Published in final edited form as:

Science. 2015 February 27; 347(6225): 1017–1021. doi:10.1126/science.1262088.

CTCF establishes discrete functional chromatin domains at the *Hox* clusters during differentiation

Varun Narendra[1,2], **Pedro P. Rocha**[3], **Disi An**[4], **Ramya Raviram**[3], **Jane A. Skok**[3], **Esteban O. Mazzoni**[4,*], and **Danny Reinberg**[1,2,*]

[1]Howard Hughes Medical Institute

[2]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY 10016, USA

[3]Department of Pathology, New York University School of Medicine, New York, NY 10016, USA

[4]Department of Biology, New York University, New York, New York, 10003, USA

Abstract

Polycomb and trithorax group proteins encode the epigenetic memory of cellular positional identity by establishing inheritable domains of repressive and active chromatin within the *Hox* clusters. Here, we demonstrate that the CCCTC-binding factor (CTCF) functions to insulate these adjacent yet antagonistic chromatin domains during embryonic stem cell differentiation into cervical motor neurons. Deletion of CTCF binding sites within the *Hox* clusters results in the expansion of active chromatin into the repressive domain. CTCF functions as an insulator by organizing *Hox* clusters into spatially disjoint domains. Ablation of CTCF binding disrupts topological boundaries such that caudal *Hox* genes leave the repressed domain and become subject to transcriptional activation. Hence, CTCF is required to insulate facultative heterochromatin from impinging euchromatin to produce discrete positional identities.

Precise expression of *Hox* genes is required for cells to maintain their relative position within a developing embryo *(*1-4*)*. For example, motor neurons (MNs) rely on *Hox* gene expression for the formation of position-dependent neural circuits that control voluntary movement *(*5-7*)*. High concentration of retinoic acid (RA) signaling induces rostral *Hox* gene expression (*Hox1* to *Hox5*) and thus cervical identity to differentiating MNs *(*8*)*. The *in vivo* development of MNs with a cervical positional identity can be faithfully recapitulated *in vitro* by exposing differentiating embryonic stem cells (ESCs) to RA and a sonic hedgehog signaling agonist (smoothened agonist, SAG) (Fig. S1a) *(*9*)*. ESC-derived MNs exposed to RA activate the rostral portion of the *HoxA* cluster (*Hoxa1-6*), while *Hoxa7-13* remain repressed (Fig. 1a, Table S1) *(*10, 11*)*. The transcriptional partitioning of the *HoxA*

*Correspondence and requests for materials should be addressed to D.R. (danny.reinberg@nyumc.org) or E.O.M. (eom204@nyu.edu).

All sequencing data has been deposited to the GEO as Series GSE60240, which is currently private but accessible to the referees through the link below. All data will be made immediately available to the public upon publication: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=ohkxouosfnuvhwt&acc=GSE60240

V.N., E.O.M. and D.R. conceived the project, designed the experiments and wrote the paper. V.N. performed most of the experiments and the bioinformatic analysis. D.A. performed the immunocytochemistry. P.R., R.R and J.S. advised on the 4C-seq procedure and analysis.

cluster is mirrored at the level of chromatin. As previously described, H3K27me3 — the catalytic product of polycomb repressive complex 2 (PRC2) activity — decorates the entire *HoxA* cluster in ESCs *(*11*)* (Fig. 1b, top). Upon differentiation into MNs, H3K4me3 and RNA polymerase II (RNAPII) access the rostral segment of the cluster, whereas H3K27me3 becomes restricted to the caudal segment *(*11*)* (Fig. 1b). Within the *HoxA* cluster, MNs display two clear discontinuities in H3K4me3 and H3K27me3 density – at the intergenic region between *Hoxa5* and *Hoxa6* (C5|6), and between *Hoxa6* and *Hoxa7* (C6|7) (Fig. 1c). Interestingly, the DNA sequence underlying each of these discontinuities contains a highly conserved binding site for CTCF *(*12*)* (Fig. 1c and Fig. S1b,c) that is constitutively occupied in both ESCs and differentiated MNs (Fig. 1b and Fig. 3a, top). CTCF-demarcated chromatin boundaries were observed at the *HoxC* and *HoxD* clusters as well (Fig. 1c and Fig. S2), and have recently been identified in the orthologous bithorax complex in *Drosophila melanogaster (*13*)*.

CTCF has been suggested to function as a chromatin barrier insulator by restricting the spread of heterochromatin, though this remains in dispute *(*14-16*)*. Therefore, we tested whether CTCF can perform *Hox* gene barrier insulation during differentiation to produce functional MN circuits. We employed the CRISPR genome-editing tool *(*17, 18*)* in ESCs to disrupt CTCF binding sites that localize to chromatin boundaries within *Hox* clusters. We first generated a 9bp homozygous deletion within the core CTCF motif between *Hoxa5* and *Hoxa6* (Δ5|6) (Fig. 1d), and did not detect any mutations at potential off-target cleavage sites (Table S2). The 9bp deletion results in a total abrogation of CTCF occupancy (Fig. 1e). The neighboring CTCF binding site (C6|7) also shows a dramatic reduction in binding, suggesting an interdependence (Fig. 1e) *(*19, 20*)*. Importantly, Δ5|6 ESCs exhibit no defect in their ability to differentiate into MNs (Fig. S3). To examine the transcriptional consequence of deleting CTCF binding sites within the *HoxA* cluster in response to patterning signals during cell differentiation, we performed RNA-seq on wild-type (WT) and Δ5|6 cells at two stages: ESCs and differentiated MNs. In ESCs, all *HoxA* genes are repressed in both lines (Fig. 2a, left and Table S1). Upon differentiation, *Hoxa1-6* are activated in the wild-type setting, whereas *Hoxa7-13* remain repressed, mirroring the distribution of active and repressive chromatin across the cluster. *Hoxa1-6* are equivalently activated in WT and Δ5|6 MNs. However, *Hoxa7* – the gene located immediately caudal to the affected C6|7 site – is upregulated more than 25-fold relative to the WT control. *Hoxa9* shows very modest expression in Δ5|6 MNs, whereas *Hoxa10-13* remain fully repressed (Fig. 2a, right and Table S1). Furthermore, while *Hoxa6* – the gene located between the deleted C5|6 and C6|7 site – is equivalently expressed in terminally differentiated WT and Δ5|6 MNs, it is transcriptionally activated earlier in differentiating Δ5|6 cells than in WT cells, unlike the rostral *Hoxa5* control (Fig. S4). Thus, CTCF occupancy regulates the spatial and temporal activation of the *HoxA* cluster. Demonstrating that CTCF boundary activity is not restricted to a single *Hox* cluster, deletion of a 13bp sequence within a binding site at the *HoxC* chromatin boundary results in the equivalent transcriptional activation of genes located caudal to the site of mutation (Fig. S5).

Hoxa7-specific transcriptional activation in Δ5|6 MNs suggests that the intact C7|9 peak serves as a new boundary. To study if there is a relocation of the chromatin boundary during

MN differentiation in the mutant line, we investigated the chromatin state of ESCs and differentiated MNs. Site-specific ablation of CTCF does not affect the chromatin state of undifferentiated cells, as WT and Δ5|6 ESCs possess H3K27me3 distributed across the entire *HoxA* cluster (Fig. 2b, top). However after differentiation, Δ5|6 MNs exhibit a 50% reduction in H3K27me3 levels specifically within the region delimited by C5|6 and C7|9 (Fig. 2b and Fig. S6a,c). In agreement with C7|9 serving as the new boundary element in Δ5|6 MNs, H3K27me3 density recovers to wild-type levels immediately caudal to the C7|9 peak. Moreover, deletion of C5|6 results in a complementary expansion of H3K4me3 and RNAPII up to the C7|9 boundary (Fig. 2b and Fig. S6b,d). The Δ5|6 mutation does not produce pleiotropic effects, as chromatin boundaries are not disrupted in *trans* within the *HoxC and HoxD* clusters (Fig. S6c,d,e). Likewise, ablation of the C5|6 CTCF binding event within the *HoxC* cluster ($\Delta 5|6_{HoxC}$) results in an equivalent chromatin boundary relocation (Fig. S5). Thus, CTCF does not function within the *Hox* clusters according to the traditional definition of a chromatin insulator - to restrict the spread of repressive chromatin into adjacent euchromatin – but rather to restrict *in cis* the exposure of polycomb repressed genes to trithorax activity.

CTCF-dependent insulation occurs via its ability to mediate looping interactions between non-adjacent segments of DNA *(*21*)*. Accordingly, CTCF is enriched at boundaries between topologically associated domains (TADs) *(*15, 22, 23*)*. To test how CTCF-mediated looping may regulate the dynamic spatial reorganization of the *HoxA* cluster during differentiation, we performed 4C-seq in WT and Δ5|6 cells using viewpoints located within either the transcriptionally active (4C.Hoxa5-A) or repressive (4C.Hoxa10) domains of the *HoxA* cluster (Fig. 3). In WT and Δ5|6 ESCs, the strong interaction signal of both 4C-seq viewpoints extends to the perimeter of the *HoxA* cluster, suggesting an organization of the locus as a single architectural domain which the C5|6 binding site does not alter (Fig. 3a and Fig. S7a,b). As expected, in WT cells this domain partitions during differentiation into two at roughly the C6|7 position, mirroring the distribution of H3K4me3 in the rostral domain and H3K27me3 in the caudal domain (24, 25). This is demonstrated by the strong interactions with the 4C.Hoxa5-A viewpoint that occur almost exclusively within the rostral domain (Fig. 3b), and interactions with the 4C.Hoxa10 viewpoint that are restricted to the caudal domain (Fig. 3c). Unlike as is the case in ESCs, deletion of the C5|6 CTCF binding site affects the spatial organization of the *HoxA* cluster in MNs. The Δ5|6 mutation repositions the topological boundary in MNs to the intact C7|9 site, matching the *de novo* chromatin boundary and thereby evicting *Hoxa7* from the caudal repressed domain and into the rostral active domain (Fig. 3b,c and Fig. S8a,b). Thus, the elimination of a CTCF binding site causes a structural reorganization of the *HoxA* cluster that results in an aberrant chromatin boundary and altered gene expression.

This data argues that in response to RA signaling, the most rostral CTCF binding event forges a topological boundary within the *HoxA* cluster that can insulate active from repressive chromatin and thus maintain proper gene expression. This predicts that eliminating the C7|9 CTCF binding site in Δ5|6 MNs would cause aberrant activation of *Hoxa7-10*, and caudal regression of the topological boundary to the C10|11 position. Using the CRISPR genome-editing tool in Δ5|6 ESCs, we mutated the C7|9 CTCF binding site. Δ5|

6:7|9 ESCs harbor a 21bp deletion spanning the C7|9 motif on one allele. The other allele contains a 3bp deletion and an 18bp insertion that disrupts the motif (Fig. S9). *Hoxa7-10* are highly upregulated in the double mutant MNs relative to the WT control (Fig. 4a,b, Fig. S4 and Table S1). *Hoxa9-10* are the most significantly upregulated genes in the polyA-selected transcriptome, whereas *Hoxa11-13* remain transcriptionally silent. Importantly, this phenotype is specific to CTCF ablation, as deletion of a YY1 binding motif adjacent to the C7|9 site does not result in the transcriptional activation of caudal genes (Fig. S10) (26). The transcriptional profile of Δ5|6:7|9 MNs suggests an underlying caudal boundary shift. Accordingly, 4C-seq using the active Hoxa5-B viewpoint shows a shift of the topological boundary from C6|7 to the intact C10|11 position in Δ5|6:7|9 MNs (Fig. 4c and Fig. S8c), allowing for a parallel expansion of H3K4me3 onto the *Hoxa10* gene (Fig. 4d). Conversely, H3K27me3 density progressively decreases relative to the WT control in a rostral direction from the C10|11 CTCF site.

These results indicate that in response to patterning signals during differentiation, CTCF partitions the *Hox* clusters into insulated architectural domains, upon which trithorax and polycomb activities are superimposed in a mutually exclusive fashion to establish discrete *Hox* transcriptional programs. In agreement with our findings, deletion of a CTCF binding site at the boundary of a polycomb domain containing the *Tcfap2e* locus resulted in its transcriptional activation (27). It remains to be tested whether the expansion of H3K4me3 activity we observe in the *Hox* clusters is the result of aberrant enhancer contacts with caudal genes or an alternative local mechanism of trithorax expansion. Our 4C-Seq results agree with previous studies, which have shown that the caudal and rostral domains of *HoxA* cluster in differentiated cells are incorporated into separate adjacent TADs, the border of which aligns with the chromatin boundary. Our findings thus imply that CTCF is functionally required to delimit TAD boundaries, though a high-resolution all-vs-all (Hi-C) approach will be required to confirm this claim.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Luis Alejandro Rojas, Courtney Leek, Anupriya Singhal and Drs. Shengjiang Tu, Roberto Bonasio and Lynne Vales for thoughtful discussions and revision of the manuscript. We also thank the NYU Genome Technology Center for help with sequencing. This work was supported by grants from the US National Institutes of Health (R37-37120 and GM-64844 to D.R., T32 GM007238 to V.N., R01HD079682 to E.O.M., and GM086852, GM112192 to J.S.). D.A. was supported by the Project ALS foundation. P.R. is an NCC post-doctoral fellow. JS is an LLS scholar.

REFERENCES AND NOTES

1. Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. Science. 2010; 330(6004): 612–6. [PubMed: 21030644]
2. Ingham PW. Differential expression of bithorax complex genes in the absence of the extra sex combs and trithorax genes. Nature. 1983:306. [PubMed: 6877352]
3. Lewis EB. A gene complex controlling segmentation in Drosophila. Nature. 1978; 276(5688):565–70. [PubMed: 103000]

4. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011; 469(7330):343–9. [PubMed: 21248841]

5. Dasen JS, Tice BC, Brenner-Morton S, Jessell TM. A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. Cell. 2005; 123(3):477–91. [PubMed: 16269338]

6. Jung H, Lacombe J, Mazzoni EO, Liem KF, Grinstein J, Mahony S, et al. Global control of motor neuron topography mediated by the repressive actions of a single hox gene. Neuron. 2010; 67(5): 781–96. [PubMed: 20826310]

7. Wu Y, Wang G, Scott SA, Capecchi MR. Hoxc10 and Hoxd10 regulate mouse columnar, divisional and motor pool identity of lumbar motoneurons. Development (Cambridge, England). 2007; 135(1): 171–82.

8. Liu JP, Laufer E, Jessell TM. Assigning the positional identity of spinal motor neurons: rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. Neuron. 2001; 32(6):997–1012. [PubMed: 11754833]

9. Wichterle H, Lieberam I, Porter JA, Jessell TM. Directed differentiation of embryonic stem cells into motor neurons. Cell. 2002; 110(3):385–97. [PubMed: 12176325]

10. Mazzoni EO, Mahony S, Closser M, Morrison CA, Nedelec S, Williams DJ, et al. Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. Nature neuroscience. 2013; 16(9):1219–27.

11. Mazzoni EO, Mahony S, Peljto M, Patel T, Thornton SR, McCuine S, et al. Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. Nature neuroscience. 2013; 16(9): 1191–8.

12. Rhee HS, Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell. 2011; 147(6):1408–19. [PubMed: 22153082]

13. Bowman SK, Deaton AM, Domingues H, Wang PI, Sadreyev RI, Kingston RE, et al. H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex. Elife. 2014; 3:e02833. [PubMed: 25082344]

14. Cuddapah S, Jothi R, Schones DE, Roh T-YY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome research. 2008; 19(1):24–32. [PubMed: 19056695]

15. Phillips-Cremins JE, Corces VG. Chromatin insulators: linking genome organization to cellular function. Molecular cell. 2013; 50(4):461–74. [PubMed: 23706817]

16. Recillas-Targa F, Pikaart MJ, Burgess-Beusse B, Bell AC, Litt MD, West AG, et al. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99(10): 6883–8. [PubMed: 12011446]

17. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337(6096):816–21. [PubMed: 22745249]

18. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature protocols. 2013; 8(11):2281–308.

19. Pant V, Kurukuti S, Pugacheva E, Shamsuddin S, Mariano P, Renkawitz R, et al. Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. Molecular and cellular biology. 2004; 24(8):3497–504. [PubMed: 15060168]

20. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Molecular cell. 2004

21. Splinter E, Heath H, Kooren J, Palstra R-JJ, Klous P, Grosveld F, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development. 2006; 20(17):2349–54. [PubMed: 16951251]

22. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485(7398):376–80. [PubMed: 22495300]

23. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012; 485(7398):381–5. [PubMed: 22495304]

24. Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D. Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. eLife. 2013:3.

25. Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. The dynamic architecture of Hox gene clusters. Science (New York, NY). 2011; 334(6053):222–5.

26. Vella P, Barozzi I, Cuomo A, Bonaldi T, Pasini D. Yin Yang 1 extends the Myc-related transcription factors network in embryonic stem cells. Nucleic Acids Res. 2012; 40(8):3403–18. [PubMed: 22210892]

27. Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, et al. Control of cell identity genes occurs in insulated neighborhoods in Mammalian chromosomes. Cell. 2014; 159(2):374–87. [PubMed: 25303531]

28. van de Werken HJ, Landan G, Holwerda SJ, Hoichman M, Klous P, Chachik R, et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nature methods. 2012; 9(10):969–72. [PubMed: 22961246]

29. Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, et al. Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res. 2009; 37(18):e123. [PubMed: 19620212]

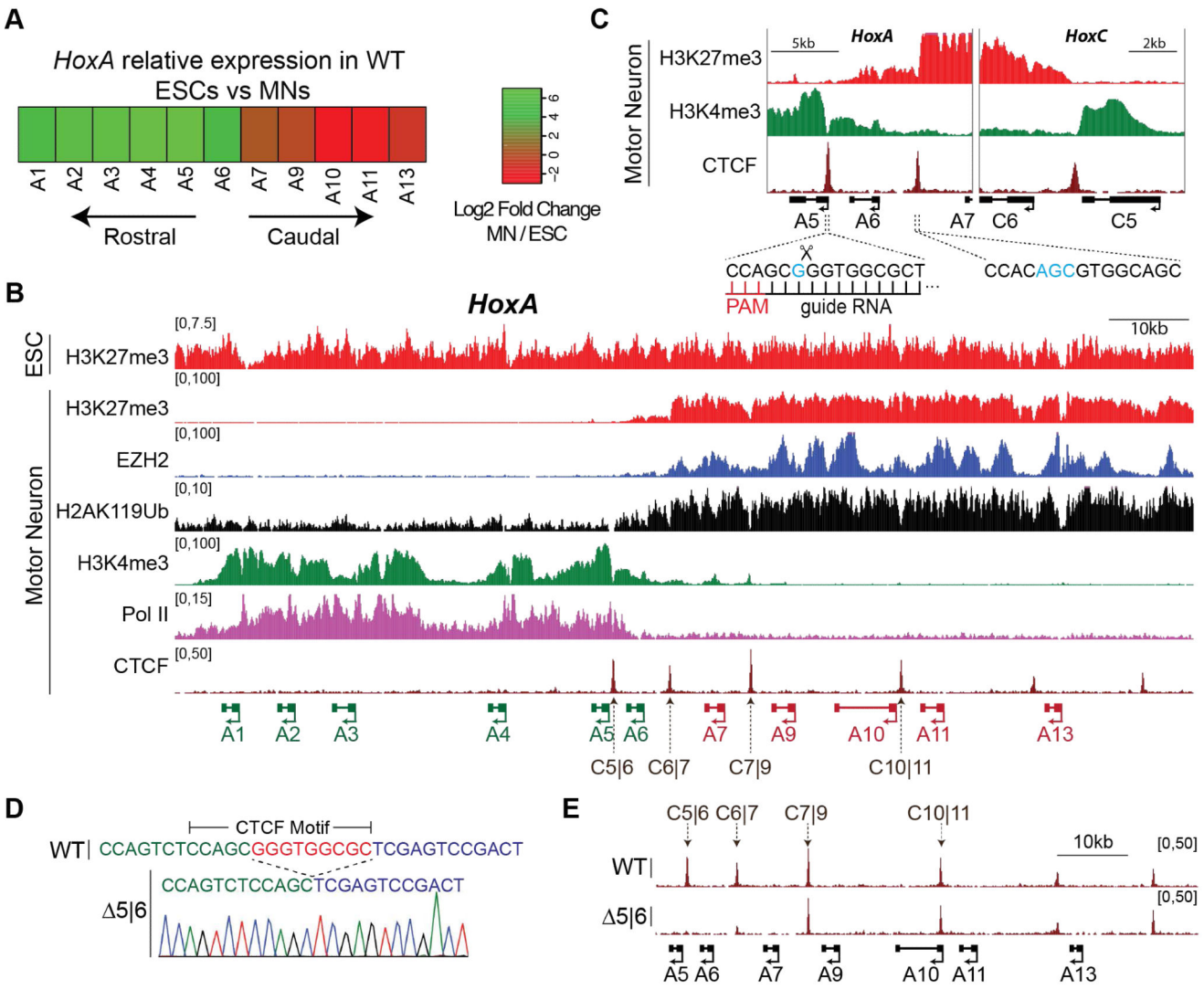

Fig. 1. CTCF localizes to a *HoxA* chromatin boundary in motor neurons

A. Heatmap of *HoxA* relative expression (log2) between WT ESCs and MNs.

B. Normalized ChIP-seq read densities for the indicated proteins/modifications in ESCs and MNs from two merged biological replicates. Genes that are activated during differentiation are annotated in green; repressed in red.

C. Zoomed-in view of H3K4me3 and H3K27me3 boundaries, along with CTCF peaks and their underlying binding motifs. Blue highlights nucleotides that diverge from the consensus motif. gRNA used to target C5|6 is shown.

D. Sequencing chromatogram of Δ5|6 line depicts a 9bp deletion overlapping the CTCF core motif.

E. Normalized ChIP-seq read densities for CTCF in WT and Δ5|6 MNs from 2 merged biological replicates. The deleted CTCF binding site (C5|6) is boxed, as well as the neighboring site (C6|7).

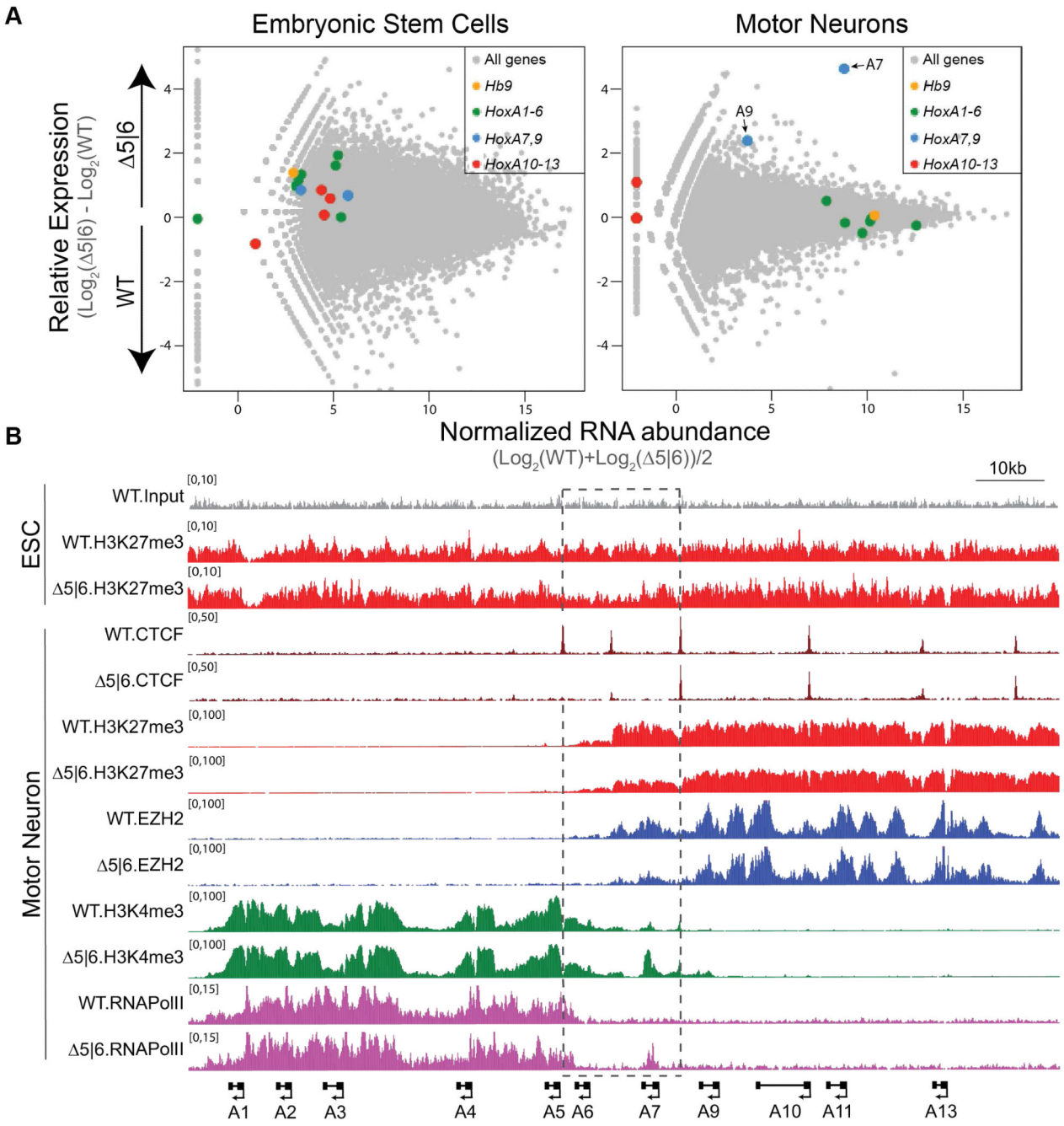

Fig. 2. Chromatin boundary is disrupted upon deletion of C5|6 CTCF motif

A. RNA-seq MA plot of WT vs. Δ5|6 ESCs (left) and MNs (4 days after RA/SAG, right). MN data is representative of two biological replicate experiments, ESC data of one. Mean abundance is plotted on the *x*-axis and enrichment is plotted on the *y*-axis. Hb9 is a marker of motor neurons.

B. Normalized ChIP-seq read densities for the indicated protein/modifications along the *HoxA* cluster in ESCs and MNs (4 days after RA/SAG) from two biological replicates.

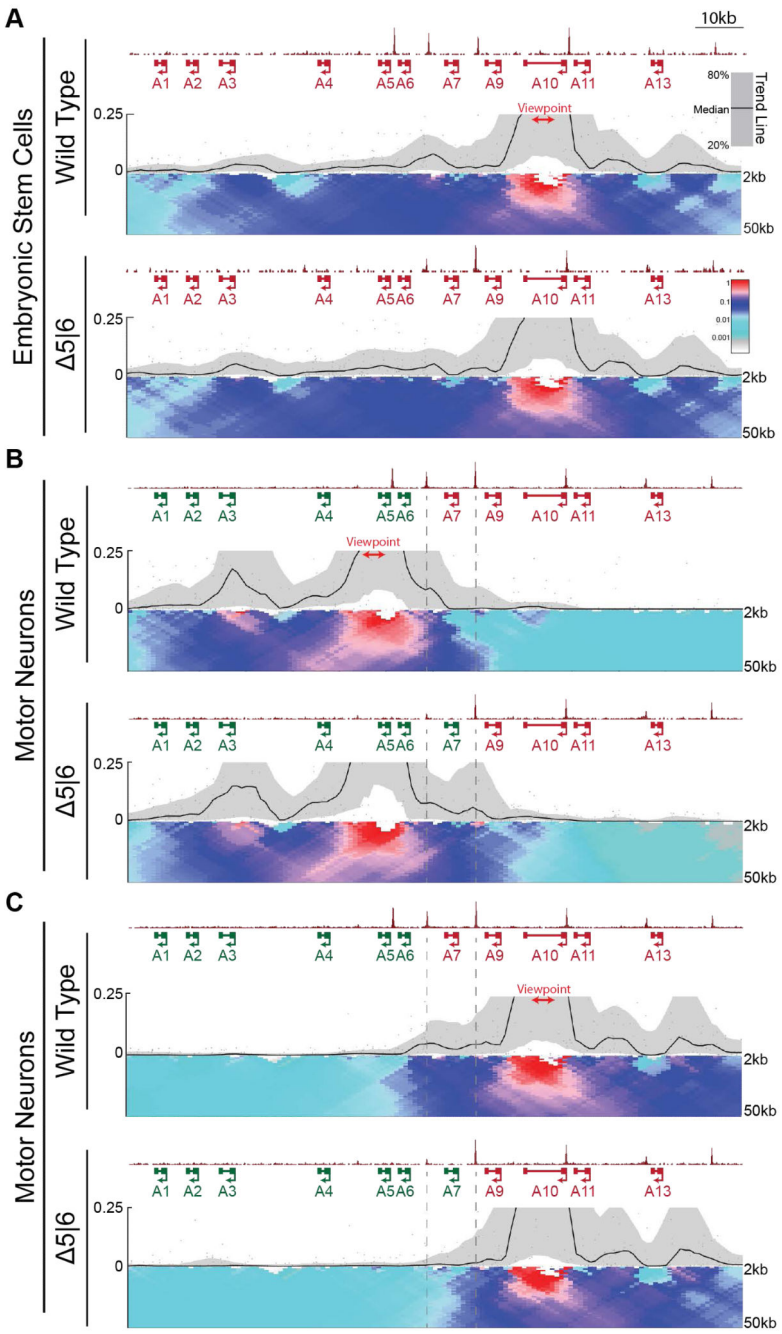

Fig. 3. Loss of CTCF alters topological architecture of the *HoxA* locus

A,B. Normalized ChIP-seq read densities for CTCF and 4C contact profiles in WT and Δ5|6 ESCs **(A)** and MNs **(B,C)** using a viewpoint (red) in either the rostral **(B,** 4C.Hoxa5-A**)**, or caudal **(A,C,** 4C.Hoxa10**)** segment of the cluster. ChIP signal is merged across two biological replicates, and 4C signal across three replicates. The median and 20th/80th percentile of sliding 5kb windows determine the main trendline. Color scale represents enrichment relative to the maximum attainable 12-kb median value. Dotted lines highlight the region between C6|7 and C7|9.

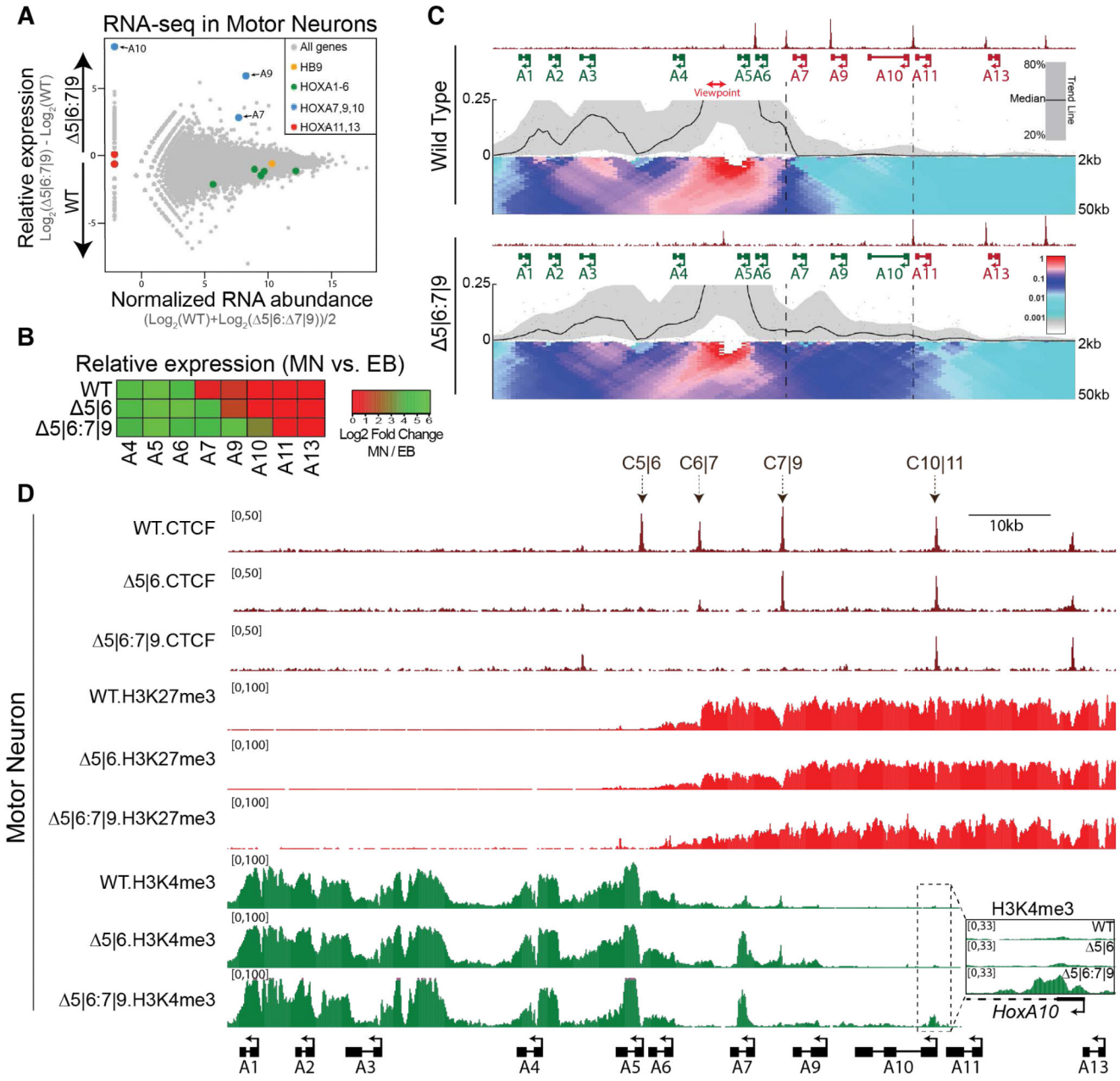

Fig. 4. Compound C5|6:7|9 deletion causes a further caudal spread of active transcription within the *HoxA* locus

A. RNA-seq MA plot of WT vs. Δ5|6:7|9 MNs. Mean abundance is plotted on the *x*-axis and enrichment is plotted on the *y*-axis.

B. Heatmap of *HoxA* relative expression in MNs (Day 4) vs. EBs (embryoid bodies, Day 0) across two biological replicates (single replicate in the Δ5|6:7|9 line).

C. Normalized ChIP-seq read densities for CTCF and 4C contact profiles in WT and Δ5|6:7|9 MNs using the 4C.Hoxa5-B viewpoint (red) from two biological replicates. The median and 20th/80th percentile of sliding 5kb windows determine the main trendline. Color scale

represents enrichment relative to the maximum attainable 12-kb median value. Dotted lines highlight the region between C6|7 and C10|11.

D. Normalized ChIP-seq read densities for the indicated protein/modifications along the *HoxA* cluster in MNs (four days after RA/SAG). A magnified view of the boxed region is presented on the right.