Genes & Diseases (2024) 11, 26–29

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/en/journals/genes-diseases

RAPID COMMUNICATION

Distinct enhancer-promoter modes determine *Sox2* regulation in mouse pluripotent cells

Sox2 is a key transfer factor for maintaining pluripotency and self-renewal in embryonic stem cells (ESCs),[1] though the mechanism of its transcriptional regulation in ESCs has not been fully addressed. Distal enhancer-promoter interactions are vital for Sox2 transcription activity in mammals. However, how these diverse interactions individually influence Sox2 gene regulation in mouse ESCs remains unclear. Previous studies found that three distal enhancers (termed E1, E2, and E3) interact with *Sox2* promoter[2] (Fig. 1A; Fig. S1A, and Table S1). E1 forms an mESC-specific interaction, while E2 functions in mouse neural progenitor cells (mNPCs) and has not been shown to directly interact with *Sox2* in mESCs. E3 can interact with *Sox2* in mouse ESCs (mESCs), mouse neural stem cells (mNSCs), and mNPCs. Therefore, there are three classes of E-P associations in mESCs: specific, indirect, and common. In mESCs, E1 and E3 display enhancer activity, while E2 shows weak activity as measured by fluorescence reporter assays (Fig. 1B).

Sox2 knockout leads to cell death *in vitro* and early mortality after implantation *in vivo*, hindering functional studies.[3] Additionally, *Sox2* is expressed in multiple cells/tissues (Fig. S1B). To circumvent this problem, three mESCs E-KO cell lines were generated with CRISPR/Cas9 system (Fig. S2A, B); E1-KO and E2-KO cells showed normal clone morphology, like wild-type (WT) mESCs, whereas E3-KO cells showed loose clone morphology (Fig. 1C). Deleting each of these three enhancers results in viable clones with significant *Sox2* down-regulation (Fig. 1D), indicating that E-P association can impact *Sox2* gene expression differently.

To explore the biological effects of individual E-P association on *Sox2* expression, a full transcriptome analysis was performed of each E-KO cell (Fig. S2C–G). The differentially expressed gene (DEG) analysis revealed hundreds of affected genes, compared with WT mESCs (Table S2, and supplementary methods for details). Results of DEGs overlap analysis indicated that large amounts of DEGs commonly existed in two or three E-KO cells (Fig. S10). We also examined the unique effects of each enhancer (Fig. S14A–D). The coefficient within replicates is more than 0.98, varying across different groups (~0.94–0.97; Fig. 1E). Together, they show a mild targeted effect of E-KOs in the overall cellular transcriptional profile. In contrast, *Sox2* conditional knockout profoundly impacts the cellular transcriptional profile and ESC pluripotency[4](Fig. S7A).

E1 is located distally downstream of *Sox2* and specifically interacts with its promoter in mESCs (Fig. 1A). Gene enrichment analysis of altered biological processes and signaling pathways in E1-KO cells demonstrated the main effect on the cell cycle and differentiation process (Fig. S3A, and Table S3), while KEGG pathway analysis of DEGs revealed an impact on cell proliferation-related pathways (Fig. S3B, and Table S4). We next assessed cell cycle analysis by cytometry. GO analysis of up and down DEGs showed that they were enriched in developmental differentiation and metabolic processes, respectively (Fig. S11). Overall, E1-KO resulted in significant G1 phase arrest, compared with WT mESCs ($P < 0.01$, *t*-test) (Fig. S4A–E), suggesting that E1 regulates the transition of G1 to S phase in mESCs. Similar results were not observed with the EdU assay (Fig. S4F, G). We next assessed stem cell self-renewal by colony formation assays, which demonstrated a significantly lower clonogenicity rate of E1-KO cells (Fig. 1F).

To examine the impact of E1-KO on differentiation, we performed a parallel test on nerve cell directional differentiation using E1-KO and WT cells. At early and late differentiation stages, the expression of *Nestin* (a marker of nerve cell differentiation) was significantly lower in E1-KO than in WT mESCs (Fig. 1G), indicating that this association is needed for proper neural differentiation. Immunofluorescence experiments also proved that E1 knockout can significantly affect the efficiency of mESC differentiation

Peer review under responsibility of Chongqing Medical University.

https://doi.org/10.1016/j.gendis.2022.11.019

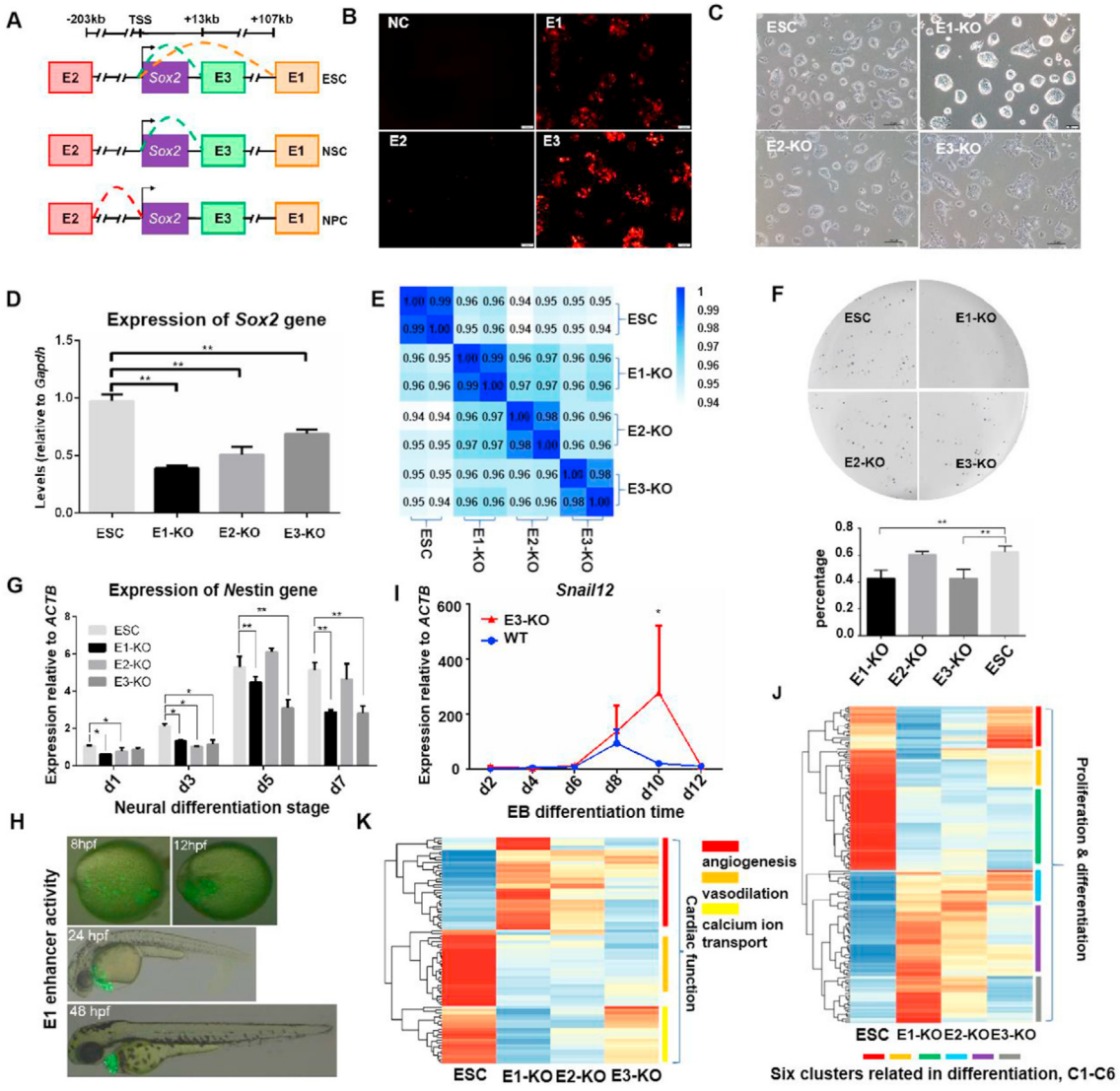

Figure 1 Different patterns of enhancer-promoter (E—P) associations and activities of Sox2 enhancers influenced gene function. **(A)** Three enhancers of the *Sox2* gene show different kinds of E-P associations. The top bar shows markers concerning the transcription start site. Enhancers E1, E2, and E3 are shown in orange, red, and green, respectively. Associations between the enhancers and promoters are indicated by colored dotted lines. **(B)** Enhancer activity assay in mESCs using the pGL4.23-mCherry vector. Red fluorescence indicates positive activity. NC, empty pGL4.23-mCherry vector; E1, E2, E3, vectors containing the corresponding enhancer fragment. **(C)** Cell morphology of E-KO and WT mESCs. **(D)** *Sox2* gene expression in E-KO and wild type (WT) mESCs cells. Values were an average of three biological replicate experiments, with each experiment performed in triplicate. Error bars represent SEM. $^{**}P < 0.01$, *t*-test. **(E)** Pearson correlation coefficients among all samples were calculated based on the values of log 10 (FPKM+1). Different colors represent the squared correlation coefficient (r^2), which ranges from 0.94 to 1. **(F)** The rate of clone formation (the number of clones/inoculated cells) of E1-KO cells decreased significantly compared with WT cells. **(G)** Expression of *Nestin* gene during neuron differentiation of E-KO cells. **(H)** E1 displays enhancer activity around the zebrafish heart at 24 h and 48 h post-fertilization. **(I)** *Snail 12* expression during EB formation of E3-KO cells. **(J)** Gene expression heatmap of genes involved in proliferation and differentiation. **(K)** Gene expression heatmap of genes involved in cardiac function.

into nerve cells (Fig. S12A, B). These indicate that E1 association is required for mESC maintenance by influencing the cell cycle and differentiation.

Several DEGs associated with cardiac function were also identified by GO analysis. These were involved in angiogenesis, cardiac muscle contraction, and vasodilation (Fig. S3A, and Table S3). Pathway analysis also revealed enrichment in cardiomyopathy and cGMP-PKG (Fig. S3B, and Table S4). In order to verify the effect of E1 knockout on myocardial differentiation, we counted the EB beating of E1-KO and wild-type at the late stage of differentiation. The results showed that E1 knockout significantly inhibited myocardial differentiation (Fig. S13A, B). Then we used the zebrafish system to verify the enhancer activity of E1 *in vivo*, E1 showed enhancer activation 24 h post-fertilization in the heart primordium that migrates to the heart at 48 h

post-fertilization (Fig. 1H). No signals were observed in other organs. This indicates that E1 is involved in the early development of the cardiovascular system and the regulation of cardiac function *in vivo*.

E2 is a distal upstream regulatory element that lacks the active histone markers (H3K27ac) and does not directly associate with the *Sox2* promoter in mESCs (Fig. 1A).[2] Despite its weak enhancer activity (Fig. 1B), E2-KO leads to significant changes in the cellular transcriptional profile, and significant *Sox2* down-regulation (Fig. 1D).

GO and KEGG analyses of DEGs from E2-KO demonstrated that metabolism, cytoskeleton, and nervous system functions were the main affected processes, but other metabolic processes were also involved (Fig. S3C, D and Table S3, 4). Additionally, it showed enhancer activity in the nervous system, suggesting a potential regulatory role in nervous system development, neuron apoptosis, and axon regeneration (Table S3). Hence, we conducted a directional nerve cell differentiation assay and demonstrated that *Nestin* is significantly down-regulated on the third day of neural differentiation (Fig. 1G). This suggests that the nerve cell differentiation efficiency is influenced by E2 at the early differentiation stage. DEGs in E2-KO were enriched in nervous system development or neuron differentiation represented by the down-regulation of *Mbp*, *Acvr1b*, *Vcan*, Otx2, *Rb1*, *Trappc9*, *S1pr1*, and *Notch 3* genes (Fig. S5).

E3 interacts with the *Sox2* promoter in mESCs, mNSCs, and mNPCs,[2] and was therefore classified as a common E-P association. GO analysis of DEGs from E3-KO cells indicated that E3 influences ion transport and homeostasis, and amino acid metabolism, similar to E1 and E2 (Fig. S3E and Table S3). This implies they could coordinate *Sox2* function; however, pathway analysis also showed an E3-specific regulatory role in glycolysis/gluconeogenesis and pyruvate metabolism (Fig. S3F). E3 DEGs were significantly enriched in signaling pathways positively regulating the pluripotency of stem cells (Fig. S6 and Table S5), confirmed by the relatively loose clone morphology of E3-KO cells (Fig. 1C).[5] Additionally, the embryoid body differentiation assay demonstrated that the late mesodermal marker (snail 12) and two endoderm marker genes (gata 4, gata 6) are significantly up-regulated in E3-KO cells during the middle and late differentiation stages (Fig. 1I; Fig. S5D).

Sox2 is the key node of the core transcription regulatory network for pluripotent maintenance in mESCs.[2] As such, there is a potential relationship between changes in the transcriptome induced by E-KO and the *Sox2* regulatory network. To explore this relationship, the ChIP-seq of Sox2[4] was integrated with DEGs from E-KO cells for analysis (Fig. S7B). Enrichment analysis found that 15.33%—18.53% of the DEGs in each E-KO cell were Sox2 target genes (Table S2); and DEGs of each E-KO line were enriched in networks associated with pluripotency (plurinetwork, wp1763) including double KO cells (Fig. S7A). These results indicate that down-regulated *Sox2* is involved in the transcriptome fluctuation of E-KO mESCs.

To explore potential synergistic effects, we generated E1 and E3 double KO (E1&E3-dKO) in mESCs, which resulted in even looser clone morphology, and further *Sox2* gene down-regulation (Fig. S8A, B). DEGs analysis between single and double KO revealed 253 DEGs (~32%) only present in E1&E3-dKO cells, many directly involved in Sox2 regulation (Fig. S8C, E); and GO and pathway analyses showed enrichment in cell motility (regulation of cell adhesion, microtubule-based movement), pluripotency or cell differentiation (Notch signaling) (Fig. S8D).

To understand the overlap of these three distal enhancers (E1, E2, and E3) on their impact on different biological processes, we next performed gene expression heatmap cluster analysis (Fig. 1J, K; Fig. S9 and Table S6). Clustering based on cell proliferation and differentiation genes revealed 6 clusters (C1—C6). In E1-KO cells, C1—C3 contained down-regulated DEGs, while C4—C6 contained up-regulated DEGs, compared with WT mESCs. E2-KO cells showed a similar profile to E1, but were driven by C4; E3-KO cells showed up-regulation of C1 (Fig. 1J and Table S7). Clustering analysis of cardiac function-related genes shows specific subdivisions of biological processes (Table S8). Collectively, up-regulated DEGs primarily promote angiogenesis, while down-regulated DEGs inhibit calcium ion transport and vasodilation (Fig. 1K). Finally, clustering analysis of metabolism-related genes demonstrated that DEGs were divided into three categories (Fig. S9 and Table S9). For E3-KO cells, up-regulated DEGs were enriched in lipid metabolic and oxidation—reduction processes, while down-regulated DEGs were enriched in the cholesterol metabolic process, fatty acid metabolic process, and oxidation-reduction process. Overall, E3-KO cells showed a more diverse transcriptional profile than E1- and E2-KO clones.

In conclusion, complicated E-P associations can cooperatively regulate *Sox2* function in various patterns. This study offers an alternative strategy to explore spatiotemporal gene expression and transcription control. Future studies should seek to obtain a detailed understanding of E-P associations, which will be required to comprehensively understand gene functions.

Conflict of interests

The authors delcare that they have no competing interests.

Funding

This work was supported by the National Key Research and Development Program of China (No. 2018YFA0903201); the National Natural Science Foundation of China (No. 32202653 and 31970592); the Agricultural Science and Technology Innovation Program (33-AGIS-07); the China Postdoctoral Science Foundation (No. BX2021367 and 2021M703543); the Guangdong Basic and Applied Basic Research Foundation (No. 2022A1515010766); the Shenzhen Science and Technology Program (No. KCXFZ2020122-1173205015 and RCBS20210609104512021). It is also sup-

ported by the Thousand Talents Plan for Young Professionals [to Y.Z.].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gendis.2022.11.019.

References

1. Masui S, Nakatake Y, Toyooka Y, et al. Pluripotency governed by Sox 2 via regulation of Oct 3/4 expression in mouse embryonic stem cells. *Nat Cell Biol*. 2007;9(6):625—635.
2. Zhang Y, Wong CH, Birnbaum RY, et al. Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature*. 2013;504(7479):306—310.
3. Avilion AA, Nicolis SK, Pevny LH, et al. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev*. 2003;17(1):126—140.
4. Adachi K, Nikaido I, Ohta H, et al. Context-dependent wiring of Sox 2 regulatory networks for self-renewal of embryonic and trophoblast stem cells. *Mol Cell*. 2013;52(3):380—392.
5. Ying QL, Wray J, Nichols J, et al. The ground state of embryonic stem cell self-renewal. *Nature*. 2008;453(7194):519—523.

Lei Huang [a,1], Xiusheng Zhu [a,1], Qing Li [a,1], Dashuai Kong [a,d], Qitong Huang [a,b], Jing Luo [a], Siyuan Kong [a], Yanling Peng [a], Yubo Zhang [a,c,*]

[a] *Animal Functional Genomics Group, Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Key Laboratory of Livestock and Poultry Multi-omics of MAPA, Shenzhen Agricultural Genome Research Institute, Chinese Academy of Agricultural Sciences, Shenzhen, Guangdong 518120, China*
[b] *Animal Breeding and Genomics, Wageningen University & Research, Wageningen 6708PB, Netherlands*
[c] *College of Life Science and Engineering, Foshan University, Foshan, Guangdong 528231, China*
[d] *School of Life Sciences, Henan University, Kaifeng, Henan 475004, China*

*Corresponding author. Animal Functional Genomics Group, Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Key Laboratory of Livestock and Poultry Multi-omics of MARA, Shenzhen Agricultural Genome Research Institute, Chinese Academy of Agricultural Sciences, Shenzhen, Guangdong 518120, China.
E-mail address: ribon_001@163.com (Y. Zhang)

25 April 2022
Available online 28 December 2022

[1] These authors contributed equally to this work.