Dear Editor,

Similar to higher-order folding of polypeptide chains into functional proteins, linear DNA molecules are spatially folded in a hierarchical and dynamic manner into three-dimensional (3D) functional chromatin structures in eukaryotic nuclei (Huang and Wu, 2016; Rowley and Corces, 2018). This dynamic folding is closely related to many nuclear processes such as DNA replication and repair, chromosomal translocation, recombination, and segregation, as well as RNA transcription, splicing, and transport. In particular, dynamic long-distance chromatin looping interactions, which result in close spatial contacts between distal enhancers and target promoters, are thought to play a role in controlling precise spatiotemporal as well as cell-type specific gene expression during animal development (Rowley and Corces, 2018). Mammalian genomes contain numerous noncoding regulatory elements that regulate these dynamic long-distance chromatin looping interactions. Specially, one type of genetic elements, known as insulators, plays a delicate role to ensure proper activation of target promoters by distal enhancers (Huang and Wu, 2016). In mammals, the most prominent insulator-binding protein is CCCTC binding factor (CTCF), an architectural protein with 11 zinc-fingers essential for 3D genome organization (Yin et al., 2017; Wu et al., 2019).

CTCF dynamically and directionally binds to hundreds of thousands genomic sites and this binding is pivotal for its multivalent role in many cellular and developmental processes (Guo et al., 2015; Yin et al., 2017). In particular, CTCF, in collaboration with its associated cohesin complex, determines V(D)J recombination of the Bcr gene clusters (Huang and Wu, 2016; Rowley and Corces, 2018). This dynamic folding is closely related to many nuclear processes such as DNA replication and repair, chromosomal translocation, recombination, and segregation, as well as RNA transcription, splicing, and transport. In particular, dynamic long-distance chromatin looping interactions, which result in close spatial contacts between distal enhancers and target promoters, are thought to play a role in controlling precise spatiotemporal as well as cell-type specific gene expression during animal development (Rowley and Corces, 2018).

Mammalian genomes contain numerous noncoding regulatory elements that regulate these dynamic long-distance chromatin looping interactions. Specially, one type of genetic elements, known as insulators, plays a delicate role to ensure proper activation of target promoters by distal enhancers (Huang and Wu, 2016). In mammals, the most prominent insulator-binding protein is CCCTC binding factor (CTCF), an architectural protein with 11 zinc-fingers essential for 3D genome organization (Yin et al., 2017; Wu et al., 2019).

Genetic evidence for asymmetric blocking of higher-order chromatin structure by CTCF/cohesin

To further investigate the role of enhancers and insulators as well as their locations and relative orientations in 3D genome architecture and gene regulation, we designed a...
CTCF determines chromatin organization of clustered protocadherins

Protein & Cell
Figure 1. Genetic dissection of the Pcdha HS5-1 enhancer architecture by CRISPR DNA-fragment editing. (A) Schematics of the three human Pcdh gene clusters (hPcdh α, β, and γ) and the inversion experiments mediated by CRISPR DNA-fragment editing. The orientation of CBS arrays is indicated by horizontal arrowheads. The DNase I hypersensitive sites within super-enhancers are indicated by vertical arrows. The HS5-1 enhancer and its various CRISPR inversions are highlighted below. Two tandem CBSs in a reverse orientation (HS5-1a and HS5-1b) at the editing sites are also highlighted. SLC25A2 and TAF7 are two non-Pcdh genes. CBS, CTCF-binding site; HS, hypersensitive site; SE, super-enhancer; TF, transcription factor. (B) Shown are 4C chromatin-interaction profiles in the wild-type control (Ctr) and CRISPR cell lines with inversion of the CBS HS5-1b only (Inv I) or the HS5-1b combined with the middle region (Inv II) using the HS5-1 enhancer as a viewpoint. The significance of interactions (P value) is shown under the read’s density. The log2 ratios between each inversion cell clone and wild-type cells are also indicated. (C) Shown are 4C chromatin-interaction profiles in wild-type control and CRISPR cell lines with inversion of the CBS HS5-1a only (Inv III) or the HS5-1a combined with the middle region (Inv IV) using the HS5-1 enhancer as a viewpoint. (D) Heatmap shows the alteration of expression patterns of members of the three Pcdh gene clusters resulting from CRISPR inversion of different enhancer and insulator (CBS) elements by RNA-seq experiments.

series of dual sgRNAs to dissect the HS5-1 enhancer architecture (Fig. 1A) (Shou et al., 2018). We made use of human HEC-1-B cells and screened single-cell CRISPR clones for inversion of the single CBS HS5-1a or HS5-1b site as well as its combination with the middle enhancer region of the HS5-1 super-enhancer (Fig. S3). Remarkably, inversion of the CBS HS5-1b only or with the middle enhancer region results in a significant decrease of long-distance DNA-looping interactions between the HS5-1 enhancer and Pcdha promoters, and a corresponding increase of long-distance DNA-looping interactions with the Pcdhβ promoters (Fig. 1B). However, inversion of the CBS HS5-1a only or combined with the middle enhancer region results in no significant alteration of long-distance DNA-looping interactions with either the Pcdha or Pcdhβ promoters (Fig. 1C). Thus, the relative orientations of the single CBS HS5-1b at the Pcdha chromatin domain CCD or subTAD boundary determine the chromatin-looping directions between the distal enhancer and its target promoters.

To assess functional consequences of the altered chromatin looping, we performed RNA-seq experiments and found that each CRISPR inversion cell line with disrupted architecture of the HS5-1 enhancer displays altered patterns of Pcdh gene expression (Fig. 1D). Finally, a series of CRISPR inversions of progressive numbers of CBSs in the β-globin locus demonstrate that only those inversions covering the CBS at a TAD boundary (CBS15) switch chromatin-looping directions (Figs. S4 and S5). We conclude that single CBSs at domain boundaries determine chromatin looping orientation and function as insulators.

The super-enhancer of the Pcdh β clusters, which is located downstream of Pcdhβ cluster, contains tandem reverse CBSs and one forward CBS (Figs. 1A, S1A and S1F) (Yokota et al., 2011; Guo et al., 2015). We analyzed the published transcription-factor binding landscape in the Pcdhβ regulatory region of SK-N-SH cells and found that it is enriched with high levels of active marks of P300, Pol2, and H3K4me3, as well as CTCF and its associated cohesin complex (Fig. S2C) (Guo et al., 2015). Similar to the SK-N-SH cells, this region is enriched with high levels of P300, Pol2, H3K4me1, H3K27ac, and CTCF, as well as Med1, Oct4, Sox2, and Nanog in mouse ES cells (Fig. S2D). Moreover, this region is also marked with H3K27ac and H3K4me1 in mouse brain tissues, suggesting that it is an active super-enhancer in vivo (Guo et al., 2015). Finally, systematic characterization of the region revealed specific enrichments of CTCF proteins. Together, this suggests that the CTCF-enriched regulatory region is a super-enhancer regulating the 3D chromatin architecture of the Pcdhβ clusters.

Next, we made use of the murine neuroblastoma cell line Neuro2A to investigate the chromatin architecture of the super-enhancer and its regulation of the Pcdhβ promoters. We performed RNA-seq and circularized chromosome conformation capture (4C) experiments and found that HS7L, HS5-1βL, and HS18-20 form specific long-distance chromatin interactions with the promoter regions of the expressed Pcdhβ genes in Neuro2A cells (Fig. 2A and 2B). These
results indicated that the active Pcdhβ promoters, for example Pcdh β11/12, β17, β22, yb1, yb8, and yc3, are in spatial contacts with the super-enhancer (Fig. 2A and 2B).

To study how the super-enhancer downstream of the Pcdhβ cluster regulates the higher-order chromatin architecture of the Pcdhβ clusters in vivo, we generated the super-enhancer CBSs b-e inversion CRISPR mice by DNA-fragment editing (Fig. 2C and 2D). We then performed 5C (chromosome conformation capture carbon copy) and RNA-seq experiments using the CBSs b-e inverted mouse brain tissues and found that there is a significant decrease of long-distance chromatin interactions between the super-enhancer and members of the Pcdhβ cluster (Fig. 2E). Consistent with the 5C experiments, RNA-seq experiments revealed that there is a significant decrease of expression levels of the Pcdhβ genes (Fig. 2F). Intriguingly, the expression levels of Pcdhya7 are significantly increased (Fig. 2F).

The variable regions of the three Pcdh clusters contain forward-oriented tandem arrays of CBSs (Fig. S1A–E). To investigate the roles of these variable-region CBSs in the regulation of higher-order chromatin structure of the clustered Pcdh genes, we generated large deletions of forward-oriented tandem-arrayed CBSs by Cre/LoxP-mediated site-specific trans-allelic recombination in mice in vivo (Fig. S6A). Specifically, we bred mice to generate delα and delαβ mouse lines with the deletion of 24 or 45 forward-oriented tandem CBSs, respectively, but with the Pcdha1 CSE intact (Fig. S6B and S6C). Conformation capture 4C experiments with HS5-1, γα3, or HSS-1bL as a viewpoint (VP) revealed that there is a significant increase of long-distance chromatin interactions between the viewpoint and Pcdha1 CSE (Fig. S6D–F), suggesting that these large tandem-arrayed promoter CBSs function as insulators for the Pcdha1 promoter. By contrast, there appears no significant alteration of long-distance chromatin interactions between the viewpoint and CBSs located downstream of the large deletions (Fig. S6E and S6F). These data demonstrated that the tandem-arrayed promoter CBSs function as insulators and that these insulators alter long-distance chromatin interactions of upstream but not downstream CBSs, thus functioning in an asymmetric manner in vivo.

To investigate whether the insulation activity of CBSs is orientation-dependent in vitro, we generated a series of luciferase reporter constructs (Fig. S6G). Interestingly, the pair of tandem CBSs in the forward-reverse orientation flanking the luciferase reporter has the strongest enhancer-blocking activity (Fig. S6H). This demonstrates that CBSs function as enhancer-blocking insulators in an orientation-dependent manner in the luciferase reporter assay in vitro, similar to functioning in the none-orientation-independent manner in vivo.

Recent studies revealed that the locations and relative orientations of CBSs genome-wide determine the 3D chromosomal architecture (Rowley and Corces, 2018). In particular, the orientation and relative locations of a repertoire of Pcdh and β-globin CBSs determine the directional looping between the distal enhancer and its target promoters (Guo et al., 2015). The HSS-1 enhancer is flanked by two reverse CBSs (Fig. 1A) and there is no evidence to date that single CBS sites can determine chromatin-looping directions between enhancers and promoters. To this end, we inverted reverse-oriented single CBSs at chromatin domain boundaries and found that single CBSs determine the direction of chromatin looping in the Pcdh and β-globin gene clusters.

G

Cohesin loading

Cohesin easy sliding through

Cohesin difficult sliding through

Insulator (CBS) orientation

Sliding direction

Blocking sliding

Figure 2. continued.
CTCF determines chromatin organization of clustered protocadherins

(Figs. 1 and S4). In addition, deletion of large regions containing forward-oriented tandem-CBS arrays in mice in vivo demonstrated asymmetric influences of chromatin looping between super-enhancers and target promoters (Fig. S6). Finally, our CBSs b-e inversion in mice in vivo is consistent with previous deletion experiments that this region mainly enhances promoter activities of members of the Pcdhβ cluster (Yokota et al., 2011). Thus, the super-enhancer is pivotal to Pcdhβ expression and both deletion and inversion of its CBSs abolish long-distance chromatin looping to the distal Pcdhβ cluster beyond the proximal Pcdhγ cluster. In addition, the super-enhancer also regulates members of the Pcdhβ cluster by forming asymmetric chromatin looping interactions (Figs. 2B, 2E, S6E, and S6F). We proposed an asymmetric blocking model of cohesin sliding by insulator-bound CTCF proteins (Fig. 2G). Further experiments are needed to systematically dissect regulatory mechanisms and 3D chromatin architectures of the Pcdhβy clusters.

FOOTNOTES

This work was supported by Grants from MOST (2017YFA0504203, 2018YFC1004504) and the National Natural Science Foundation of China (31630039, 91640118, and 31470820) to Q.W. Q.W. is a Shanghai Subject Chief Scientist.

Yujia Lu, Jia Shou, Zhilian Jia, Yonghu Wu, Jinhuai Li, Ya Guo, and Qiang Wu declare that they have no conflict of interest.

All institutional and national guidelines for the care and use of laboratory animals were followed.

Yujia Lu1, Jia Shou1, Zhilian Jia1, Yonghu Wu1, Jinhuai Li1,2, Ya Guo1,3, Qiang Wu1✉

1 MOE Key Lab of Systems Biomedicine, Center for Comparative Biomedicine, State Key Lab of Oncogenes and Related Genes, Shanghai Cancer Institute, Joint International Research Laboratory of Metabolic & Developmental Sciences, Institute of Systems Biomedicine, Xin Hua Hospital, Shanghai Jiao Tong University, Shanghai 200240, China
2 Present address: Ming Wai Lau Centre for Reparative Medicine, Karolinska Institutet, Hong Kong, China
3 MRC London Institute of Medical Sciences, Imperial College London, London W12 0NN, UK
✉ Correspondence: qiangwu@sjtu.edu.cn, qwu123@gmail.com (Q. Wu)

OPEN ACCESS

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

Canzio D, Nwakeze CL, Horta A, Rajkumar SM, Coffey EL, Duffy EE, Dufﬁe R, Monahan K, O’Keeffe S, Simon MD et al (2019) Antisense IncRNA transcription mediates DNA demethylation to drive stochastic protocadherin α promoter choice. Cell 177 (639–653):e615
Guo Y, Monahan K, Wu H, Gertz J, Varley KE, Li W, Myers RM, Maniatis T, Wu Q (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. Proc Natl Acad Sci USA 109:21081–21086
Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y et al (2015) CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. Cell 162:900–910
Huang HY, Wu Q (2016) CRISPR double cutting through the labyrinthine architecture of 3D genomes. J Genet Genomics 43:273–288
Jain S, Ba Z, Zhang Y, Dai HQ, Alt FW (2018) CTCF-binding elements mediate accessibility of RAG substrates during chromatin scanning. Cell 174(102–116):e114
Jiang Y, Loh YE, Rajaraajan P, Hirayama T, Liao W, Kassim BS, Javidfar B, Hartley BJ, Kleofas L, Park RB et al (2017) The methyltransferase SETDB1 regulates a large neuron-specific chromatin domain. Nat Genet 49:1239–1250
Monahan K, Rudnick ND, Kehayova PD, Paul F, Newberry KM, Myers RM, Maniatis T (2012) Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin-alpha gene expression. Proc Natl Acad Sci USA 109:9125–9130
Mountoufaris G, Canzio D, Nwakeze CL, Chen WV, Maniatis T (2018) Writing, reading, and translating the clustered protocadherin cell surface recognition code for neural circuit assembly. Annu Rev Cell Dev Biol 34:471–493
Ribich S, Tasic B, Maniatis T (2006) Identiﬁcation of long-range regulatory elements in the protocadherin-alpha gene cluster. Proc Natl Acad Sci USA 103:19719–19724
Rowley MJ, Corces VG (2018) Organizational principles of 3D genome architecture. Nat Rev Genet 19:789–800
Shou J, Li J, Liu Y, Wu Q (2018) Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. Mol Cell 71(498–509):e494
Wu Q, Zhang T, Cheng JF, Kim Y, Grimwood J, Schmutz J, Dickson M, Noonan JP, Zhang MQ, Myers RM et al (2001) Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. Genome Res 11:389–404
Yujia Lu et al.

Wu Q, Guo Y, Lu Y, Li J, Wu Y, Jia Z (2019) Tandem directional CTCF sites balance protocadherin promoter usage. bioRxiv. https://doi.org/10.1101/525543

Yin M, Wang J, Wang M, Li X, Zhang M, Wu Q, Wang Y (2017) Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites. Cell Res 27:1365–1377

Yokota S, Hirayama T, Hirano K, Kaneko R, Toyoda S, Kawamura Y, Hirabayashi M, Hirabayashi T, Yagi T (2011) Identification of the cluster control region for the protocadherin-beta genes located beyond the protocadherin-gamma cluster. J Biol Chem 286:31885–31895

Yujia Lu and Jia Shou have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13238-019-00656-y) contains supplementary material, which is available to authorized users.