Induce Pluripotency via Specific distal enhancer-promoter associations

Xiusheng Zhu^{1†}, Lei Huang^{1†}, Dongwei Li^{1†}, Jing Luo¹, Qitong Huang¹, Xueyan Wang¹ and Yubo Zhang^{1*}

¹Animal Functional Genomics Group, Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China

*Corresponding author Email: ribon_001@163.com.

†These authors contributed equally to this work.

Abstract

Induced pluripotent stem cell (iPSC) technology promises to be an inexhaustible source of any type of cell needed for therapeutic and research purposes. It is unclear that how distal enhancer-promoter associations/ 3D chromatin conformation involving in the capacity of self-renewal and pluripotency maintenance. In this study, we have selected a few defined enhancer-promoter associations. After screening of enhancer specificity and activity individually, we design the different combinations and transfect these enhancers into the MEF cells. We simultaneously transfect 7 determined enhancers which represents various specific distal chromatin associations into a GFP tracing MEF cell line. We observe that the MEF cells start generating iPS-like clones at day 22. Importantly, our validations with three germ layer marker genes and *in vitro* experiments have further confirmed the pluripotency of these clones. Here, our study proposes a potential *de novo* method of a low-genetic risk iPS generation by introducing spatiotemporal distal chromatin associations. This result also paves out the way on utilizing 3D genomic information to alter cell identity and reprogramming for potential therapeutic strategy.

Introduction

Induced Pluripotent Stem Cells (iPS) is a promising technology and derived from adult cells that have been reprogrammed back into an embryonic-like pluripotent state and it enables the development of an unlimited source of any type of cell needed for therapeutic purposes and other applications. It opens up unprecedented opportunities in regenerative medicine, disease modelling and drug discovery. Therefore, researchers have rapidly improved the techniques to generate iPSCs and lots of different methods have been developed since then. There are still a few shortcomings to limit its application, such as production efficiency, safety, carcinogenicity and etc¹⁻⁸. Among them, one big concern is the permanent effects by introducing the endogenous gene/TFs would disturb the transcriptional factory in the cells. This might have some unpredicted consequences and inscrutable effects in the long run or next generation individual/cells⁹.

Three-dimensional (3D) genomics is emerging as a powerful tool to study the 3D structure and transcriptional regulation of eukaryotic genomes during the last decade¹⁰. Scientists have started realizing the interplay between transcription and genome conformation is a driving force for cell-fate decisions¹¹. Through 3D genomics, we can bring the role of genome conformation in transcriptional regulation to the fore and understand the basic structure of DNA, the growth and development of organisms and the occurrence of diseases. However, at present, the researches still remain on describing the physical interactions/associations within various cell/organs. It is a major challenge to set up the solid connection between defined 3D genomic information and cell-fate decisions.

In our previous study, we have identified the physical transcriptional network existing in mouse embryonic stem cells (mESCs)¹². The similar result is also described in mouse neural stem cells (mNSCs). So we recognize that the OSKM or Yamanaka factors (Oct4, Sox2, Klf4, and cMyc) are all connected via different chromatin associations. We find these enhancer-promoter associations show strong specificity comparing with other cells which is alike in OSKM's gene expression patterns. Additionally, Nanog has also been reported to be essential for the

establishment of both embryonic stem cells (ESCs) from blastocysts and iPSCs from somatic cells¹³, which is also involved in the network via specific interactions. In this study, we have selected these defined mES specific associations to explore the relationship between cell-fate decision and these specific chromatin associations. And we choose the classical iPSCs technical mode as our research roadmap.

Results

Activation of Endogenous Pluripotent Gene

Oct4, *Sox2*, *Klf4* and *Nanog* were selected as targets because of their central roles in pluripotency maintenance and induction. In previous study, we have identified the specific chromatin associations of them¹². Further, we verify their activities by considering other factors such as H3K27ac and H3K4me1 modification in pluripotent stem cells, and their potential to form promoter-enhancer loops mediated by Mediator complex¹⁴⁻¹⁶ (Fig.1A). Finally, we choose 9 candidate enhancers, namely OE1 (chr17:35622426-35623338), OE2 (chr17:35639453-35641591) (*Oct4* enhancers); KE1 (chr4:55527598-55528553), KE2 (chr4:55663709-55664620) (*Klf4* enhancer); SE1 (chr3:34656590-34657598), SE2 (chr3:34342406-34344100), SE3 (chr3:34562477-34563357) (*Sox2* enhancer); NE1 (chr6:122652070-122654308), NE2 (chr6:122619081-122620092) (*Nanog* enhancers) as candidates for the following experiments.

First, in order to test their transcriptional activities, we individually transfect these enhancers into the MEFs through a retroviral system¹⁷. We apply a transgenic *Oct4* promoter driving GFP expression MEFs, OG2 MEF cells, as our model system. After 96 hours, we use qPCR to detect the expression level changes of these endogenous target genes. As expected, most of them can activate/enhance the expression of endogenous target genes at various levels (Fig. 1B). To reduce virus damage to cells and obtain higher activation efficiency, we further test the gene expression level of different enhancer combinations of individual pluripotent gene. The results show that the two enhancer combinations of *Oct4* gene have significantly higher transcriptional activation levels than individual enhancer. For *Sox2* gene, the

combined activation of transcription levels of three enhancers is significantly higher than that of a single enhancer and SE1&SE2, SE2&SE3 combination. But there is no significant difference with the SE1&SE3. For the *Klf4* gene, the transcriptional activation level of the two enhancer combinations for the target gene is significantly higher than that of KE6, but there is no significant difference to KE5. The combination of two enhancers of *Nanog* gene has significantly higher transcriptional activation levels for target genes than individual one (Fig. 1C). Therefore, we choose OE1, OE2, KE1, SE1, SE3, NE1, NE2 as the candidates and use them to induce MEFs and to determine the relationship between the pluripotency and chromatin associations.

To double check its activity of these 7 chose enhancers, we construct them on a luciferase vector which has been widely used as a golden standard for this kind of validation experiments. The results verify the activity, and their expression levels are all significantly higher than the control group (Fig. 1D). These results are consistent with previous ones in zebrafish *in vivo*(ref). These indicate that exogenous enhancers can enhancer/activate transcription of the endogenous target genes.

Induce Pluripotency via Specific distal interactions

We next try to determine whether pluripotency can be fully reactivated and established in MEFs following classical iPSs procedure. As mentioned before, in order to detect the phenomena more conveniently, we use OG2 MEF cells for the whole process. With OG2 MEF cells, when the endogenous *Oct4* is activated, the cells show a strong EGFP signal and can easily capture the signal via microscopes¹⁸.

To determine whether these specific distal interactions are sufficient to induce pluripotency with OG2 MEFs, we follow the efficient retroviral transfection iPSs protocol¹⁷ to integrate these enhancers into the MEFs. 48 hours after transfection, the MEFs medium has been replaced with iCD1 medium, which is recorded as day 0. Then we daily track and observe the morphological changes of these cells. In the study, reprogramming clusters start appearing on the 11th day (Fig. 2A). On 18th day, the emergence of clones can be observed (Fig. 2B). In order to determine more details

regarding to the changes of Oct4 and Sox2 gene during the reprogramming process, we monitor their expression levels at different time points. The results showed that Oct4 gene expression has been constantly increasing during the reprogramming process. But, for Sox2 gene, it is lower on the 18th day than on the 11th day (Fig. 2C), which indicates the expression pattern of them is a little different during the process. The number of clones increased significantly on the 22nd day. To determine its pluripotency, we determine the signals of the EGFP expression under a fluorescence microscope, and the results present that they are all EGFP-positive ones (Fig. 2D). These data indicates that the pluripotency has been induced.

Next, in order to further confirm their differentiation ability *in vitro*, we pick some clones and culture them in suspension in a petri dish for 7 days. After the suspension culture process, these clones can form embryoid bodies (EBs) (Fig. 3A). In order to test whether they can differentiate into three primitive layers (endoderm, ectoderm, mesoderm), We continue to adhere to these EBs. In standard protocols, immunofluorescence experiments can be used to detect the specific expression pattern of three germ layers. Here, marker genes *SOX17* (endoderm), *β -Tubulin* (ectoderm), and *SMA* (mesoderm) are applied to determine the ability (Fig. 3B). In the results, it confirms that these EBs show specific expression pattern and have the potential to differentiate into other types of cells. Together, these results indicate that these specific distal interactions can induce the MEFs into a pluripotent status and form the iPSs like clones. Further, these clones can differentiate into three primitive layers and have the potential to differentiate into various types of cells *in vitro*.

Discussion

For putative regenerative applications, safety is always the foremost consideration for scientist. In traditional iPSs protocol, expression of OSKM factors is a key risk for sustained pluripotency, uncontrolled proliferation and tumorigenesis. It is known that TFs are coding genes and usually common to several cell types. They play a general role in gene regulation to determine the phenotypic characteristics of a cell¹¹. This would bring unpredicted consequences and inscrutable effects and could

limit its application. The distal enhancer-promoter chromatin associations present robust spatiotemporal specific in various studies^{12,19,20}. The majority of enhancers are non-coding elements. Thus, manipulating them without disturb the gene/TF may dramatically reduce the risk and increase the safety. In this study, we choose the defined mESCs specific interactions of OSKM and follow the standard iPSs protocol. In the study, these specific interactions activate/enhance endogenous gene expression and result in iPS-like cells. Our method will lower the risk for iPSs application and provide an alternative route for iPSs generation.

3D chromatin architecture has been recognized in influencing gene regulation, cell fate decisions and evolution in the last decade. Also, it is known that it orchestrates transcriptional regulation act in a complex, multilayered and interconnected fashion. Different non-random spatial chromatin organization has been described based on various methodologies. However, all these studies are descriptive and hypothetical. Many of them imply or suspect a role for 3D genome conformation in the transcription factor-driven control of gene regulation and, consequently, cell-fate decisions. ChIA-PET is a unique technology which can bring along the function information of interested TFs with distal chromatin associations. In our previous study, we have identified physical connected transcription networks in mESCs and mNPCs. In current study, we apply these specific interactions to induce the iPSs following the standard protocol. Therefore, we have built a solid connection between 3D chromatin conformation and cell-fate decision here. We realize that cell identity can be determined by the spatiotemporal distal chromatin associations and a spatially organized genome. If that was the case, our method could propose some alternative therapeutic strategies by altering the cell-fate decision in cancer or other deadly diseases or ESCs/iPSs free route for regeneration therapies.

Methods

Cell culture

Plat-E cells purchased from ATCC, OG2 MEFs were from Dr. Baoming Qin's laboratory, Plat E cells and MEFs were cultured in DMEM supplemented with 10% FBS and non-essential amino acid (NEAA).

All iPS-like cells and mouse ES cells were cultured in N2B27-2i/LIF medium (Neural Basal, F12, 1% Glutamax, 1% NEAA, 1% Sodium pyruvate, 0.1mM β -me, 1%B27, 0.5%N2, 0.5% P/s(optional), mLif(1:8000), 3uM CHIR99021(1:4000), 1 uM PD0325901), MEFs are induced in iCD1 medium (Delicell, 820250).

Plasmid Construction

Pmx-mP-mCherry based on the modified pMX-GFP retroviral vector (Cell Biolabs, USA), is used in the study. Mini-promoter and mCherry (miniP-mCherry) fragments were obtained from modified pGL4.23 luciferase reporter vector (Promega, USA). Briefly, PGL4.23 was digested with XbaI/NcoI restriction enzymes to remove the original coding sequences of luc2 reporter gene and replace it with mCherry. The miniP-mCherry fragment was then amplified with primers containing NcoI or NotI restriction enzymes. Synthetic DNA fragment was cloned into NcoI and NotI digested pMX-GFP vector. The resulting screening vector was named pmx-mchery. Then double chop with Hind III and xhoI to get a linearized vector that can be attached to the input library.

Retroviral Preparation and Transduction

Prepare Plat-E cells one day in advance, digest them with 0.25% pancreatin, and count them. Inoculate about 8×10^6 /100mm plate of cells, wait until the cell density is about 80%-90%, transfect with lipo3000, change 10ml DMEM 10% FBS medium after 12 hours, and perform MEFs cell plating after 24 hours. The plating density was 5×10^4 /well. After 12 hours, the supernatant of Plat-E cells was collected and filtered into a centrifuge tube with a 0.45 μ M filter membrane. After adding polybrene with a final concentration of 8 μ g/ml, the MEFs medium was replaced with fresh medium.

Add the virus supernatant, add 1 ml of each enhancer virus supernatant (if it is a concentrated virus solution, reduce the corresponding multiple), Pmx-mCherry virus infection MEFs as a control group. A second virus transfection was carried out after 24 hours, and the medium was changed to iCD1 mouse induction medium after another 24 hours.

Gene expression analysis

The cells were digested with 0.25% trypsin, and washed twice with PBS to collect the cells. The total RNA of the cells was extracted with the kit (Vazyme, RC112). Each reaction was reversed with 1 μ g total RNA, and SYBR Green SuperMix (Vazyme, R323) on the 7500 Fast Real-Time PCR System (Applied Bio-systems). The pluripotent gene primer sequence is as follows: mOct4-F: ATGAGGCTACAGGGACACCTT, mOct4-R: GTGAAGTGGGGGCTTCCATA; mKlf-F: GTGCCCCGACTAACCGTTG, mKlf-R: GTCGTTGAACTCCTCGGTCT; mSox2-F: GCGGAGTGGAACTTTTGTCC, mSox2-R: CGGGAAGCGTGTACTTA TCCTT; mNanog-F: AGGACAGGTTTCAGAAGCAGA, mNanog-R: CCATTGCTAG TCTTCAACCACTG; mGAPDH-F: AACTTTGGCATTGTGGAAGG, mGAPDH-R: A CACATTGGGGGTAGGAACA.

In vitro differentiation experiment

Suspension culture of mouse ips-like cells in N2B27-2i/LIF medium for 7 days, and then continue to adhere to culture to differentiate into cells of different germ layers.

Immunofluorescent Staining

Remove the cell culture medium, wash 3 times with PBS, and fix with 4% paraformaldehyde for 30 min at room temperature; wash 3 times with blocking solution (PBS containing 100 mmol/L glycine and 0.3% BSA), then add 1% Tritonx-100 Permeabilization at room temperature for 15 minutes; then add 5% bovine serum albumin (BSA) and block for 2 hours at room temperature; wash 3 times with TBP

(Tritonx-BSA-PBS), add primary antibody and incubate overnight at 4 °C; rewarm at room temperature the next day 20 min, wash with TBP 3 times, add secondary antibody and incubate for 1 h at room temperature; wash 3 times with TBP, then observe under a fluorescence microscope.

Enhancer activity verification

The Pmx-mP-mCherry plasmid was double digested with endonucleases Hind III and XhoI, and the enhancer sequence to be verified was ligated to the vector by homologous recombination(Vazyme, Nanjing, China). Transfect the constructed plasmid into the cells(Invitrogen, Cat. No. L3000075), and observe the expression under a fluorescent microscope 24 hours later.

Competing interests There is no conflict of interest.

Acknowledgments Thanks to the cell line provided by Baoming Qin Laboratory, Guangzhou Institute of Health, Chinese Academy of Sciences. Thanks to Dr. Ben Huang's laboratory of Guangxi University for their help in the in vitro differentiation experiment. Thanks to Fei Fang for his help in cell culture experiments. Dr. Jun Yin reviews out manuscript.

Funding National Natural Science Foundation of China [31970592];National Key Research and Development Program of China [2018YFA0903201]; The Agricultural Science and Technology Innovation Program; The Elite Young Scientists Program of Chinese Academy of Agricultural Sciences [CAASQNYC-KYYJ-41]; National Natural Science Foundation of China (No. 32002173); Natural Science foundation of Guangdong Province, China [2018A0303130009].

Author contributions Y.Z., X.Z. and L.H. conduct and design the experiments. X.Z., D.L performed the experiments. X.W. help X.Z performed the experiments. Q.H, and J.L is responsible for photo editing and typesetting. The manuscript is written by X.Z. and Y.Z. All authors read and approved the final manuscript.

References

- 1 Kaji, K. *et al.* Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* **458**, 771-775, doi:10.1038/nature07864 (2009).
- 2 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
- 3 Miyoshi, N. *et al.* Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell stem cell* **8**, 633-638, doi:10.1016/j.stem.2011.05.001 (2011).
- 4 Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell stem cell* **7**, 618-630, doi:10.1016/j.stem.2010.08.012 (2010).
- 5 Seki, T. *et al.* Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell stem cell* **7**, 11-14, doi:10.1016/j.stem.2010.06.003 (2010).
- 6 Jia, F. *et al.* A nonviral minicircle vector for deriving human iPS cells. *Nature methods* **7**, 197-199, doi:10.1038/nmeth.1426 (2010).
- 7 Yu, J. *et al.* Human induced pluripotent stem cells free of vector and transgene sequences. *Science* **324**, 797-801, doi:10.1126/science.1172482 (2009).
- 8 Woltjen, K. *et al.* piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458**, 766-770, doi:10.1038/nature07863 (2009).
- 9 Ma, T., Xie, M., Laurent, T. & Ding, S. Progress in the reprogramming of somatic cells. *Circ Res* **112**, 562-574, doi:10.1161/CIRCRESAHA.111.249235 (2013).
- 10 Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nature reviews. Molecular cell biology* **20**, 535-550, doi:10.1038/s41580-019-0132-4 (2019).
- 11 Stadhouders, R., Filion, G. J. & Graf, T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345-354, doi:10.1038/s41586-019-1182-7 (2019).
- 12 Zhang, Y. *et al.* Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature* **504**, 306-310, doi:10.1038/nature12716 (2013).
- 13 Schwarz, B. A., Bar-Nur, O., Silva, J. C. & Hochedlinger, K. Nanog is dispensable for the generation of induced pluripotent stem cells. *Curr Biol* **24**, 347-350, doi:10.1016/j.cub.2013.12.050 (2014).
- 14 Sethi, A. *et al.* Supervised enhancer prediction with epigenetic pattern recognition and targeted validation. *Nature methods* **17**, 807-814, doi:10.1038/s41592-020-0907-8 (2020).
- 15 Wang, S. P. *et al.* A UTX-MLL4-p300 Transcriptional Regulatory Network Coordinately Shapes Active Enhancer Landscapes for Eliciting Transcription. *Molecular cell* **67**, 308-321 e306, doi:10.1016/j.molcel.2017.06.028 (2017).
- 16 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).
- 17 Chen, J. *et al.* Rational optimization of reprogramming culture conditions for the generation of induced pluripotent stem cells with ultra-high efficiency and fast kinetics. *Cell research* **21**, 884-894, doi:10.1038/cr.2011.51 (2011).
- 18 Szabo, P. E., Hubner, K., Scholer, H. & Mann, J. R. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech Dev* **115**, 157-160,

- doi:10.1016/s0925-4773(02)00087-4 (2002).
- 19 Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84-98, doi:10.1016/j.cell.2011.12.014 (2012).
- 20 Kubo, N. *et al.* Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nature structural & molecular biology* **28**, 152-161, doi:10.1038/s41594-020-00539-5 (2021).

Figure legends:

Figure.1 Gene-specific interaction enhancers activate endogenous gene expression

A: Describes the specific interaction between Oct4, Klf4, Sox2, Nanog enhancers and target genes, as well as the overlap of histone modifications, P300, and Mediator complex.

B: The activation of exogenous specific enhancers on the transcription level of pluripotent genes. *: $p < 0.05$; **: $p < 0.01$.

C. The effect of combined activation of different enhancer combinations on the transcription level of target genes. *: $p < 0.05$; **: $p < 0.01$, ns: There is no significant difference.

Figure 2. The specific interaction enhancer induces pluripotency network in MEFs.

A. The MEFs induced on the 11th day showed cell aggregation.

B. Clone-like cells appeared on the 18th day.

C. The expression trend of the pluripotent genes Oct4 and Sox2 at different time points.

D. The expression level of EGFP in clone-like cells observed by fluorescence microscope.

Figure 3. In vitro differentiation experiment of iPS-like cells.

A. a. EBs formed by cloning-like cells in suspension culture for 7 days. b. EBs began to differentiate gradually to form many types of cells.

B. Immunofluorescence experiments observed that EBs can differentiate spontaneously and express the marker genes of the three germ layers.

A **B** **C** **D**