

***Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of
H3K27me3 domains**

Kevin Van Bortle, Edward Ramos, Naomi Takenaka, Jingping Yang, Jessica E. Wahi, and
Victor G. Corces

SUPPLEMENTAL MATERIAL

Supplemental Table and Figure Legends

Table S1. H3K27me3 domains are enriched for developmental genes. H3K27me3 domains were determined by ChIP-seq in *Drosophila* Kc cells, and ultimately compared with publicly available H3K27me3 ChIP-seq data in late embryos and the requirement of overlapping Polycomb signal obtained from publicly available ChIP-chip data in *Drosophila* Kc cells (Celniker et al. 2009). GO analysis was performed by submitting domain-containing genes to DAVID using all annotated *Drosophila* Flybase gene IDs as background.

Table S2. Knockdown of insulator proteins has no effect on gene expression at genes flanking the *even skipped* gene. RNAi knockdown of insulator proteins results in moderate changes in expression for CG12134, *even skipped*, and *TER94*.

Table S3. List of primers used for qRT-PCR including sequential ChIP experiments (Figure S2) and H3K27me3 level analyses in insulator knockdowns (Figure 6).

Table S4. List of primers used for RNAi against insulator proteins CP190, dCTCF, SU(HW), and BEAF-32. Multiple amplicons targeting each gene for knockdown were used with the exception of BEAF-32, which only used one.

Table S5. List of H3K27me3 domains called in *Drosophila* Kc cells by comparing H3K27me3 ChIP-seq data with publicly available Polycomb data (Celniker et al. 2009). A total of 127 domains were found with an average size of ~ 87 Kb.

Figure S1. ChIP-seq reveals previously published consensus sequences and novel dCTCF secondary sequences. (A-C) Position weight matrices for dCTCF, BEAF-32, and SU(HW), respectively. Consensus sequences determined by MEME-ChIP (Machanick and Bailey 2011)

using previously reported ChIP-seq data (Wood et al., 2011). Enrichment for secondary dCTCF motif identified by removing dCTCF sites containing primary sequence, and independently confirmed using Weeder 1.3 (Pavesi et al., 2004). (D) Previously reported dCTCF consensus sequences in humans and *D. melanogaster* (Holohan et al., 2007). (E-F) Previously reported BEAF-32 and SU(HW) consensus sequence in *D. melanogaster* (Negre et al., 2010).

Figure S2. dCTCF tandemly aligns with BEAF-32 and SU(HW) in individual cells. (A) ChIP-seq profile for insulator proteins dCTCF (black), BEAF-32 (blue), and SU(HW) (green) at sites where dCTCF aligns with either BEAF-32, SU(HW), or both. (B) Sequential ChIP-PCR: primary ChIP for dCTCF was subsequently re-chromatin immunoprecipitated for either BEAF-32 (blue) or SU(HW) (green), or pre-immune sera (grey). Enriched signal for insulator proteins suggests these insulators bind concurrently.

Figure S3. dCTCF motifs 2 and 3 represent low occupancy binding sites compared to the conserved core consensus. (A) Heatmap for dCTCF sites containing motifs 1, 2, or 3 sorted from highest to lowest read intensity. (B) Average read intensities for dCTCF at motifs 1, 2, and 3 reveal lower occupancies at motifs 2 and 3, suggesting these represent lower affinity target sites.

Figure S4. dCTCF sites are enriched for motif 2 when aligned with BEAF-32, or BEAF-32 and SU(HW). Represented as percentage of dCTCF sites with motifs 1, 2, and 3 when independent, aligned with BEAF-32, aligned with SU(HW), or aligned with both BEAF-32 and SU(HW).

Figure S5. dCTCF and BEAF-32 co-localize with MOD(MDG4) isoforms different than MOD(MDG4)2.2. (A) Heatmap and average read intensities for MOD(MDG4) and MOD(MDG4)2.2 at SU(HW) sites. Strong enrichment for MOD(MDG4)2.2 is observed as expected. (B-C) Heatmap and average read intensities for MOD(MDG4) and MOD(MDG4)2.2

at BEAF-32 and dCTCF sites respectively. Enrichments are reversed, with stronger read intensities for MOD(MDG4), consistent with strong co-localization to sites where MOD(MDG4), but not MOD(MDG4)2.2, is present (Figure 3).

Figure S6. dCTCF aligns with BEAF-32 and/or SU(HW) at H3K27me3 domain borders. (A) Percentage of dCTCF sites aligned with BEAF-32 (blue), SU(HW) (green), and both (black) at all aligned dCTCF sites. (B) Percentage of dCTCF sites aligned with BEAF-32 (blue), SU(HW) (green), and both (black) within 5 Kb of H3K27me3 domain borders.

Figure S7. Efficacy of insulator knockdown. (A) Western blot probing for CP190 in repeated knockdown experiments. α -H3 was used as a loading control. (B) Western blot probing for dCTCF in repeated knockdown experiments. (C) Western blot probing for dCTCF/SU(HW) and dCTCF/BEAF-32 in repeated combinatorial knockdown experiments. (D) Western blot probing for MOD(MDG4) and MOD(MDG4)/CP190 in repeated combinatorial knockdown experiments. (E) Western blot probing for H3K27me3 levels in each knockdown condition.

Figure S8. H3K27me3 levels and gene expression in response to insulator knockdown. (A) H3K27me3 levels are specifically reduced in Pc domains containing dCTCF when dCTCF is knockdown. Average read intensities for H3K27me3 in WT *Drosophila* Kc cells (black), compared to cells in which dCTCF is knockdown (red) at Pc domain borders. Comparison of H3K27me3 reduction in Pc domains containing dCTCF to those without dCTCF. (B) Gene expression within H3K27me3 is largely unaffected by insulator knockdown. Percentage of genes up- or down-regulated by >2 fold under each insulator condition. Comparison between genome-wide average (black) and H3K27me3 containing genes (red) in each condition.

Figure S9. The *even skipped* gene provides a model locus for insulator alignment, localization to H3K27me3 domain borders, and H3K27me3 depletion in insulator knockdown experiments. ChIP-seq profiles for insulator proteins dCTCF, BEAF-32, SU(HW), CP190, MOD(MDG4), L(3)MBT, Chromator, and histone H3K27me3 surrounding the *even skipped* gene. H3K27me3 levels are depleted after dCTCF knockdown, though *even skipped* expression remains low (Table S2). These results suggest insulators do not align simply to form a robust barrier, as H3K27me3 does not spread, nor is the expression of domain flanking genes affected (Table S2).

Table S1. Gene ontology analysis for biological processes of genes within H3K27me3 domains

	<u>P-value</u>
Regulation of metabolic process	1.5 E-33
Multicellular organismal development	1.5 E-24
Anatomical structure development	1.1 E-23
Embryonic development	2.7 E-23
Pattern specification process	5.0 E-21
Anatomical structure morphogenesis	4.5 E-17
Cellular developmental process	1.4 E-10
Post-embryonic development	4.2 E-10
Anatomical structure formation involved in morphogenesis	4.6 E-07

VanBortle_TableS2

Table S2. Effect of insulator knockdown on gene expression (fold change) at *eve* locus

<i>knockdown:</i>								
	CP190	dCTCF	BEAF-32	Su(Hw)	dCTCF + CP190	BEAF-32 + CP190	Su(Hw) + CP190	dCTCF + BEAF-32
CG12134	1.09	-1.06	1.08	-1.15	-1.04	-1.08	-1.05	-1.07
<i>even skipped</i>	-1.32	-1.42	-1.01	-1.86	-1.52	-1.80	-1.81	-1.48
<i>TER94</i>	1.13	-1.48	1.18	-1.57	-1.19	-1.29	-1.30	-1.09

Table S3. List of primers used for qRT-PCR experiments

<u>Sequential ChIP (Fig. S1)</u>		
chr2L	F	TTCTTTTTCCGAAACGAAGC
	R	GATCCACAGATTCAGCCACA
chrX	F	AAAAGCACGCCAATAGTTGA
	R	CTCTCGCAGCTTCAGCTTTC
chr2R	F	GCGAGCATAGCCAAGTTAGG
	R	GCTGGAAATCGGCACTTTTA
<u>H3K27me3 levels (Fig. 7)</u>		
1	F	ATATCACCTCGTTGCGGTTC
	R	CTCCTCGCACTCGTAGCTCT
2	F	TCTAGCCATCGCCATCTTCT
	R	CCCTAATCCCTTCGACATCA
3	F	GACCTGCCGCATTGTTATCT
	R	AGATCGCGCAGATTCTTGAT
4	F	GTTGCGAAACTAGGGCAAAC
	R	CCGAAAATTGTATCCTCCA
5	F	TAGGCCGCGAATGTTTTTAG
	R	AATGGCATAAGCGGAAACTG
6	F	GCCACTCGGAGAGTCTTCTG
	R	TCACCGACTCCCTGATTCTC
7	F	GAAGGGATGGAATGGTGCTA
	R	GCCGATCCCATCTTATCTGA
8	F	TACTTGTGACGGAGCGTCTG
	R	CCTTCATGTGGAAGGGGTTA
9	F	AATGTGGGCTATGGCTTGTC
	R	TCTGTATCCGTGCTGCTGTC
10	F	CCGGAATTCCCCCTATTTTA
	R	GAGAGCGGGTAGATTGTTGC
11	F	CGAGTTCCTAATCCGAAACG
	R	CTGGTGGGCAGCTAGGTACT
12	F	GTGTTCGCCATTCTTTCCAT
	R	GCACACTCGATCTGAAACCA

Table S4. List of primers used for RNAi

	Primer	Sequence (T7 - gene specific target)	Length	GC	Tm
B-gal-1a	B-gal genomic forward	taatacgactcactataggg-ggtttccgcgaggt	34	50	74.12
B-gal-1b	B-gal genomic reverse	taatacgactcactataggg-gtcgcacagcgtgtac	36	50	73.87
B-gal-2a	B-gal genomic forward	taatacgactcactataggg-agcgaaaacggtctg	35	45.72	72.03
B-gal-2b	B-gal genomic reverse	taatacgactcactataggg-gctcagctggtccg	34	52.95	74.32
SuHw-A3'	SuHw genomic reverse	taatacgactcactataggg- aaacaacaacgggtcatcaatg	41	39.03	73.86
SuHw-A5'	SuHw genomic forward	taatacgactcactataggg- ctacatgattcttcaaattatct	44	34.1	70.84
SuHw-B3'	SuHw genomic reverse	taatacgactcactataggg- aactgtggccacgacctaac	40	47.5	74.78
SuHw-B5'	SuHw genomic forward	taatacgactcactataggg- tgtgcggttaattgtacgcat	40	42.5	74.05
CP190-A5'	CP190 genomic forward	taatacgactcactataggg- cctggctgtgcctgaga	37	51.36	76.7
CP190-A3'	CP190 genomic reverse	taatacgactcactataggg- ctggtagacttatgtccgaaa	41	41.47	72.47
CP190-B3'	CP190 genomic reverse	taatacgactcactataggg- gaatgcggaattggtgaatc	40	42.5	75.48
CP190-B5'	CP190 genomic forward	taatacgactcactataggg- agccgcgagatgacttagt	40	45	74.55
CP190 RNAi	from el R	TAATACGACTCACTATAGGGAG AGTAAACGGACGACCCATTAGC ATTC			
CP190 RNAi	from el S	TAATACGACTCACTATAGGGAG AGTCTGCTCTGGTTCCTGCTC			
CTCF-A5'	CTCF genomic forward	taatacgactcactataggg- GAGCCCGACATCAGTTCAAT			
CTCF-A3'	CTCF genomic reverse	taatacgactcactataggg-			

		GAGCACTTGAAGGATGGCTC			
CTCF-B5'	CTCF genomic forward one off target	taatacgactcactataggg- CCTACGCATCGCAGGAC	•	•	•
CTCF-B3'	CTCF genomic reverse one off target	taatacgactcactataggg- TTCTGGTGGCTGATACTCTT	•	•	-
BEAF-32	BEAF genomic forward	taatacgactcactataggg- AGGGTCGTGTTAG	•	•	
BEAF-32	BEAF genomic reverse	taatacgactcactataggg- GAAAAGAAATCAG	•	•	

Table S5. List of H3K27me3 domains in *Drosophila* Kc cells

1	chr2L	491076	607040	H3K27me3_Kc	0	+	
2	chr2L	1945176	1973640	H3K27me3_Kc	0	+	
3	chr2L	2378476	2491040	H3K27me3_Kc	0	+	
4	chr2L	3531276	3618540	H3K27me3_Kc	0	+	
5	chr2L	3657876	3698240	H3K27me3_Kc	0	+	
6	chr2L	3815476	3862040	H3K27me3_Kc	0	+	
7	chr2L	5356476	5513040	H3K27me3_Kc	0	+	
8	chr2L	6528076	6557040	H3K27me3_Kc	0	+	<i>eyes absent</i>
9	chr2L	6800276	6850040	H3K27me3_Kc	0	+	
10	chr2L	7263676	7391540	H3K27me3_Kc	0	+	
11	chr2L	8745276	8887940	H3K27me3_Kc	0	+	
12	chr2L	9577276	9612640	H3K27me3_Kc	0	+	
13	chr2L	11284976	11485740	H3K27me3_Kc	0	+	
14	chr2L	12568876	12691340	H3K27me3_Kc	0	+	
15	chr2L	14370340	14596740	H3K27me3_Kc	0	+	
16	chr2L	15087276	15128040	H3K27me3_Kc	0	+	
17	chr2L	15291476	15495940	H3K27me3_Kc	0	+	
18	chr2L	16381576	16488740	H3K27me3_Kc	0	+	
19	chr2L	18714876	18828140	H3K27me3_Kc	0	+	
20	chr2L	18860576	18938740	H3K27me3_Kc	0	+	
21	chr2L	20767676	20787240	H3K27me3_Kc	0	+	
22	chr2L	21766976	21965340	H3K27me3_Kc	0	+	
23	chr2L	22004976	22041740	H3K27me3_Kc	0	+	
24	chr2R	1596081	1640245	H3K27me3_Kc	0	+	
25	chr2R	2380281	2488381	H3K27me3_Kc	0	+	
26	chr2R	3297181	3327745	H3K27me3_Kc	0	+	
27	chr2R	3883181	3938245	H3K27me3_Kc	0	+	
28	chr2R	4446681	4539445	H3K27me3_Kc	0	+	
29	chr2R	4892981	4971281	H3K27me3_Kc	0	+	
30	chr2R	5067481	5123145	H3K27me3_Kc	0	+	
31	chr2R	5860581	5876245	H3K27me3_Kc	0	+	<i>even skipped</i>
32	chr2R	7353781	7462545	H3K27me3_Kc	0	+	
33	chr2R	7580381	7703945	H3K27me3_Kc	0	+	
34	chr2R	8769581	8946845	H3K27me3_Kc	0	+	
35	chr2R	10084681	10102345	H3K27me3_Kc	0	+	
36	chr2R	10323381	10367745	H3K27me3_Kc	0	+	
37	chr2R	10661981	10721345	H3K27me3_Kc	0	+	
38	chr2R	12691181	12716145	H3K27me3_Kc	0	+	
39	chr2R	13693681	13727145	H3K27me3_Kc	0	+	
40	chr2R	14539645	14563545	H3K27me3_Kc	0	+	

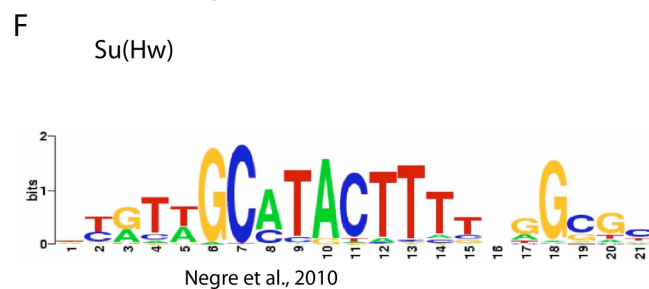
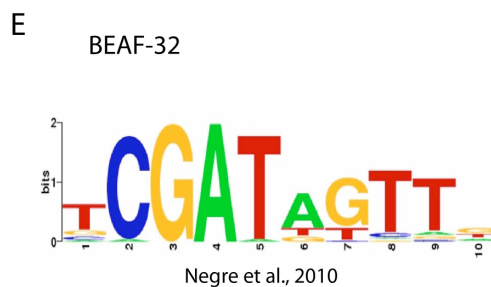
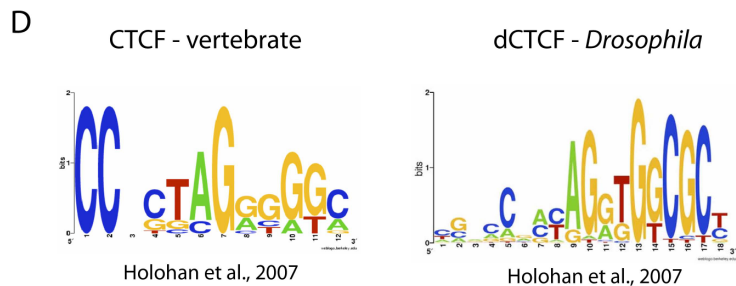
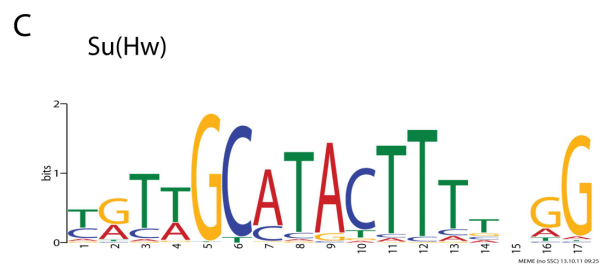
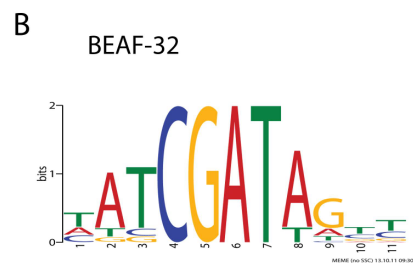
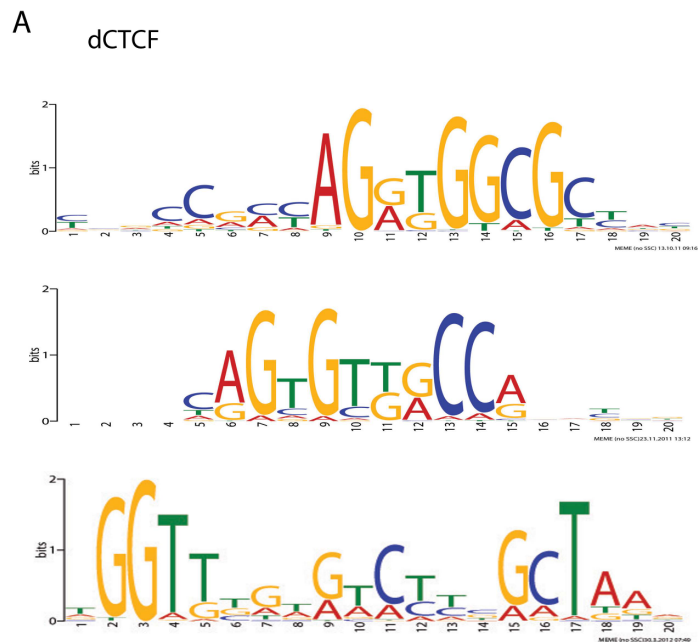
41	chr2R	15135181	15167845	H3K27me3_Kc	0	+
42	chr2R	16773481	16878345	H3K27me3_Kc	0	+
43	chr2R	17410381	17429845	H3K27me3_Kc	0	+
44	chr2R	17448581	17494745	H3K27me3_Kc	0	+
45	chr2R	17686345	17783245	H3K27me3_Kc	0	+
46	chr2R	18091881	18184345	H3K27me3_Kc	0	+
47	chr2R	18870981	18993745	H3K27me3_Kc	0	+
48	chr2R	19444581	19475645	H3K27me3_Kc	0	+
49	chr2R	19491981	19569945	H3K27me3_Kc	0	+
50	chr2R	20220281	20273445	H3K27me3_Kc	0	+
51	chr2R	20685581	20748545	H3K27me3_Kc	0	+
52	chr2R	20919681	21130045	H3K27me3_Kc	0	+
53	chr3L	367329	435593	H3K27me3_Kc	0	+
54	chr3L	1096029	1134193	H3K27me3_Kc	0	+
55	chr3L	1164329	1251493	H3K27me3_Kc	0	+
56	chr3L	6753129	6934193	H3K27me3_Kc	0	+
57	chr3L	6974229	6998193	H3K27me3_Kc	0	+
58	chr3L	7933529	7961593	H3K27me3_Kc	0	+
59	chr3L	8993929	9043093	H3K27me3_Kc	0	+
60	chr3L	10219629	10330093	H3K27me3_Kc	0	+
61	chr3L	10896929	11031593	H3K27me3_Kc	0	+
62	chr3L	12412329	12464093	H3K27me3_Kc	0	+
63	chr3L	12566429	12723193	H3K27me3_Kc	0	+
64	chr3L	14087829	14182393	H3K27me3_Kc	0	+
65	chr3L	14558229	14597793	H3K27me3_Kc	0	+
66	chr3L	18226229	18447093	H3K27me3_Kc	0	+
67	chr3L	19640429	19678493	H3K27me3_Kc	0	+
68	chr3L	20544629	20710493	H3K27me3_Kc	0	+
69	chr3L	21442329	21475893	H3K27me3_Kc	0	+
70	chr3L	21572429	21605593	H3K27me3_Kc	0	+
71	chr3L	21770429	21813193	H3K27me3_Kc	0	+
73	chr3R	648636	729700	H3K27me3_Kc	0	+
74	chr3R	2487736	2891100	H3K27me3_Kc	0	+
75	chr3R	3956836	4054400	H3K27me3_Kc	0	+
76	chr3R	4096836	4164100	H3K27me3_Kc	0	+
77	chr3R	6243636	6415700	H3K27me3_Kc	0	+
78	chr3R	6428500	6504700	H3K27me3_Kc	0	+
79	chr3R	7071736	7204500	H3K27me3_Kc	0	+
80	chr3R	8050236	8166000	H3K27me3_Kc	0	+
81	chr3R	8872136	8903500	H3K27me3_Kc	0	+
82	chr3R	9672036	9780900	H3K27me3_Kc	0	+
83	chr3R	11811336	11866300	H3K27me3_Kc	0	+
84	chr3R	12201136	12257300	H3K27me3_Kc	0	+

85	chr3R	12299136	12453000	H3K27me3_Kc	0	+
86	chr3R	12484436	12855100	H3K27me3_Kc	0	+
87	chr3R	13372736	13437200	H3K27me3_Kc	0	+
88	chr3R	13856636	13982100	H3K27me3_Kc	0	+
89	chr3R	14038236	14096136	H3K27me3_Kc	0	+
90	chr3R	14179436	14219200	H3K27me3_Kc	0	+
91	chr3R	16465236	16502636	H3K27me3_Kc	0	+
92	chr3R	16896636	16914800	H3K27me3_Kc	0	+
93	chr3R	17204336	17394900	H3K27me3_Kc	0	+
94	chr3R	18266700	18294800	H3K27me3_Kc	0	+
95	chr3R	20902536	20922500	H3K27me3_Kc	0	+
96	chr3R	22694936	22712900	H3K27me3_Kc	0	+
97	chr3R	22986900	23066700	H3K27me3_Kc	0	+
98	chr3R	24371236	24424300	H3K27me3_Kc	0	+
99	chr3R	25330536	25416800	H3K27me3_Kc	0	+
100	chr3R	25509036	25542700	H3K27me3_Kc	0	+
101	chr3R	26430436	26613400	H3K27me3_Kc	0	+
102	chr3R	26725436	26771700	H3K27me3_Kc	0	+
103	chr3R	26871736	26924300	H3K27me3_Kc	0	+
104	chr4	397780	451944	H3K27me3_Kc	0	+
105	chr4	518380	546480	H3K27me3_Kc	0	+
106	chr4	719880	741644	H3K27me3_Kc	0	+
107	chr4	819180	869544	H3K27me3_Kc	0	+
108	chr4	997180	1025244	H3K27me3_Kc	0	+
109	chr4	1102980	1130244	H3K27me3_Kc	0	+
110	chrX	444403	505967	H3K27me3_Kc	0	+
111	chrX	2020503	2042367	H3K27me3_Kc	0	+
112	chrX	4269603	4540667	H3K27me3_Kc	0	+
113	chrX	5206303	5218567	H3K27me3_Kc	0	+
114	chrX	5423303	5506267	H3K27me3_Kc	0	+
115	chrX	5888203	5928567	H3K27me3_Kc	0	+
116	chrX	7006003	7172467	H3K27me3_Kc	0	+
117	chrX	7396003	7568567	H3K27me3_Kc	0	+
118	chrX	8501903	8570767	H3K27me3_Kc	0	+
119	chrX	8617203	8729767	H3K27me3_Kc	0	+
120	chrX	9179603	9207267	H3K27me3_Kc	0	+
121	chrX	9583803	9669467	H3K27me3_Kc	0	+
122	chrX	16002103	16144167	H3K27me3_Kc	0	+
123	chrX	17200603	17311867	H3K27me3_Kc	0	+
124	chrX	17647903	17714167	H3K27me3_Kc	0	+
125	chrX	18133703	18213467	H3K27me3_Kc	0	+
126	chrX	18421003	18522203	H3K27me3_Kc	0	+
127	chrX	19427103	19451567	H3K27me3_Kc	0	+

Ods-site homeobox

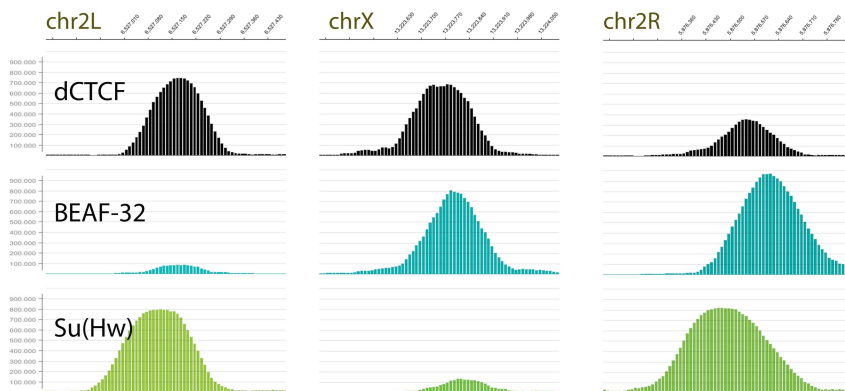
128	chrX	20405703	20630367	H3K27me3_Kc	0	+
-----	------	----------	----------	-------------	---	---

VanBortle_FigS1

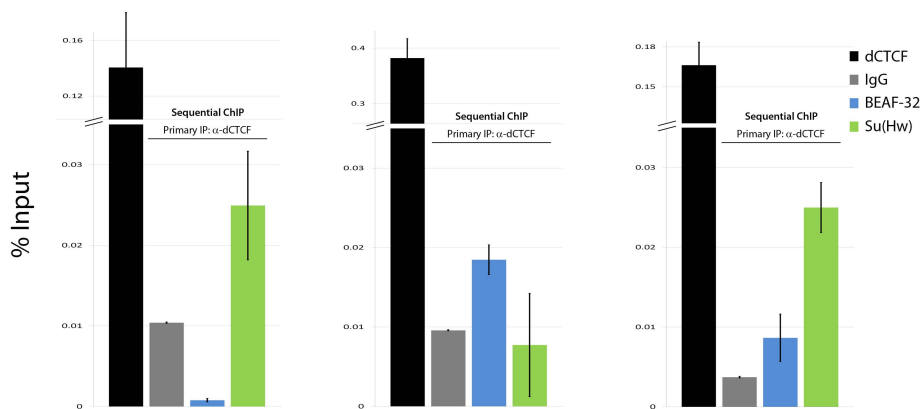


VanBortle_FigS2

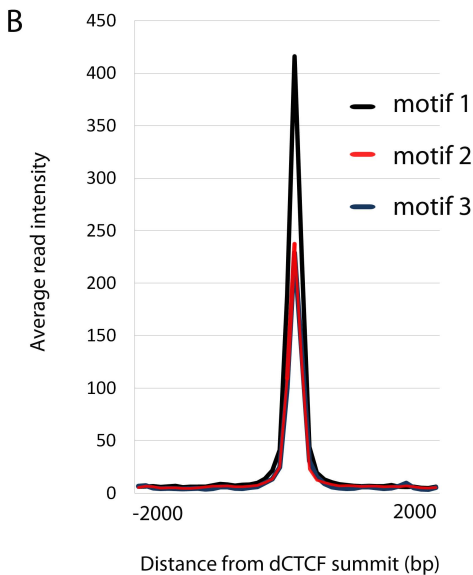
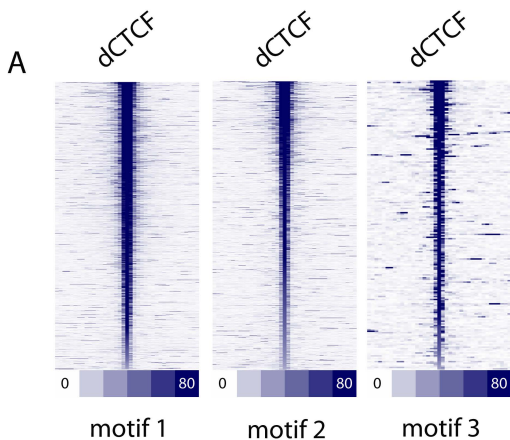
A



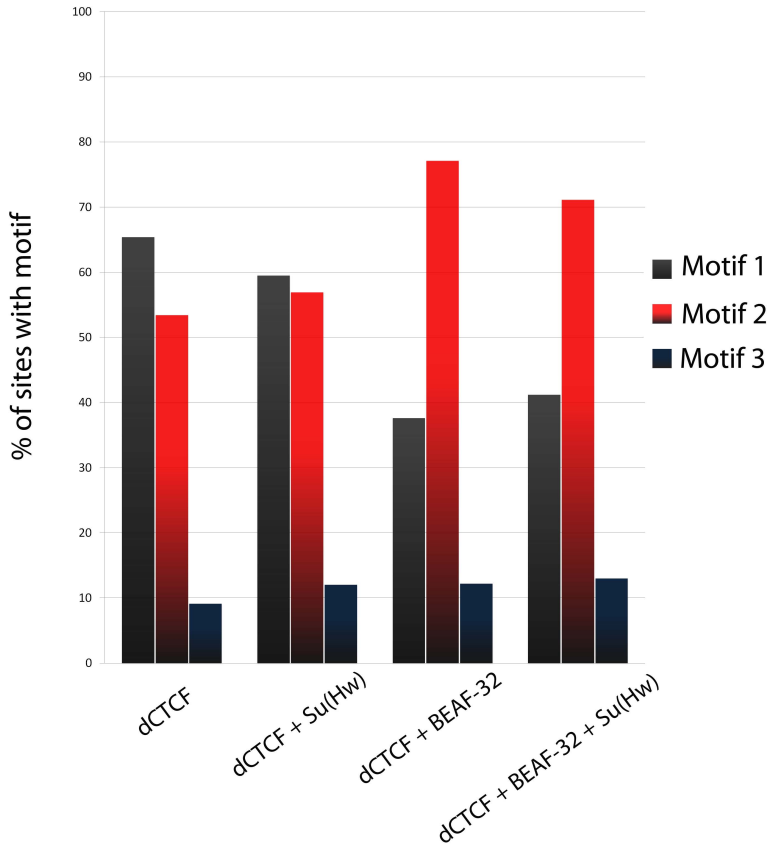
B



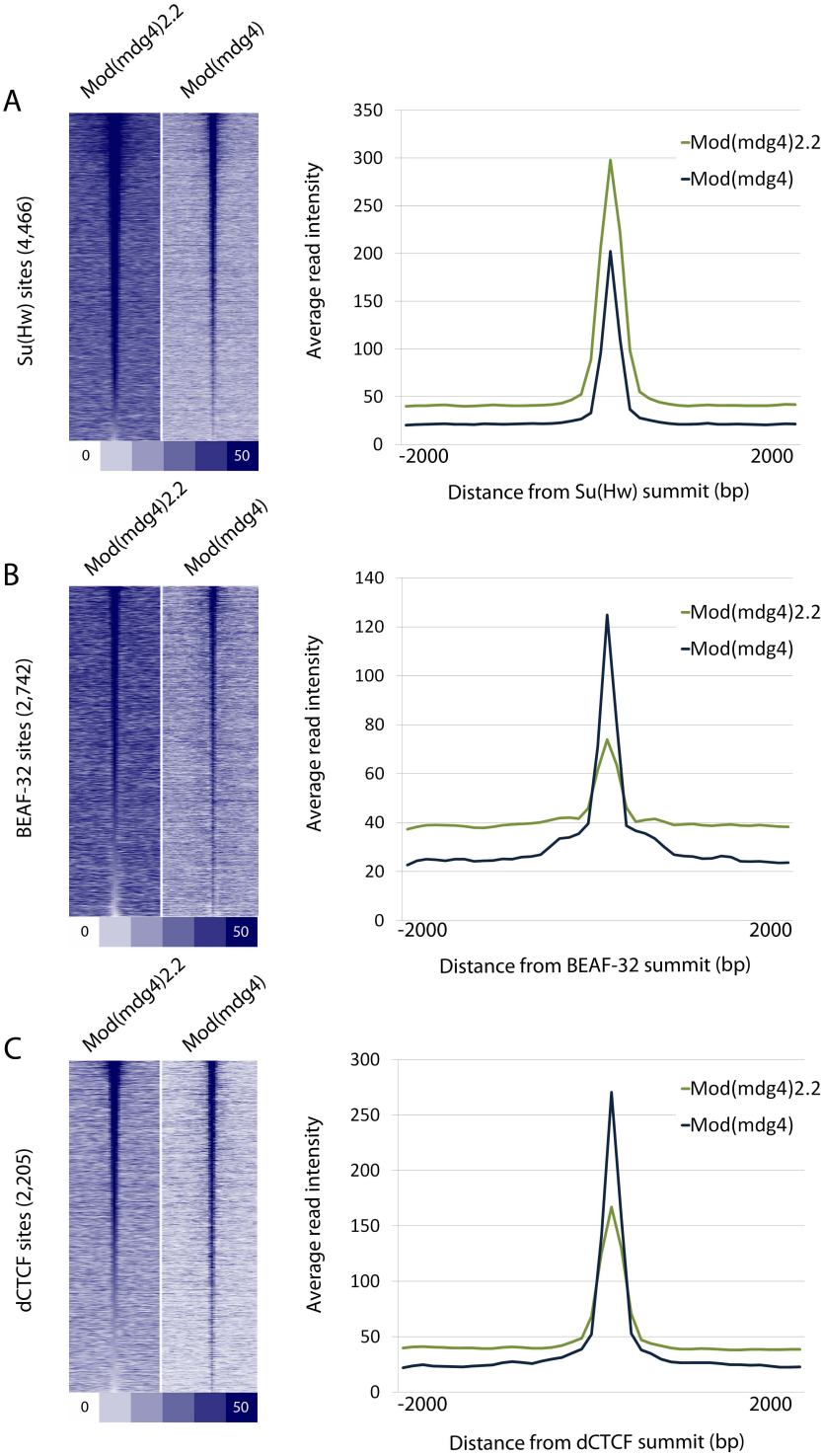
VanBortle_FigS3



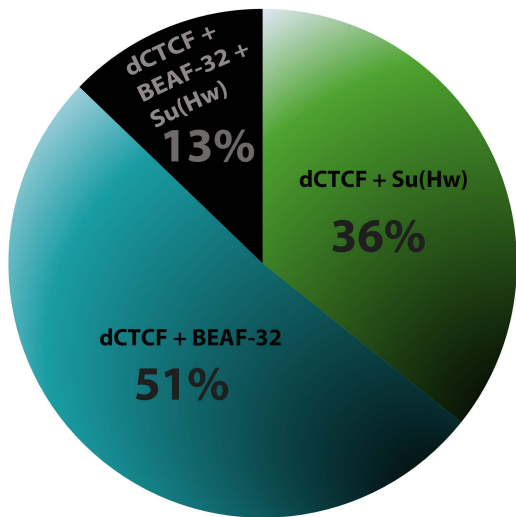
VanBortle_FigS4



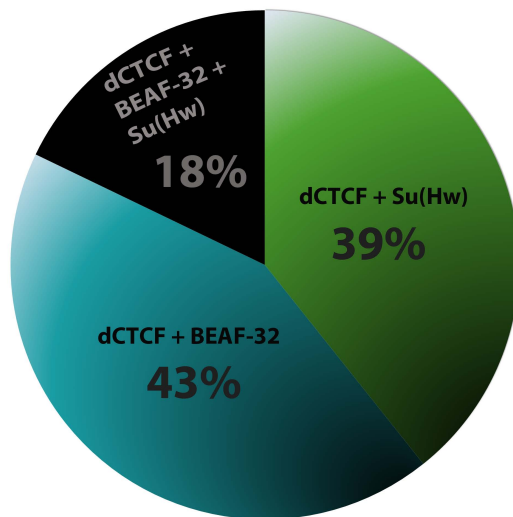
VanBortle_FigS5



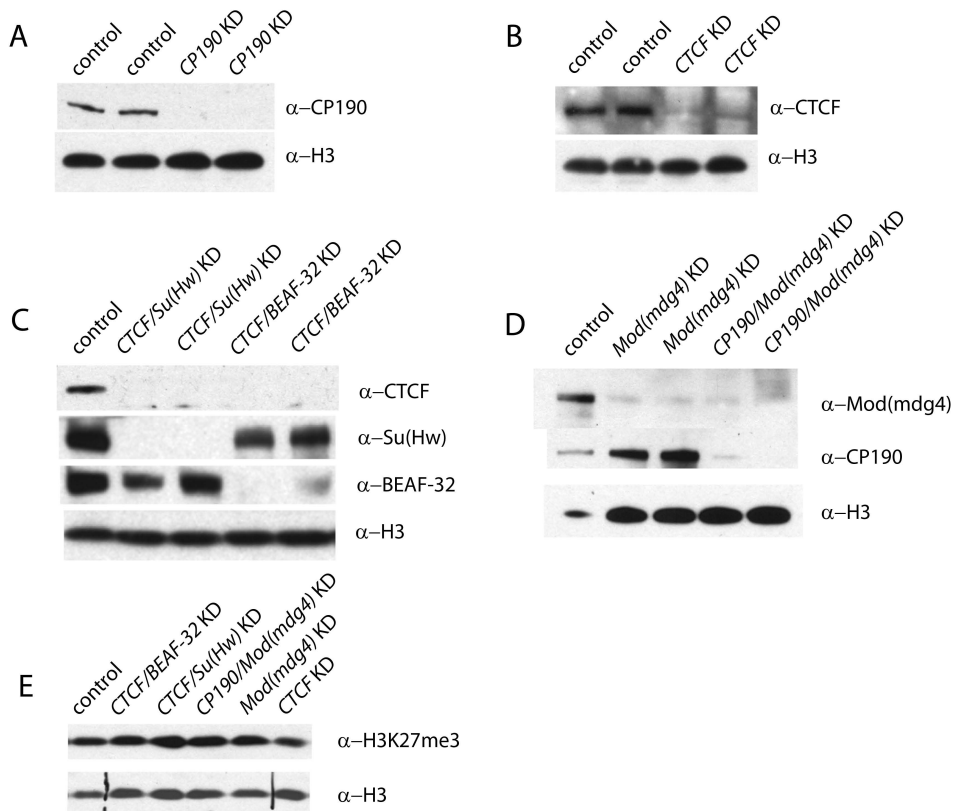
Genome-wide



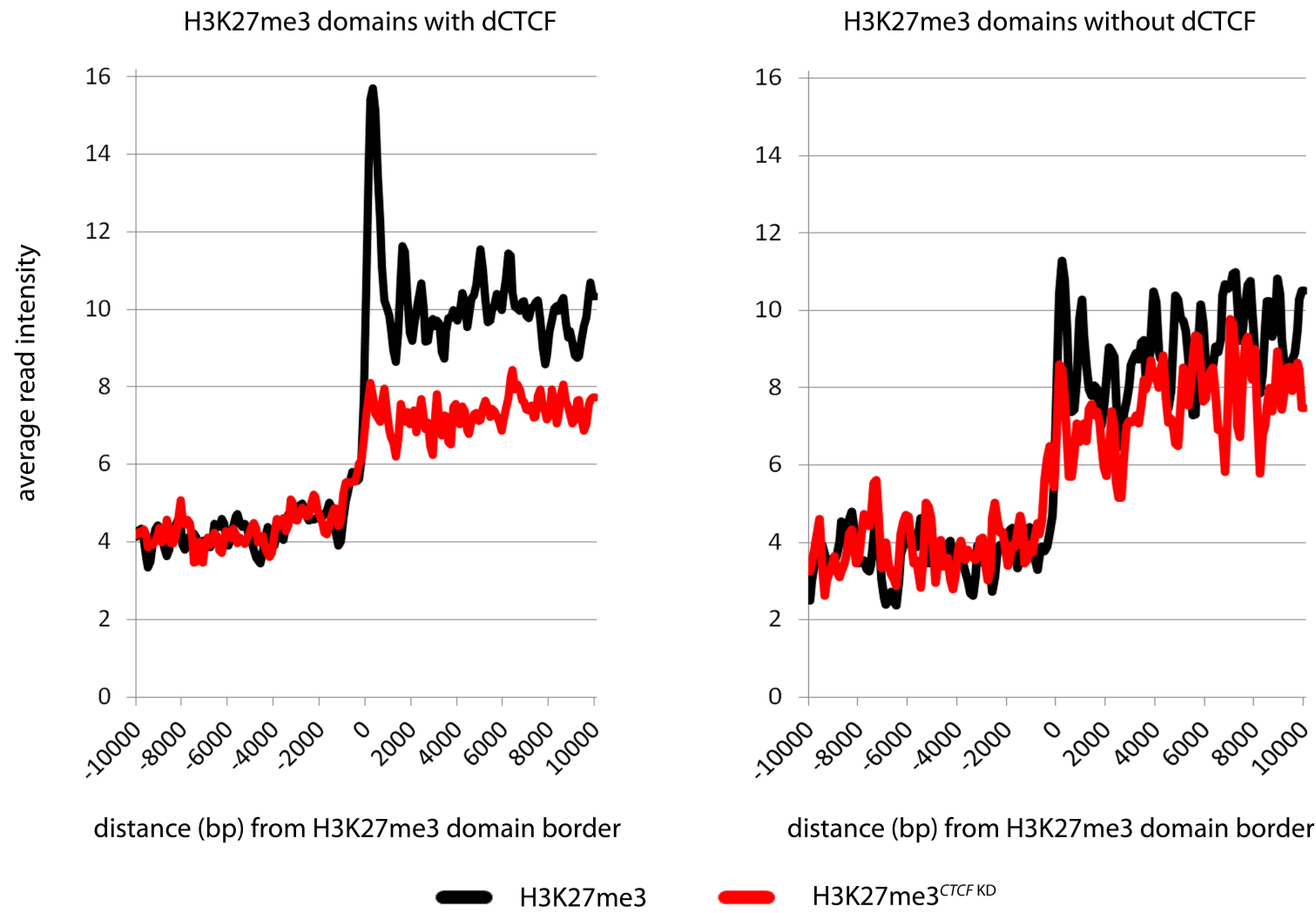
H3K27me3 Domain Borders



VanBortle_FigS7



A



B

