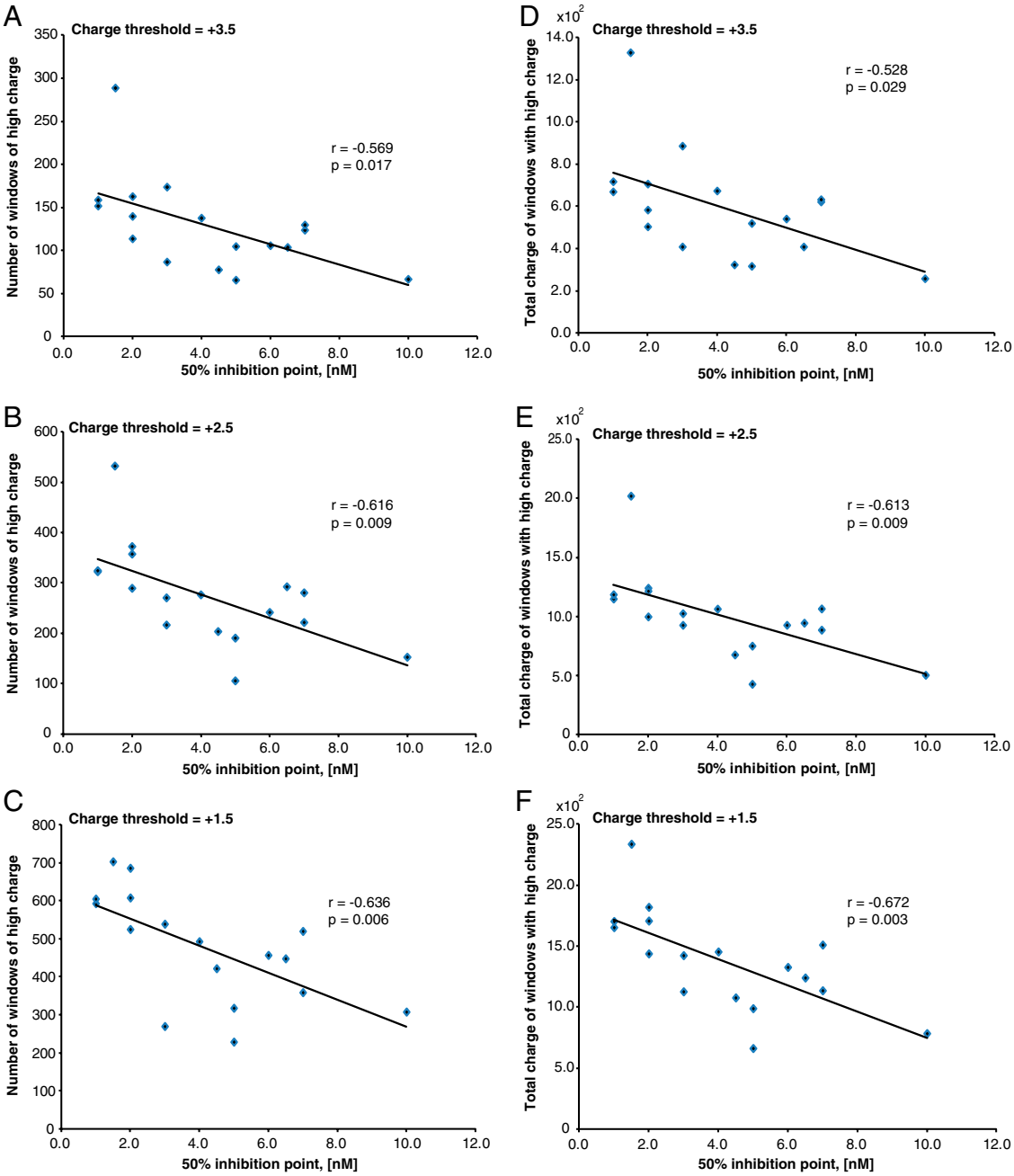



**Fig. S4.** The number of highly charged windows is correlated with the activity of repressive PSC-CTRs. (*A*, *B*, *C*) The number of 25-amino acid windows with charge greater than +3.5, +2.5, or +1.5 in each PSC-CTR was plotted against its 50% inhibition point (REA activity). A significant correlation is observed, suggesting that highly active PSC-CTRs tend to have a higher number of highly charged patches. This correlation is robust over the range of tested charge thresholds. (*D*, *E*, *F*) The total charge of all 25-amino acid windows with charge greater than +3.5, +2.5, or +1.5 in each PSC-CTR was plotted against the 50% inhibition point (REA activity). A significant correlation is observed, suggesting that the charged patches in highly active PSC-CTRs tend to exhibit a higher magnitude of positive charge. This correlation is robust over the range of tested charge thresholds.

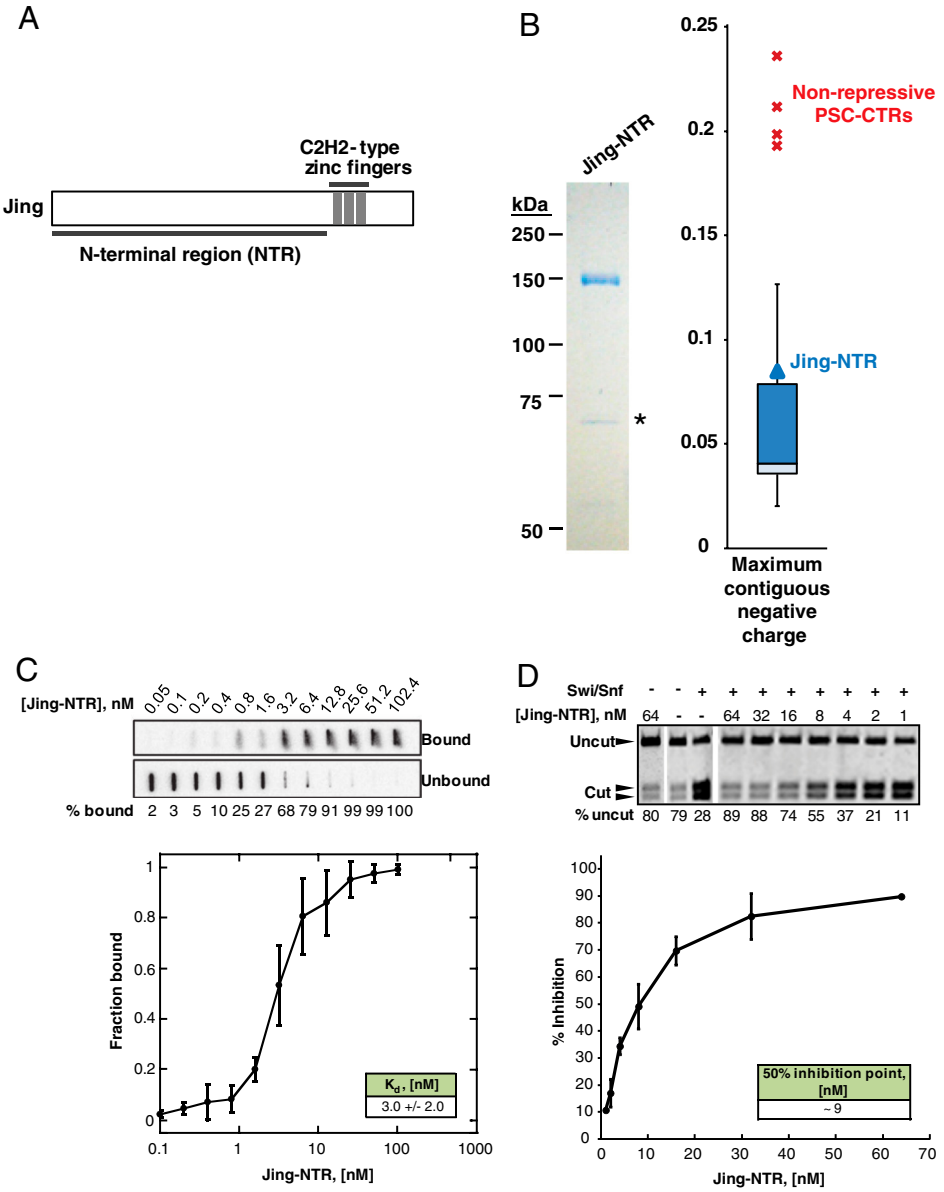



**Fig. S5.** Biochemical analysis of Jing. (*A*) Schematic representation of the N-terminal region (NTR) and conserved C2H2-type zinc fingers of *Drosophila melanogaster* Jing. (*B*) Coomassie-strained gel of *D. melanogaster* Jing-NTR purified from Baculovirus-infected Sf9 cells. The asterisk denotes an approximate 70 kDa band likely to be HSC70. Box and whisker plot depicts maximum contiguous negative charge of PSC-CTRs, relative to Jing (blue triangle). Box and whisker plot was constructed as described in Fig. S2. (*C*) Double filter binding assays to determine the affinity of Jing-NTR for free DNA. Error bars denote standard deviation. (*D*) REA assays to quantify inhibition of chromatin remodeling by Jing-NTR. Error bars denote standard deviation.

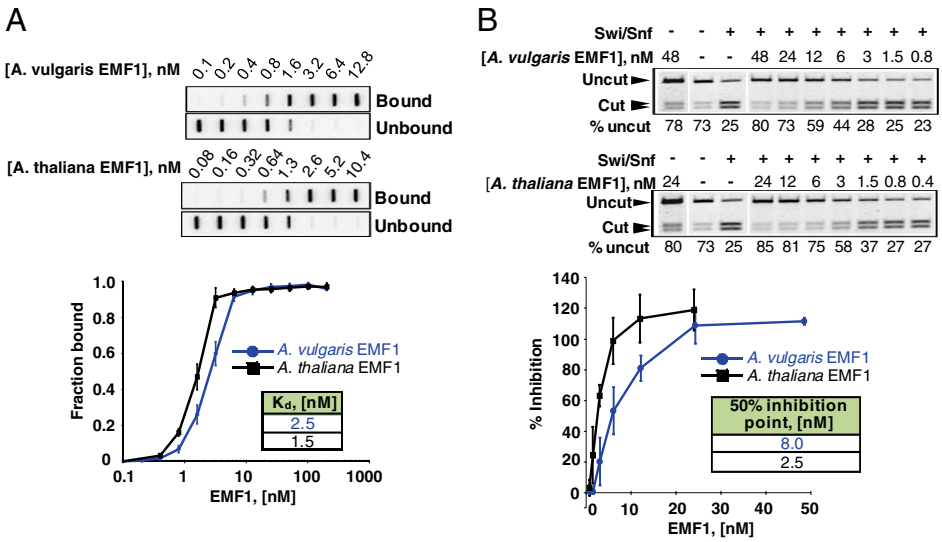



**Fig. S6.** Plant EMF1 binds tightly to DNA and inhibits chromatin remodeling. (*A*) Double filter binding assays to determine the affinity of *Arabidopsis thaliana* EMF1 and *Aquilegia vulgaris* EMF1 for free DNA. (*B*) REA assays to quantify inhibition of chromatin remodeling by *A. thaliana* EMF1 and *A. vulgaris* EMF1. All error bars denote standard deviation.

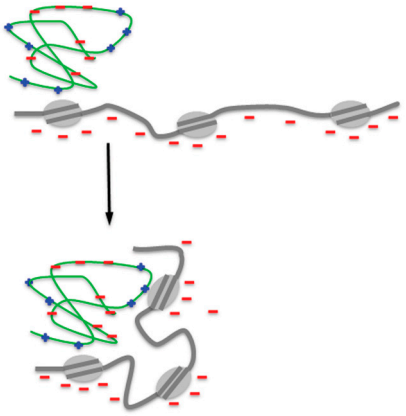
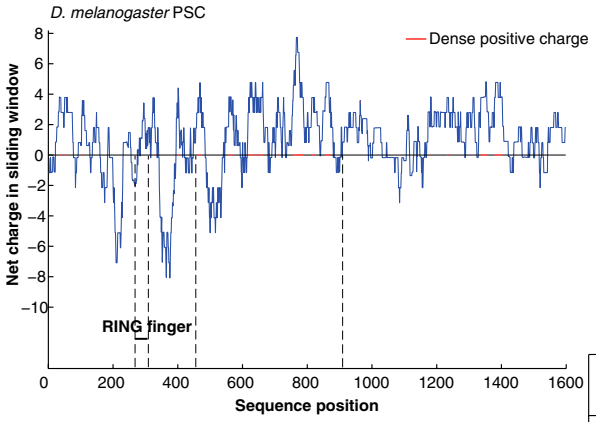


**Fig. S7.** Speculative model for chromatin compaction by PSC-CTRs. Sections of PSC-CTR with dense positive charge bind tightly to the negatively charged chromatin array. In contrast, sections of PSC-CTR that have local negative charge can move away from chromatin, out of the range over which electrostatic interactions are significant. The presence of multiple positively charged patches separated by regions of net negative charge may allow binding to noncontiguous chromatin or DNA segments, facilitating chromatin compaction. In the first step, a single positively charged patch binds chromatin. In the second step, a second positively charged patch engages a nonadjacent nucleosome (although binding to a more proximal site could also occur), holding the chromatin segments together. The PSC-CTR may be able to access distal sites through mechanism similar to the "fly-casting" or "monkey bar" mechanisms suggested to explain how disordered domains can have a high capture radius for initial encounter with their binding partners (1, 2). Alternatively, chromatin in solution is dynamic (3) so that the PSC-CTR could capture a second nucleosome during fluctuations of the chromatin fiber. We draw an extreme example of two segments of the PSC-CTR bound to distal sites on chromatin and separated by a long protein segment; in fact, we expect multiple binding sites and smaller protein loops. The diagram is intended to be conceptual—we are not able to make structural predictions about the configuration of the protein or chromatin.


1 Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc Natl Acad Sci USA* 97:8868–8873.
2 Vuzman D, Azia A Levy Y (2010) Searching DNA via a "Monkey Bar" mechanism: The significance of disordered tails. *J Mol Biol* 396:674–684.
3 Grigoryev SA, Arya G, Correll S, Woodcock CL, Schlick T (2009) Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc Natl Acad Sci USA* 106:13317–13322.




| Region of PSC | Activity | Predicted intrinsic disorder (%) | Amino acid composition distance | Max contig negative charge |
|---|---|---|---|---|
| 1-1603 | ++ | 85.5 | 0.046 | 0.037 |
| 456-909 | + | 100.0 | 0.085 | 0.122 |
| 1-910 | ++ | 80.8 | 0.092 | 0.062 |
| 456-1603 | ++ | 94.9 | 0.000 | 0.052 |

**Fig. S8.** Sequence analysis of D. melanogaster PSC subdomains. Charge plot of *D. melanogaster* PSC, with red horizontal lines denoting dense stretches of positive charge (defined as regions with charge greater than +3.5). Charge plots were generated from consecutive 25-amino acid sliding windows. Black horizontal line denotes the RING finger, while black horizontal bars denote various PSC regions that were previously characterized experimentally (1). "Activity" of each region was based on their ability to bind chromatin and inhibit chromatin remodeling, as previously reported (1). The metrics "Predicted intrinsic disorder," "Amino acid composition distance," and "Max contig negative charge" are computed as described in the "Criteria for prediction of repressive activity" subsection in *Materials and Methods*. The maximum contiguous negative charge value of *D. melanogaster* PSC 456–909 is highlighted in red to indicate that it is greater than that of other subdomains, and is closer to (although still below) the threshold of 0.15 for classification of a PSC-CTR as non-repressive.

1 Lo SM, Francis NJ (2010) Inhibition of chromatin remodeling by polycomb group protein posterior sex combs is mechanistically distinct from nucleosome binding. *Biochemistry* 49:9438–9448.

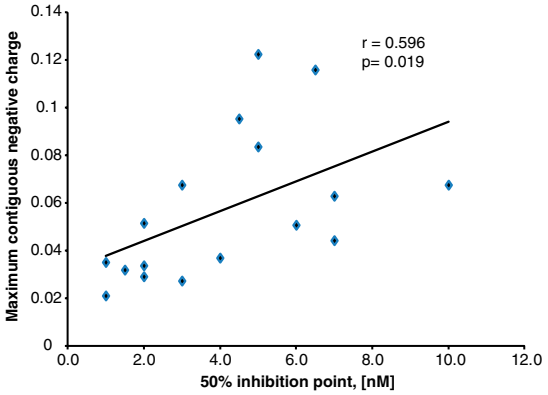



**Fig. S9.** Maximum contiguous negative charge is inversely correlated with repressive activity. The maximum contiguous negative charge (calculated as described in *Materials and Methods*) for each PSC-CTR was plotted against its 50% inhibition point. A significant correlation between maximum contiguous negative charge and 50% inhibition points is observed, indicating that PSC-CTRs with higher maximum contiguous negative charge tend to have lower repressive activity.

A

| | *C. elegans* PSC | | Mouse M33 |
|---|---|---|---|
| | N-terminal region | C-terminal region | Full length |
| Predicted intrinsic disorder (%) | 84.4 | 22.8 | 78.2 |
| Amino acid composition distance | 0.149 | 0.190 | 0.089 |
| Max contiguous negative charge | 0.095 | 0.149 | 0.078 |

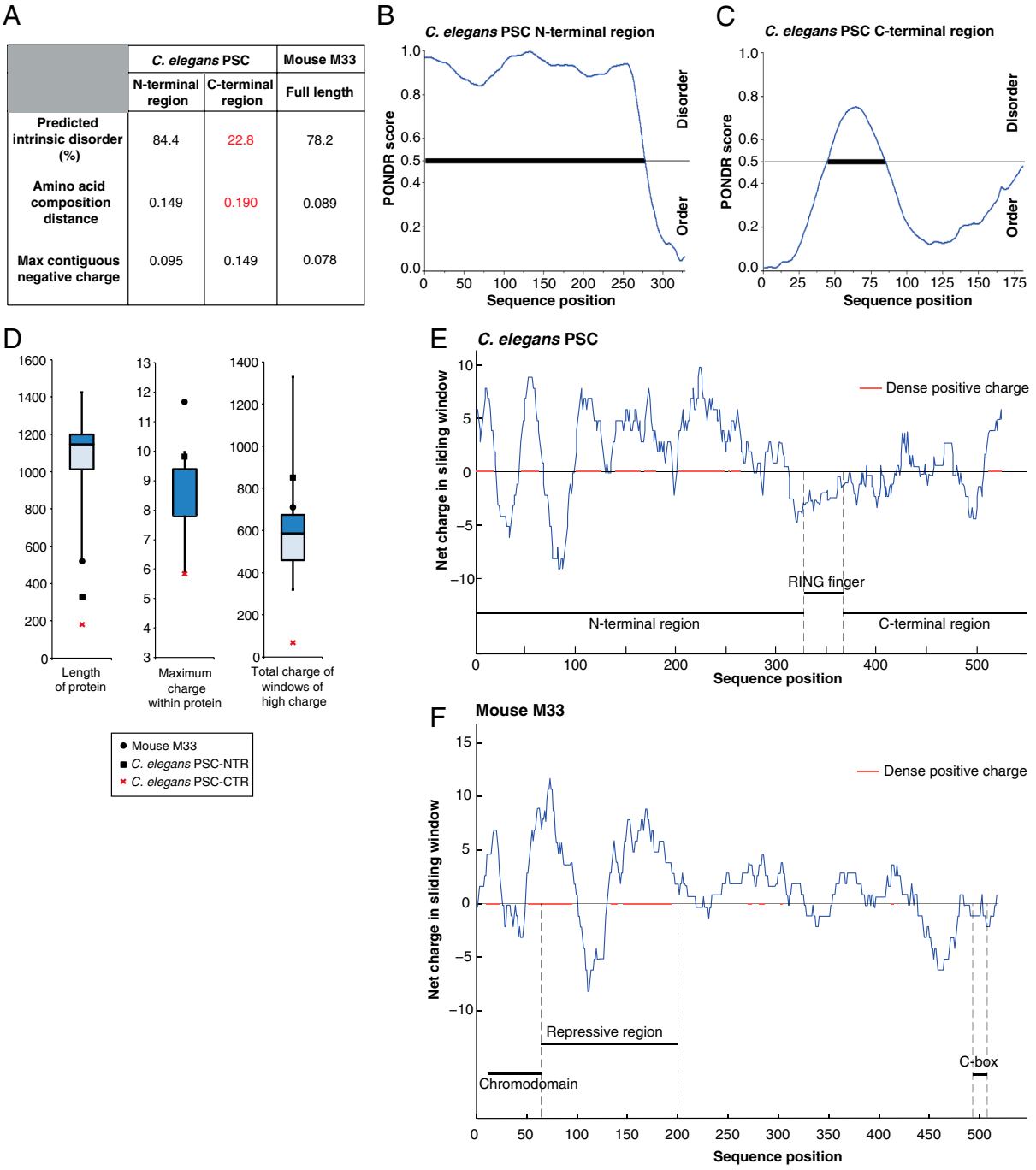



**Fig. S10.** Sequence and charge analysis of Caenorhabditis elegans PSC and mouse M33. (*A*) Summary of predicted intrinsic disorder, amino acid composition distance (with respect to *D. melanogaster PSC-CTR*), and maximum contiguous negative charge values for *C. elegans* PSC-NTR/CTR, and mouse M33. Metrics are computed as described in "Criteria for prediction of repressive activity" subsection in *Materials and Methods*. Values in red do not fulfill the respective criteria for repressive activity (described in *Materials and Methods*). Note that *C. elegans* PSC-NTR and mouse M33 exhibit lower maximum contiguous negative charge than the CTR. (*B*, *C*) PONDR VL3-BA prediction of intrinsic disorder in the N-terminal region (NTR) and C-terminal region (CTR) of *C. elegans* PSC. Only 22.8% of the CTR is predicted to be intrinsically disordered, compared to 84.4% in the NTR. (*D*) Comparison of charge and length properties of *C. elegans* PSC and mouse M33 with repressive PSC-CTRs. Repressive PSC-CTRs are depicted with box and whisker plots. The ends of the whiskers respectively represent the maximum and minimum data points; upper and lower bounds of the box represent the upper and lower quartiles; horizontal line through the box is the median. "Length of protein" is given by the number of amino acids in each protein. "Maximum charge within protein" represents the highest charge attained among all 25-amino acid windows in each protein sequence. "Total charge of windows of high charge" represents the sum of charge in all 25-amino acid windows with charge greater than +3.5 in each protein sequence. (*E*) Charge plot of *C. elegans* PSC, with red horizontal lines denoting dense stretches of positive charge (defined as regions with charge greater than +3.5). Black horizontal lines denote the respective domains: N-terminal region (NTR), C-terminal region (CTR), and RING finger. The charge plot was generated from consecutive 25-amino acid sliding windows. (*E*) Charge plot of mouse M33, with red horizontal lines denoting dense stretches of positive charge (defined as regions with charge greater than +3.5). Black horizontal lines denote the respective domains: chromodomain, C-box, and a repressive region defined by Grau *et al* (1) to be important for PSC-CTR-like activity. Note that the repressive region contains dense patches of positive charge. The charge plot was generated from consecutive 25-amino acid sliding windows.

1 Grau DJ, et al. (2011) Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev* 25:2210–2221.

**Table S1. Annotated PSC-CTR sequences from metazoan and plant genomes. (*A*) All PSC-CTRs are listed, with their respective lengths, maximum contiguous negative charge, and amino acid composition distance (defined as the square root of the sum of the squared distances for each amino acid from *D. melanogaster* PSC-CTR.). (*B*) Filtered (predicted repressive) PSC-CTRs are listed, with their respective length, maximum contiguous negative charge, amino acid composition distance, and predicted extent of structural disorder. Detailed prediction criteria are described in *Materials and Methods*.Table S1 (XLSX)**

**Table S2. Tissue sources, protein and primer sequences and accession numbers. (*A*) Tissue samples and cell lines used for nucleic acid isolation and subsequent PCR amplification of PSC, Jing, and EMF1 genes. (*B*) Sequences of experimentally tested proteins and the Homology Region (HR) of *D. melanogaster* PSC. Strain-specific sequence polymorphisms may be present in the cloned sequences. There is a slight overlap in sequence between *D. melanogaster* PSC-HR used for TBLASTN searches, and *D. melanogaster* PSC-CTR purified in this study. This PSC-CTR sequence is two amino acids shorter than that previously used (1–3). The database accession number for EMF1 (AcoGoldSmith_v1.000332m) corresponds to that *Aquilegia coerulea* EMF1, a very closely related species to *Aquilegia vulgaris*). Note that the *D. melanogaster* HR was not cloned; rather, it was used in BLAST searches to query for PSC homologues. (*C*) Primers used for amplification of PSC-CTR, Jing, and EMF1 genes.Table S2 (XLSX)**

1 Lo SM, Francis NJ (2010) Inhibition of chromatin remodeling by polycomb group protein posterior sex combs is mechanistically distinct from nucleosome binding. *Biochemistry* 49:9438–9448.
2 Francis NJ, Kingston RE, Woodcock CL (2004) Chromatin compaction by a polycomb group protein complex. *Science* 306:1574–1577.
3 King IF, et al. (2005) Analysis of a polycomb group protein defines regions that link repressive activity on nucleosomal templates to in vivo function. *Mol Cell Biol* 25:6578–6591.