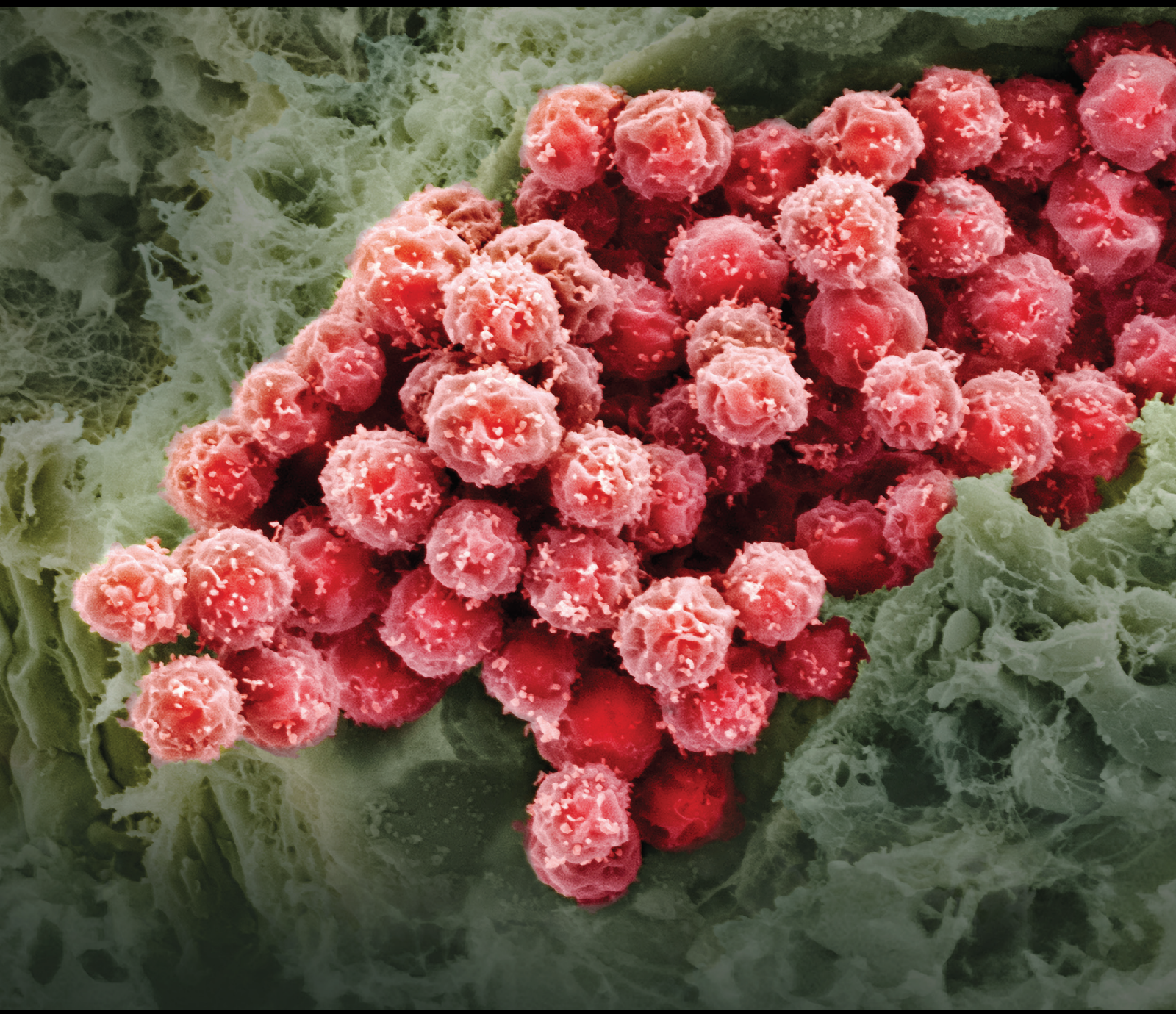
Stem Cells International

Chromatin Regulation in Stem Cell Differentiation

Lead Guest Editor: Qiang Wu
Guest Editors: Wensheng Zhang, Yuin-Han Loh, and Ernesto Guccione



Chromatin Regulation in Stem Cell Differentiation

Stem Cells International

Chromatin Regulation in Stem Cell Differentiation

Lead Guest Editor: Qiang Wu
Guest Editors: Wensheng Zhang, Yuin-Han Loh, and Ernesto Guccione

Contents

Hindawi
Stem Cells International
Volume 2021, Article ID 6660936, 8 pages
https://doi.org/10.1155/2021/6660936

Review Article

The Regulation and Functions of Endogenous Retrovirus in Embryo Development and Stem Cell Differentiation

Yangquan Xiang and Hongqing Liang

Division of Human Reproduction and Developmental Genetics, Women's Hospital, and Institute of Genetics, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Correspondence should be addressed to Hongqing Liang; lianghongqing@zju.edu.cn

Received 29 October 2020; Accepted 19 February 2021; Published 28 February 2021

Academic Editor: Qiang Wu

Endogenous retroviruses (ERVs) are repetitive sequences in the genome, belonging to the retrotransposon family. During the course of life, ERVs are associated with multiple aspects of chromatin and transcriptional regulation in development and pathological conditions. In mammalian embryos, ERVs are extensively activated in early embryo development, but with a highly restricted spatial-temporal pattern; and they are drastically silenced during differentiation with exceptions in extraembryonic tissue and germlines. The dynamic activation pattern of ERVs raises questions about how ERVs are regulated in the life cycle and whether they are functionally important to cell fate decision during early embryo and somatic cell development. Therefore, in this review, we focus on the pieces of evidence demonstrating regulations and functions of ERVs during stem cell differentiation, which suggests that ERV activation is not a passive result of cell fate transition but the active epigenetic and transcriptional regulation during mammalian development and stem cell differentiation.

1. Introduction

ERVs belong to a Class family of retrotransposon elements in the genome. Together with DNA transposons, they are known as transposable elements (TEs), which are derived from DNA fragments able to transpose within the genome. Due to their capacities to hop around and copy themselves in the genome, TEs are considered one of the main driving forces in reconstructing the genome during mammalian evolution. To date, TEs have mostly lost the ability to transpose [1, 2], considering that the transposition events might lead to genome instability. ERVs and other family members of TEs used to be considered as "junk DNA," but with the technological advancement in genome-wide expression and epigenetic profiling, we started to appreciate more on their functional contribution to development and diseases. We now understand that the complexity of the mammalian genome is not achieved through a significant increase of the protein-coding sequences, but by the vast expansion of regulatory capacities imparted by the non-coding sequences. TEs occupy nearly half of the non-coding genome and thus are thought to play critical roles in shaping the complexity of mammalian gene regulatory network.

Comparing to other repeat element families, such as short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs), ERVs bear more sequence complexities and thus may play more specific regulatory functions in the genome [1, 3]. Although ERVs are the smallest class of retrotransposon family, they exhibit significant enrichment and are over-represented in cell type-specific active regulatory sequences [4]. ERVs are thought to be generated as by-products of retroviral infection and integration events in the ancestral mammalian genome. During the evolution, they were endogenized and inherited through germline transmission [5]. Most ERVs are tamed now in the host genome through mutations of their transposition

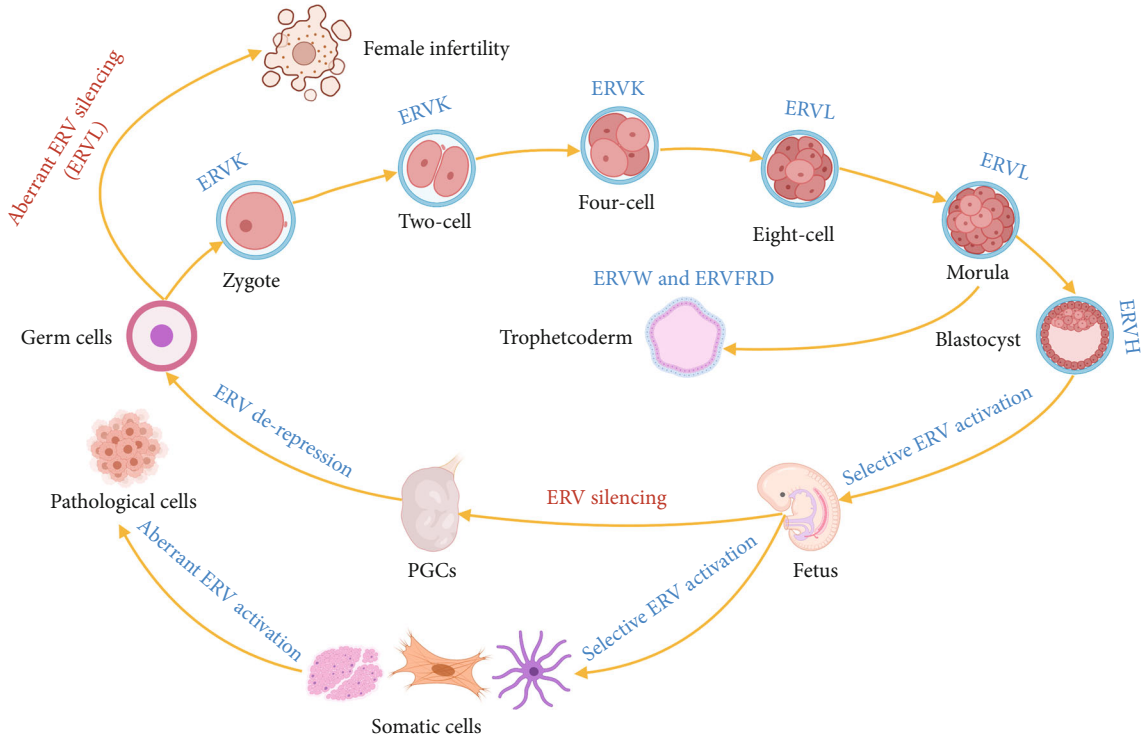


FIGURE 1: The dynamic regulation ERVs during development. During embryonic and somatic development, ERVs are selectively activated, whereas aberrant activation or silencing of ERVs results in pathological consequences.

machinery or through coevolution of host regulatory factors that repress ERV activation [5]. A full-length ERV consists of two long-terminal repeats (LTRs) flanking at both 5′ and 3′ sides, and the open reading frame (*GAG*, *POL*, and *ENV*) in the center. It should be emphasized that LTRs are the regulatory elements of ERVs [6]. The LTR regions of ERVs possess binding sites for a broad scope of transcription factors to interact with the host gene regulatory machinery and achieve precise control of ERV activity [7]. Meanwhile, exaptation of the LTRs' cis-regulatory functions (enhancer and promoter) also leads to innovations of the transcription network in the host genome. ERVs also exclusively possess the primer binding site (PBS) which can recruit complementary tRNA to prime for viral reverse transcription. PBS sequences are also found to be the binding sites for ERV silencing factors from the host [7]. Based on the similarity to tRNA sequence in the PBS region, ERVs can be further classified into several families, ERVH, ERVW, ERVK, ERVL, etc. Out of the 8% genomic constitution of ERVs in human genome, 90% exist as solitary ERVs with only the LTR sequences present and the viral protein-coding ERV-int regions shed off [3].

The expression level of ERVs is dynamically regulated in early embryogenesis, differentiated tissues, and germ cells [8]. Interestingly, the expression of different ERV subfamilies exhibited high temporal specificity during early human embryo development [8], suggesting ERVs as stringent markers for specific embryonic stages (Figure 1). Besides, many shreds of evidence also demonstrated that abnormal ERV expression may lead to different types of diseases [9–13]. ERVs can affect genome-wide transcription through multiple layers of regulation as discussed below. Thus, their activities should be tightly controlled in the mammalian genome to coordinate with proper development and cell fate decision process. The precise control of ERVs in the host genome is largely through transcriptional and epigenetic regulation. DNA methylation is considered a common regulatory mechanism to repress ERV expression. Many ERVs in human are heavily methylated and silenced in differentiated tissues but show loss of methylation and aberrant expression in cancer [14]. Apart from DNA methylation, the Krüppel-associated box domain-containing zinc finger protein (KRAB-ZFP) is known to regulate chromatin configuration surrounding ERV elements [15]. ERV elements are bound by zinc finger domains of the KRAB-ZFPs, and the KRAB domain can recruit tripartite motif-containing 28 (TRIM28), resulting in the trimethylation of histone H3 lysine9 (H3K9me3) and ERV silencing in embryonic stem cells [16]. Histone deacetylation is also involved in ERV regulation. It has been found that histone deacetylase inhibitor (HDACi) treatment led to ERV9 activation which prevented testicular cancer progression, but this did not lead to upregulation of other ERV sub-families, implying that histone deacetylation may regulate human ERV silencing in a subfamily-specific manner [15, 17]. In general, it can be envisaged that a combination of different kinds of epigenetic modifications is orchestrated to tightly control the ERV activity.

Over the last 10 years, increasing pieces of evidence are showing that LTRs may play under-recognized regulatory roles in mammalian development and diseases [9–13]. In the following sessions, we will discuss in detail about the current knowledge on the functions of ERV in chromatin and transcription regulation, how these functions are achieved,

and how they contribute to cell fate decision during mammalian embryonic development and stem cell differentiation.

2. The Functions of ERV in Gene Regulation

If chromatin regulation is a symphony, then ERV has several instruments to play. ERV recruits transcription factors, works as alternative promoters, encodes long non-coding RNAs (lncRNAs), and produces protein products to mediate cellular function. These abilities could be stemmed from intrinsic functions of ERV or could be coopted during the coevolution with the host genome. Nevertheless, the functions of ERV have become an integral part of the regulatory machinery in the genome and indispensable for the normal development and homeostasis of mammals.

2.1. The Recruitment of Transcription Factors. In-silico mapping revealed that many ERVs are enriched with transcription factor-binding sites, suggesting ERVs may act as cis-regulatory elements for transcription [4]. Putative epigenetic markers for promoter and enhancer, such as H3K4me3 and H3K27ac, are frequently seen on the LTR regions [11]. Activated ERVs are largely associated with cell type-specific open chromatin configuration. For example, in human pluripotent stem cell, HERVH sub-family is enriched with binding sites for pluripotency transcription factors such as OCT4 and KLF4, as well as active histone modifications like H3K4me3 and H3K27ac, adopting open chromatin conformation [11, 18]. In addition, DUX4, as well as its mouse homologous DUX, can bind to the ERVL sub-family in human and mouse, respectively. This leads to epigenetic activation of genes downstream of the ERVL elements, which are essential for initiating zygote genome activation (ZGA) in early human and mouse embryos [19]. Human DUX4 is kept silenced in differentiated tissues, as aberrant activation of DUX4 in muscle tissue upregulates HERVL, leading to unscheduled transcription activation of early embryonic genes which eventually resulted in facioscapulohumeral muscular dystrophy [10]. These shreds of evidence together suggested that ERVs can recruit transcription factors to actively influence the epigenetic landscape in the nearby region, thus contributing to cell type-specific gene regulation.

Moreover, ERVs can also modulate signaling pathways to coordinate cell fate change. It has been found that ERVs shaped the evolution of the transcription network underlying the interferon response [20]. For instance, one of the ERV sub-families, MER41, is enriched with interferon-induced STAT1-binding sites [20]. STAT1-bound MER41 regions were enriched with H3K27ac upon interferon stimulation. The knockout of MER41 impaired the expression of interferon-induced genes such as *AIM2* which senses cytosolic foreign DNA and activates inflammatory responses [20]. This suggests that ERV can sense the interferon signaling pathway and feedback to regulate innate immunity.

2.2. Alternative Promoters and Alternative Splicing. The LTR elements in ERVs possess the intrinsic promoter activity to drive ERVs expression. LTRs can also function as alternative promoters to drive host ORF expression. It has been estimated that up to 75% of human genes take advantage of alternative promoters to achieve tissue-specific regulation [21]. The employment of ERVs as alternative promoters not only results in stage- or tissue-specific gene expression patterns but also generates different isoforms of proteins [3, 21, 22]. Besides, ERVs are found over-represented in regions close to protein-coding sequences, suggesting that they are closely related to transcription initiation in the genome [23]. For instance, MT2 of the mouse ERVL sub-family is highly activated in mouse 2C embryo and functions as an alternative promoter to upregulate MERVL nearby genes, generating chimeric transcripts with junctions to MERVL elements [24]. An example to demonstrate is that *Zfp352* has two promoters (P1 and P2) that are active in mouse early embryo and somatic cells, respectively [25–27]. Interestingly, the active promoter of *Zfp352* in early embryos overlaps with MT2B1 repeats, indicating the ERV promoter may be critical for the early activation of *Zfp352* [25–27]. A recent large-scale transcriptomic analysis discovered that 23% of all protein-coding genes expressed in various cancer types possess at least two promoters that cause a significant tumor type-specific change in isoform expression [28]. For example, *JAZF1* prefers the 3′ full-length promoter (prmtr.40310) in KIRP cancer, whereas in KIRC cancer, a truncated promoter (prmtr.40312) is favored [28].

The presence of alternative promoters not only leads to context-dependent gene activation but also creates alternative splicing variants of the transcripts [21]. Alternative splicing can occur in the retroviral RNA itself, which has been correlated to cancer initiation[9]. For example, the open reading frame of HERVK provides a source for alternative splicing, and the spliced variants of HERVK can be detected in various cancers, some of which are cancer type-specific [29]. The differentially expressed retroviral RNA isoforms raise questions of how these isoforms are generated, and what functional differences exist between these isoforms. Apart from retroviral isoforms, ERVs are also involved in generating alternatively spliced isoforms in coding genes. For instance, the upstream MER4A can be utilized as an alternative promoter for *GTSO1*, which led to the generation of 15 isoforms of GTSO1 that may function differently under different disease contexts [30].

2.3. ERV-Derived Long Non-coding RNA. More importantly, many ERVs can encode for lncRNA. The functions of these lncRNAs can be involved in various processes like recruiting transcription factors, cooperating with epigenetic regulators or modifiers, or interacting with miRNAs [31–33].

A few studies demonstrated that the ERV-derived lncRNAs can participate in signaling transduction by regulating protein recruitment and protein degradation [34–37]. One of the ERV sub-family members, ALVE1, transcribes into lnc-ALVE1-AS1 to activate the TLR3 signaling pathway in the cytoplasm and induce antiviral innate immunity [35]. In addition, transcriptome analysis revealed that a human ERV-derived lncRNA, termed TROJAN, binds to metastasis-repressing factors and promotes their degradation through ubiquitin-associated signaling pathway [36], thus

promoting breast cancer progression. On the converse, antisense oligonucleotide repressing TROJAN slows down the breast cancer progression extraordinarily *in vivo*, suggesting that TROJAN promotes cancer invasion and can serve as a potential therapeutic target [36].

2.4. ERV-Derived Proteins. In addition to RNAs, the proteins translated from ERVs can also perform specific functions under certain contexts. These proteins are derived from the open reading frame of ERV, including *GAG*, *POL*, and *ENV*. The functions of these viral proteins are diversified [38–40]. For instance, the ENV protein from HERVK can upregulate the p-ERK1/2 and RAS signaling pathways in human pancreatic cancer, and knockdown of ENV suppressed the activity of the ERK signaling pathway [40]. Moreover, ENV proteins from HERVW and HERVFRD aid in trophectoderm cell fusion and facilitate mammalian embryo implantation into the uterus [41, 42], and the GAG protein produced by HERVK promotes prostate cancer progression by inducing androgen hormone release [38].

3. ERV in Stem Cell Differentiation

Embryonic development is initiated after fertilization, followed by zygote cleavage. In the early embryo cleavage stages, the zygotic genome is activated, accompanied by global remodeling and rewiring of the transcription network. Before the first cell fate segregation in late morula and blastocyst, cells in embryos retain the capacity to give rise to the complete embryo proper and are thus considered totipotent. In blastocyst, cells are committed to the outer layer trophectoderm and inner cell mass which gives rise to the pluripotent epiblast and differentiates into three germ layers and somatic tissues. Numerous genetic and epigenetic programs governing the embryo developmental processes have been revealed, but mostly focusing on the regulation of the coding genome. Non-coding elements such as ERVs are poorly understood in this context but are increasingly gaining attention. ERVs are extensively activated in early embryo development, with a highly restricted spatial-temporal pattern, and are drastically silenced during differentiation with exceptions of extraembryonic tissue and germlines (Figure 1). Here, we will focus on the functions and regulation of ERVs in a few key developmental stages and context to discuss the emergent roles of ERVs in chromatin regulation and stem cell differentiation.

3.1. ERV in Totipotency Regulation. During both mouse and human embryo development, ERVL subfamily is activated around ZGA but gradually silenced thereafter. It seems that ERVL is predominantly associated with the totipotent state. In mouse, transcripts from MERVL loci occupy 2% of the total mRNA in 2C embryo [24]. More than 307 genes were found to form chimeric transcripts with partial MERVL sequence [24]. These chimeric transcripts are mostly associated with metabolism and transcription regulation involved in mouse ZGA. For instance, in mouse 2C embryo, MT2-SPIN chimeric transcript excludes 3 exons at the N-terminus compared to the native isoform [43], resulting in the native and chimeric isoforms of SPIN that bear different phosphorylation sites by MAPK [43] and thus may mediate different signaling functions. MT2, together with partial MERVL-int sequence, is also a robust fluorescence reporter for 2C embryo as well as 2C-like cells in mouse embryonic stem cells (mESCs) [24]. MT2 also exhibits regulatory functions in activating distal 2C-specific genes. MT2 drives *Zscan4* cluster gene expression in mouse 2C embryo, and the upregulation of *Zscan4* can further activate MT2, resulting in DNA demethylation and open chromatin configuration to further activate 2C-specific genes nearby MT2 loci [44]. Interestingly, ectopic activation of MERVL by CRISPR activation system also resulted in the upregulation of 2C genes [45], implying that MERVL can act as a cis-regulatory element to control totipotent gene expression.

Similarly, in human, HERVL expression is also enriched in 8C stage corresponding to the time of human embryo ZGA [8]. MERVL and HERVL can be bound by mouse DUX and human DUX4, respectively, but cross-species binding is minimum, suggesting independent but converged evolution in mouse and human [19, 46]. Over-expression of *Dux* in mESCs can activate MERVL and downstream 2C genes. Similarly, human *DUX4* over-expression results in HERVL activation and simultaneously upregulation of human 8C-specific genes [19, 46].

Upon exiting from 2C stage, MERVL is rapidly silenced and its expression falls back to baseline in mouse 8C embryos. The silencing of MERVL is mediated by ZFP809. ZFP809 is a mouse-specific zinc finger protein, containing the KRAB domain at the N-terminus and seven zinc finger domains at the C-terminus [47]. The zinc finger domains allow ZFP809 to bind to the PBS sequence of MERVL, and the KRAB domain recruits TRIM28, together with NURD (histone deacetylase) and SETDB1 (histone methyltransferase), which led to condensed chromatin configuration and repression of MERVL activity [7, 47, 48]. Interestingly, it is noted that *Zfp809* produces two isoforms: a full-length protein and a truncated protein that lacks 50 residues at C-terminus. The full-length protein is selectively stable in ESCs but degraded in other cell types. Whereas the short isoform is constitutively expressed in both ESCs and differentiated cells, but the underlying impact and functional differences between the two differentially expressed isoforms remain unknown [47]. Nevertheless, a critical question that remained to be validated is whether the failure to silence MERVL will lead to the delay in the development of mouse early embryos, trapping the cells in totipotency.

3.2. ERV in Pluripotency Regulation. Upon exiting from totipotent state, cells take on the first cell fate decision to become extraembryonic trophectoderm or pluripotent epiblast. ERVL is rapidly silenced along with the exit from totipotency, while other sub-families of ERVs are upregulated [8, 11, 45]. HERVH sub-family is one of the most predominant ERVs in pluripotent stem cells. The internal sequence (ERV-int) is degenerated in a slower manner compared to other ERVs, suggesting the potential function of HERVH-int sequence in the pluripotent state [5, 6]. It is not known whether the silencing of HERVL is a prerequisite for the activation of HERVH during human embryo development. But it is possible that if HERVL is not silenced, the totipotency

transcription network will remain active, and cells might be trapped in the totipotent state. Similarly, forced activation of HERVL in pluripotent stem cells may also induce totipotent gene expression and shut down HERVH expression [45].

HERVH copies are highly enriched with the putative binding sites for pluripotent factors including KLF4, NANOG, and OCT4 [11]. In hESCs, HERVH is also enriched with H3K4me3 and H3K27ac [11], implying that they are potentially active promoters or enhancers for pluripotent gene regulation. Ectopic expression of HERVH subfamilies by CRISPR activation system can result in an extensive upregulation of genes up to 200 kb nearby of HERVH sequences [49]. Besides, a total of 128 and 145 chimeric transcripts of HERVH are detected in hiPSCs and hESCs respectively, suggesting HERVH can function as alternative promoters to activate pluripotency-related genes [11]. In contrast, native promoters of these genes are rarely active in pluripotent stem cells [11]. Although there could be potential functional distinctions between chimeric transcripts from ERV promoters and original transcripts from native promoters, the ERV-mediated activation of these genes in early embryonic development offers additional opportunity to rewire gene expression and innovate on the transcription regulation.

In addition, the lncRNAs derived from HERVH also play critical roles in pluripotency regulation. They may function as scaffold units to recruit chromatin modifiers and direct them towards specific locations [50, 51]. In detail, the HERVH lncRNAs mainly localize to the nucleus, and they can recruit chromatin modifiers such as P300 to the genomic loci of LTRs to regulate transcription of pluripotency genes nearby [52]. HERVH knockdown leads to fibroblast-like cell morphology [52] and downregulates more than 1000 genes observed, including a 50% reduction in NANOG and OCT4 expression, resulting in the partial loss of pluripotency and upregulation of differentiation markers [11]. In line with its role in hESCs, HERVH exhibited similar functions during somatic cell reprogramming [52]. HERVH expression is substantially upregulated upon ectopic expression of reprogramming factors, while depletion of HERVH during reprogramming leads to a reduction of iPSC colony-forming efficiency [52]. These shreds of evidence together indicate that HERVH is indispensable for both pluripotency establishment and maintenance.

Despite the importance of HERVH, pieces of evidence have been controversial about whether HERVH is required for naïve or primed pluripotency [11, 36, 52, 53]. Based on the LTR regions, HERVH can be further divided into several sub-families, such as LTR7Y, LTR7B, and LTR7. Some of the LTRs, like LTR7, are predominantly expressed in primed pluripotency [8], while LTR7Y may be more specific to naïve pluripotency [8]. Thus, naïve and primed pluripotency might employ different sub-families of HERVH controlled by the respective LTRs, but how this specificity is achieved requires further investigation.

3.3. ERV in Extraembryonic Tissue Differentiation. Research work has shed more light on the roles of ERVs in trophectoderm differentiation since the 1990s [54]. The roles of ERVs in extraembryonic tissue differentiation are mediated by regulating trophectoderm-specific transcription and by encoding for fusion proteins during the syncytia formation.

Many ERVs have a robust expression in placenta development [55]. Among all, HERVW, HERVFRD, and HERV3 are the top three active sub-families that encode for a high level of *ENV* gene [56, 57]. The SYNCITIN 1 translated from *the ENV* gene of HERVW lacks an immunosuppressive domain compared to full-length ENV protein. It is specifically upregulated in syncytiotrophoblast during implantation [41, 42]. The hydrophobic domain in SYNCITIN 1 enables its fusion with plasma membrane and potentially aids in uterus invasion [58]. Ectopic expression of HERVW *ENV* gene can induce cell fusion, which is reversed by neutralizing antibodies against SYNCITIN 1 [59]. In contrast, the lack of SYNCITIN 1 in primary trophoblast cells reduces the ability to form syncytia [60]. Similar to SYNCITIN 1, the SYNCITIN 2 produced by HERVFRD also promotes cell fusion upon ectopic expression in several cell lines [61]. Interestingly, ENV protein derived from HERV3 is expressed not only in syncytiotrophoblast but also in a wide range of tissues, particularly those producing hormones [62, 63]. More importantly, 1% of the Caucasian population bears a premature stop codon near the N-terminus, resulting in a non-functional short isoform of the protein. However, this does not lead to observable physiological defects in these individuals [54, 55]. Different ENV proteins from different ERV sub-families might play redundant roles. Apart from the proteins, ERVs also function as cis-regulatory elements in extraembryonic differentiation. In the mouse placenta, one of the ERV sub-families, RLTR13D5, is highly enriched with H3K27ac and H3K4me1, suggesting its potential role as an enhancer [64]. Moreover, RLTR13D5 can be functionally bound by CDX2, EOMES, and ELF5 to regulate the transcription in trophoblast stem cells and contribute to placenta development [64].

3.4. ERV in Somatic Tissue Differentiation. Despite the high activity of ERV and other TE families in early embryos, they are thought to be largely deactivated during the differentiation process. The silencing mechanism involves coevolution between the host transcription regulatory machinery and ERVs to tame ERV expression and limit their transposition. Improper silencing of ERVs is associated with loss of tissue homeostasis and pathological conditions. For example, the ENV protein derived from HERVW is highly expressed in type-1 diabetes and inhibited the secretion of insulin [65]. Transcripts and proteins of HERVK are also detected in amyotrophic lateral sclerosis brain tissue, which may contribute to the inhibition of neurite growth [66]. In human muscle cells, aberrantly expressed DUX4 binds to and induces HERVL expression, which serves as alternative promoters to alter the transcription network in facioscapulohumeral muscular dystrophy [10]. Moreover, HERV-derived lncRNA TROJAN promotes ubiquitin-associated degradation of metastasis-repressing factors and accelerates breast cancer progression [36].

In addition to the conventional view that ERV activation in differentiated tissue led to pathological conditions, more and more tissue-specific ERVs were identified, and they are thought to contribute to the cell type-specific differentiation

or tissue-specific functions [67]. For instance, during mouse gastrulation, different ERV sub-families were activated in various cell fates: erythroid has high RLTR10F activity, while mesoderm favors ERVB4 [67]. However, the exact function of these ERVs in the respective lineage remains elusive. Similarly, during the differentiation of human pluripotent stem cells (hPSCs) to cardiomyocytes *in vitro*, distinct sets of ERVs are selectively activated in different cell populations. For example, LTR32, MER57A-int, and MER45A are specifically expressed in definitive cardiomyocytes while MLT1H1, HERVIP10B-int, and LTR5A are selectively active in non-contractile cells [67]. It is noted that many ERV transcription regulators in ESCs, such as KLF-family members, are also expressed in tissue-specific cell types; thus, they may regulate ERV in the respective context.

Taken together, these pieces of evidence demonstrate two fundamental aspects of ERV in differentiation: (1) Various ERVs are now associated with tissue differentiation and specific cell lineage. (2) Aberrant ERV expression in differentiated tissues may be toxic while targeting these ERVs could provide potential therapeutic means to slow down disease progression.

3.5. ERV in Germline Formation. Although ERV activity may be largely silenced during differentiation, it is highly expressed and activated during germline formation. The first observation of ERV expression in germline cells can be dated back to 1983 when virus-like "intracisternal A particle (IAP)" was detected in mouse oocytes [68]. Up to now, more than 800 types of LTRs are detected in mouse oocyte, and they are involved in diverse functions, which aid in oocyte transcription regulation and facilitate oogenesis [69]. For instance, DICER protein is present in both mouse somatic cells and oocytes [70]. However, instead of being transcribed from native promoters, oocyte-specific DICER expression is driven by the LTR of MTC and produces an isoform lacking the N-terminal DExD helicase domain compared to full-length somatic DICER produced by its native promoter. And the deletion of LTR regions of MTC impaired oocyte-specific DICER, resulting in female sterility [70]. Many of the activated ERVs in the oocyte are passed down to the zygote as maternal factors, which are thought to be involved in ZGA [71], but their exact functions remain to be dissected in the future.

On the contrary, the progenitors of mouse germline cells, namely primordial germ cells (PGCs), show repressed ERV activity. ERV sequences are enriched with H3K9me3 and H3K27me3 that induce a repressive chromatin configuration [72]. In detail, SETDB1 as a methyltransferase protects PGCs from ERV activity. SETDB1 knockout PGCs show upregulated ERV activity, low survival rate, and postnatal hypogonadism [72]. Although this is in contrast to the general knowledge that ERVs are upregulated in germ cells, it is possible that different families of ERVs are involved in various stages of germ cell formation.

4. Conclusion and Outlook

ERVs are previously thought to coordinate with the host genome during mammalian evolution, and now they are considered as integral parts to form species and cell type-specific gene regulatory networks. The research of ERVs in stem cell fate decision and differentiation has just been unraveled, and many questions remained to be answered. Given the observed stage-specific expression pattern of ERV (Figure 1), what will be the specific function of each ERV sub-family in different developmental stages? How do different cell types achieve specific activation of ERV sub-families? What is the consequence of unscheduled activation or silencing of ERVs during early embryogenesis? Are ERVs exhibiting cell type-specific expression beyond blastocyst stages? Will ERV represent novel targets for diseases? Future studies will shed light on these questions and open up the fascinating but less charted road of ERVs.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We thank Ziying Tan for comments on the manuscript. This work was supported by the National Natural Science Foundation of China, grant number 31871372 and 31950410535.

References

[1] J. Goke and H. H. Ng, "CTRL+INSERT: retrotransposons and their contribution to regulation and innovation of the transcriptome," *EMBO Reports*, vol. 17, no. 8, pp. 1131–1144, 2016.

[2] R. K. Slotkin and R. Martienssen, "Transposable elements and the epigenetic regulation of the genome," *Nature Reviews. Genetics*, vol. 8, no. 4, pp. 272–285, 2007.

[3] E. B. Chuong, N. C. Elde, and C. Feschotte, "Regulatory activities of transposable elements: from conflicts to benefits," *Nature Reviews. Genetics*, vol. 18, no. 2, pp. 71–86, 2017.

[4] P. E. Jacques, J. Jeyakani, and G. Bourque, "The majority of primate-specific regulatory sequences are derived from transposable elements," *PLoS Genetics*, vol. 9, no. 5, article e1003504, 2013.

[5] C. Romer, M. Singh, L. D. Hurst, and Z. Izsvak, "How to tame an endogenous retrovirus: HERVH and the evolution of human pluripotency," *Current Opinion in Virology*, vol. 25, pp. 49–58, 2017.

[6] Z. Izsvak, J. Wang, M. Singh, D. L. Mager, and L. D. Hurst, "Pluripotency and the endogenous retrovirus HERVH: conflict or serendipity?," *BioEssays*, vol. 38, no. 1, pp. 109–117, 2016.

[7] D. Wolf and S. P. Goff, "Embryonic stem cells use ZFP809 to silence retroviral DNAs," *Nature*, vol. 458, no. 7242, pp. 1201–1204, 2009.

[8] J. Göke, X. Lu, Y. S. Chan et al., "Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells," *Cell Stem Cell*, vol. 16, no. 2, pp. 135–141, 2015.

[9] L. Agoni, C. Guha, and J. Lenz, "Detection of human endogenous retrovirus K (HERV-K) transcripts in human prostate cancer cell lines," *Frontiers in Oncology*, vol. 3, p. 180, 2013.

[10] J. M. Young, J. L. Whiddon, Z. Yao et al., "DUX4 binding to retroelements creates promoters that are active in FSHD

muscle and testis," *PLoS Genetics*, vol. 9, no. 11, article e1003947, 2013.

[11] J. Wang, G. Xie, M. Singh et al., "Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells," *Nature*, vol. 516, no. 7531, pp. 405–409, 2014.

[12] A. Argaw-Denboba, E. Balestrieri, A. Serafino et al., "HERV-K activation is strictly required to sustain CD133+ melanoma cells with stemness features," *Journal of Experimental & Clinical Cancer Research*, vol. 36, no. 1, p. 20, 2017.

[13] P. A. Tovo, I. Rabbone, D. Tinti et al., "Enhanced expression of human endogenous retroviruses in new-onset type 1 diabetes: potential pathogenetic and therapeutic implications," *Autoimmunity*, vol. 53, no. 5, pp. 283–288, 2020.

[14] S. Szpakowski, X. Sun, J. M. Lage et al., "Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements," *Gene*, vol. 448, no. 2, pp. 151–167, 2009.

[15] T. P. Hurst and G. Magiorkinis, "Epigenetic control of human endogenous retrovirus expression: focus on regulation of long-terminal repeats (LTRs)," *Viruses*, vol. 9, no. 6, p. 130, 2017.

[16] M. Imbeault, P. Y. Helleboid, and D. Trono, "KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks," *Nature*, vol. 543, no. 7646, pp. 550–554, 2017.

[17] U. Beyer, S. K. Kronung, A. Leha, L. Walter, and M. Dobbelstein, "Comprehensive identification of genes driven by ERV9-LTRs reveals TNFRSF10B as a re-activatable mediator of testicular cancer cell death," *Cell Death and Differentiation*, vol. 23, no. 1, pp. 64–75, 2016.

[18] A. Gaspar-Maia, A. Alajem, F. Polesso et al., "Chd1 regulates open chromatin and pluripotency of embryonic stem cells," *Nature*, vol. 460, no. 7257, pp. 863–868, 2009.

[19] P. G. Hendrickson, J. A. Doráis, E. J. Grow et al., "Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons," *Nature Genetics*, vol. 49, no. 6, pp. 925–934, 2017.

[20] E. B. Chuong, N. C. Elde, and C. Feschotte, "Regulatory evolution of innate immunity through co-option of endogenous retroviruses," *Science*, vol. 351, no. 6277, pp. 1083–1087, 2016.

[21] C. J. Cohen, W. M. Lock, and D. L. Mager, "Endogenous retroviral LTRs as promoters for human genes: a critical assessment," *Gene*, vol. 448, no. 2, pp. 105–114, 2009.

[22] P. Batut, A. Dobin, C. Plessy, P. Carninci, and T. R. Gingeras, "High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression," *Genome Research*, vol. 23, no. 1, pp. 169–180, 2013.

[23] G. J. Faulkner, Y. Kimura, C. O. Daub et al., "The regulated retrotransposon transcriptome of mammalian cells," *Nature Genetics*, vol. 41, no. 5, pp. 563–571, 2009.

[24] T. S. Macfarlan, W. D. Gifford, S. Driscoll et al., "Embryonic stem cell potency fluctuates with endogenous retrovirus activity," *Nature*, vol. 487, no. 7405, pp. 57–63, 2012.

[25] H. H. Chen, T. Y. Liu, C. J. Huang, and K. B. Choo, "Generation of Two Homologous and Intronless Zinc-Finger Protein Genes, *Zfp352* and *Zfp353* , with Different Expression Patterns by Retrotransposition," *Genomics*, vol. 79, no. 1, pp. 18–23, 2002.

[26] T. Y. Liu, H. H. Chen, K. H. Lee, and K. B. Choo, "Display of different modes of transcription by the promoters of an early embryonic gene, Zfp352, in preimplantation embryos and in somatic cells," *Molecular Reproduction and Development*, vol. 64, no. 1, pp. 52–60, 2003.

[27] S. X. Ge, "Exploratory bioinformatics investigation reveals importance of "junk" DNA in early embryo development," *BMC Genomics*, vol. 18, no. 1, p. 200, 2017.

[28] D. Demircioğlu, E. Cukuroglu, M. Kindermans et al., "A Pan-cancer transcriptome analysis reveals pervasive regulation through alternative promoters," *Cell*, vol. 178, no. 6, pp. 1465–1477.e17, 2019.

[29] L. Agoni, J. Lenz, and C. Guha, "Variant splicing and influence of ionizing radiation on human endogenous retrovirus K (HERV-K) transcripts in cancer cell lines," *PLoS One*, vol. 8, no. 10, article e76472, 2013.

[30] A. B. Conley, J. Piriyapongsa, and I. K. Jordan, "Retroviral promoters in the human genome," *Bioinformatics*, vol. 24, no. 14, pp. 1563–1567, 2008.

[31] M. C. Tsai, O. Manor, Y. Wan et al., "Long noncoding RNA as modular scaffold of histone modification complexes," *Science*, vol. 329, no. 5992, pp. 689–693, 2010.

[32] S. Panni, R. C. Lovering, P. Porras, and S. Orchard, "Non-coding RNA regulatory networks," *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1863, no. 6, article 194417, 2020.

[33] J. Durruthy-Durruthy, V. Sebastiano, M. Wossidlo et al., "The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming," *Nature Genetics*, vol. 48, no. 1, pp. 44–52, 2016.

[34] J. Wang, X. Li, L. Wang et al., "A novel long intergenic noncoding RNA indispensable for the cleavage of mouse two-cell embryos," *EMBO Reports*, vol. 17, no. 10, pp. 1452–1470, 2016.

[35] S. Chen, X. Hu, I. H. Cui et al., "An endogenous retroviral element exerts an antiviral innate immune function via the derived lncRNA lnc-ALVE1-AS1," *Antiviral Research*, vol. 170, p. 104571, 2019.

[36] X. Jin, X.-E. Xu, Y.-Z. Jiang et al., "The endogenous retrovirus-derived long noncoding RNA TROJAN promotes triple-negative breast cancer progression via ZMYND8 degradation," *Science Advances*, vol. 5, no. 3, article eaat9820, 2019.

[37] B. Zhou, F. Qi, F. Wu et al., "Endogenous retrovirus-derived long noncoding RNA enhances innate immune responses via derepressing RELA expression," *mBio*, vol. 10, no. 4, 2019.

[38] B. S. Reis, A. A. Jungbluth, D. Frosina et al., "Prostate cancer progression correlates with increased humoral immune response to a human endogenous retrovirus GAG protein," *Clinical Cancer Research*, vol. 19, no. 22, pp. 6112–6125, 2013.

[39] F. Wang-Johanning, M. Li, F. J. Esteva et al., "Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer," *International Journal of Cancer*, vol. 134, no. 3, pp. 587–595, 2014.

[40] M. Li, L. Radvanyi, B. Yin et al., "Downregulation of human endogenous retrovirus type K (HERV-K) viral ENV RNA in pancreatic cancer cells decreases cell proliferation and tumor growth," *Clinical Cancer Research*, vol. 23, no. 19, pp. 5892–5911, 2017.

[41] S. G. Black, F. Arnaud, M. Palmarini, and T. E. Spencer, "Endogenous retroviruses in trophoblast differentiation and placental development," *American Journal of Reproductive Immunology*, vol. 64, no. 4, pp. 255–264, 2010.

[42] C. Dong, M. Beltcheva, P. Gontarz et al., "Derivation of trophoblast stem cells from naïve human pluripotent stem cells," *eLife*, vol. 9, 2020.

[43] A. E. Peaston, A. V. Evsikov, J. H. Graber et al., "Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos," *Developmental Cell*, vol. 7, no. 4, pp. 597–606, 2004.

[44] M. A. Eckersley-Maslin, V. Svensson, C. Krueger et al., "MERVL/Zscan4 network activation results in transient genome-wide DNA demethylation of mESCs," *Cell Reports*, vol. 17, no. 1, pp. 179–192, 2016.

[45] D. Guallar, X. Bi, J. A. Pardavila et al., "RNA-dependent chromatin targeting of TET2 for endogenous retrovirus control in pluripotent stem cells," *Nature Genetics*, vol. 50, no. 3, pp. 443–451, 2018.

[46] A. De Iaco, E. Planet, A. Coluccio, S. Verp, J. Duc, and D. Trono, "DUX-family transcription factors regulate zygotic genome activation in placental mammals," *Nature Genetics*, vol. 49, no. 6, pp. 941–945, 2017.

[47] C. Wang and S. P. Goff, "Differential control of retrovirus silencing in embryonic cells by proteasomal regulation of the ZFP809 retroviral repressor," *Proceedings of the National Academy of Sciences*, vol. 114, no. 6, pp. E922–E930, 2017.

[48] G. Wolf, P. Yang, A. C. Füchtbauer et al., "The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses," *Genes & Development*, vol. 29, no. 5, pp. 538–554, 2015.

[49] D. R. Fuentes, T. Swigut, and J. Wysocka, "Systematic perturbation of retroviral LTRs reveals widespread long-range effects on human gene regulation," *eLife*, vol. 7, 2018.

[50] M. F. Pera and P. P. L. Tam, "Extrinsic regulation of pluripotent stem cells," *Nature*, vol. 465, no. 7299, pp. 713–720, 2010.

[51] D. J. Burgess, "HOTTIP goes the distance," *Nature Reviews Genetics*, vol. 12, no. 5, p. 300, 2011.

[52] X. Lu, F. Sachs, L. A. Ramsay et al., "The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity," *Nature Structural & Molecular Biology*, vol. 21, no. 4, pp. 423–425, 2014.

[53] J. Wang, M. Singh, C. Sun et al., "Isolation and cultivation of naive-like human pluripotent stem cells based on HERVH expression," *Nature Protocols*, vol. 11, no. 2, pp. 327–346, 2016.

[54] M. Cohen, M. Powers, C. O'Connell, and N. Kato, "The nucleotide sequence of the *env* gene from the human provirus ERV3 and isolation and characterization of an ERV3-specific cDNA," *Virology*, vol. 147, no. 2, pp. 449–458, 1985.

[55] J. R. Harris, "Placental endogenous retrovirus (ERV): structural, functional, and evolutionary significance," *BioEssays*, vol. 20, no. 4, pp. 307–316, 1998.

[56] P. J. W. Venables, S. M. Brookes, D. Griffiths, R. A. Weiss, and M. T. Boyd, "Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function," *Virology*, vol. 211, no. 2, pp. 589–592, 1995.

[57] J. L. Blond, D. Lavillette, V:. Cheynet et al., "An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor," *Journal of Virology*, vol. 74, no. 7, pp. 3321–3329, 2000.

[58] N. de Parseval, V. Lazar, J. F. Casella, L. Benit, and T. Heidmann, "Survey of human genes of retroviral origin: identification and transcriptome of the genes with coding capacity for complete envelope proteins," *Journal of Virology*, vol. 77, no. 19, pp. 10414–10422, 2003.

[59] S. Mi, X. Lee, X. P. Li et al., "Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis," *Nature*, vol. 403, no. 6771, pp. 785–789, 2000.

[60] J. L. Frendo, D. Olivier, V:. Cheynet et al., "Direct involvement of HERV-W env glycoprotein in human trophoblast cell fusion and differentiation," *Molecular and Cellular Biology*, vol. 23, no. 10, pp. 3566–3574, 2003.

[61] S. Blaise, N. de Parseval, L. Benit, and T. Heidmann, "Genome-wide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution," *Proceedings of the National Academy of Sciences*, vol. 100, no. 22, pp. 13013–13018, 2011.

[62] M. T. Boyd, C. M. R. Bax, B. E. Bax, D. L. Bloxam, and R. A. Weiss, "The human endogenous retrovirus ERV-3 is upregulated in differentiating placental trophoblast cells," *Virology*, vol. 196, no. 2, pp. 905–909, 1993.

[63] N. S. Rote, S. Chakrabarti, and B. P. Stetzer, "The role of human endogenous retroviruses in trophoblast differentiation and placental development," *Placenta*, vol. 25, no. 8-9, pp. 673–683, 2004.

[64] E. B. Chuong, M. A. K. Rumi, M. J. Soares, and J. C. Baker, "Endogenous retroviruses function as species-specific enhancer elements in the placenta," *Nature Genetics*, vol. 45, no. 3, pp. 325–329, 2013.

[65] S. Levet, J. Medina, J. Joanou et al., "An ancestral retroviral protein identified as a therapeutic target in type-1 diabetes," *JCI Insight*, vol. 2, no. 17, 2017.

[66] W. Li, M.-H. Lee, L. Henderson et al., "Human endogenous retrovirus-K contributes to motor neuron disease," *Science Translational Medicine*, vol. 7, no. 307, article 307ra153, 2015.

[67] J. He, I. A. Babarinde, L. Sun et al., *Unveiling transposable element expression heterogeneity in cell fate regulation at the single-cell level*, bioRxiv, 2020.

[68] G. G. Miller, I. V. Makarova, and A. A. Iazykov, "Detection of endogenous intracisternal type A particles in early mouse zygotes," *Biulleten'eksperimental'noi Biologii i Meditsiny*, vol. 96, pp. 94–96, 1983.

[69] V. Franke, S. Ganesh, R. Karlic et al., "Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes," *Genome Research*, vol. 27, no. 8, pp. 1384–1394, 2017.

[70] M. Flemr, R. Malik, V. Franke et al., "A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes," *Cell*, vol. 155, no. 4, pp. 807–816, 2013.

[71] A. V. Evsikov and C. Marin de Evsikova, "Gene expression during the oocyte-to-embryo transition in mammals," *Molecular Reproduction and Development*, vol. 76, no. 9, pp. 805–818, 2009.

[72] S. Liu, J. Brind'Amour, M. M. Karimi et al., "Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells," *Genes & Development*, vol. 28, no. 18, pp. 2041–2055, 2014.

Hindawi
Stem Cells International
Volume 2021, Article ID 6674283, 12 pages
https://doi.org/10.1155/2021/6674283

Research Article

Lin28 Inhibits the Differentiation from Mouse Embryonic Stem Cells to Glial Lineage Cells through Upregulation of Yap1

Juan Luo,[1] Hailin Zou,[1] Liang Deng,[2] Xiang Sun,[3,4] Ping Yuan,[5,6] and Peng Li[1]

[1]*Scientific Research Center, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518107, China*
[2]*Department of General Surgery, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518107, China*
[3]*Department of Medical Bioinformatics, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510275, China*
[4]*Center for Stem Cell Biology and Tissue Engineering, Key Laboratory for Stem Cells and Tissue Engineering, Ministry of Education, Sun Yat-sen University, Guangzhou 510275, China*
[5]*Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou 510655, China*
[6]*Guangdong Institute of Gastroenterology, Guangzhou, Guangdong 510655, China*

Correspondence should be addressed to Peng Li; lipeng56@mail.sysu.edu.cn

Received 21 October 2020; Revised 5 February 2021; Accepted 10 February 2021; Published 22 February 2021

Academic Editor: Qiang Wu

The RNA-binding protein Lin28 regulates neurogliogenesis in mammals, independently of the let-7 microRNA. However, the detailed regulatory mechanism remains obscured. Here, we established Lin28a or Lin28b overexpression mouse embryonic stem cells (ESCs) and found that these cells expressed similar levels of the core pluripotent factors, such as Oct4 and Sox2, and increased Yap1 but decreased lineage-specific markers compared to the control ESCs. Further differentiation of these ESCs to neuronal and glial lineage cells revealed that Lin28a/b overexpression did not affect the expression of neuronal marker βIII-tubulin, but dramatically inhibited the glial lineage markers, such as Gfap and Mbp. Interestingly, overexpression of Yap1 in mouse ESCs phenocopied Lin28a/b overexpression ESCs by showing defect in glial cell differentiation. Inhibition of Yap1/Tead-mediated transcription with verteporfin partially rescued the differentiation defect of Lin28a/b overexpression ESCs. Mechanistically, we demonstrated that Lin28 can directly bind to *Yap1* mRNA, and the induction of Yap1 by Lin28a in mESCs is independent of Let7. Taken together, our results unravel a novel Lin28-Yap1 regulatory axis during mESC to glial lineage cell differentiation, which may shed light on glial cell generation *in vitro*.

1. Introduction

The RNA-binding proteins Lin28a and Lin28b are homologs originally identified as developmental timing regulators in C. elegans [1, 2]. They are subsequently found to function in a wide spectrum of biological processes, developments, and diseases in mammals, including embryonic stem cell self-renewal, somatic cell reprogramming, metabolism, organismal growth, and tumorigenesis [3–7]. Lin28a/b inhibits the maturation of let-7 family members, the important players in multiple diseases and cancers via their cold-shock domain (CSD), and a pair of CCHC-type zinc finger motifs [8–10]. In addition, Lin28 directly binds active promoters and recruits Tet1 to regulate its target gene expression, demonstrating its dual binding affinity to DNA and RNA in diverse biological processes [11].

Lin28 is widely expressed in a number of tissues from embryo to adult, particularly in the nervous system [6, 12]. *In vivo* study showed that Lin28a knockout leads to reduced neural progenitor cell proliferation and a small brain in mouse, whereas knockout of both Lin28a alleles and one Lin28b allele displays similar but more severe phenotypes than the control, demonstrating the redundant and critical roles of Lin28a/b in the nervous system development [13]. In addition, *in vitro* study revealed that constitutive expression of Lin28a/b results in the promotion of neurogenesis but the block of gliogenesis, independently of the let-7 microRNA [14, 15]. However, what are the downstream targets of

Lin28 and how do they mediate Lin28 functions during neurogliogenesis remain largely unknown.

Pluripotent stem cells (PSCs), such as induced pluripotent stem cells (iPSCs) and embryonic stem cells (ESCs), can propagate *in vitro* and differentiate into all adult cells. They therefore provide useful materials to study cell differentiation and hold great promise for disease remodeling, drug discovery, and regenerative medicine [16–20]. To explore the regulatory mechanism of Lin28 during neurogliogenesis, we constitutively overexpressed Lin28a and Lin28b, respectively, in mouse ESCs and then directly differentiated them to neurons and glias *in vitro*. Overexpression of Lin28 strikingly inhibited the expression of glial lineage markers, like Gfap and Mbp, but did not affect the expression of neuronal marker βIII-tubulin. Interestingly, constitutive overexpression of Yap1 in mESCs showed a similar effect to Lin28a/b OE ESCs during *in vitro* neurogliogenesis, while inhibition of Yap1/Tead-mediated transcriptional output partially rescued these phenotypes. Furthermore, RNA immunoprecipitation and qPCR assays demonstrated Lin28 can directly bind to *Yap1* mRNA, and the induction of Yap1 by Lin28a in mESCs is independent of Let7. Our study reveals a novel Lin28-Yap1 regulatory axis in mESC to glial lineage cell differentiation *in vitro*.

2. Materials and Methods

2.1. ESC Culture and Differentiation. The mouse embryonic stem cell line used in this study was isolated from wild type C57/BL6 mouse as previously described [21, 22]. Typically, E3.5 embryos at the blastocyst stage were flushed out from the uterus and cultured on mitomycin-C treated mouse embryonic fibroblasts in a 96-well plate with N2B27 medium with 2i (0.4 μM PD0325901 and 3 μM CHIR99021) and LIF (1000 U/ml). The ICM (inner cell mass) outgrowths were treated with 0.05% Trypsin and passaged on a 24-well plate until stable ESC lines were obtained. The ESCs were maintained on feeders under the normal ES medium (DMEM supplemented with 15% FBS, 0.1 mM nonessential amino acids, 0.1 mM 2-mercaptoethanol, 2 mM Glutamine, 100 U/ml penicillin/streptomycin, and 1000 U/ml LIF). To obtain feeder-free ESCs, the ESCs were grown on a 0.1% gelatin-coated dish in 2i + LIF medium. For cell differentiation to neuronal and glial lineage cells, we integrated the differentiation protocols previously described [23, 24]. First, the mouse ESCs were trypsinized to single cells and then replaced at 1×10^4 cells per well of an ultralow adhesion 96-well plate to quickly aggregate and form uniformly sized embryoid bodies in KSR medium (high glucose DMEM supplemented with 15% knockout serum replacement, 0.1 mM nonessential amino acids, 0.1 mM 2-mercaptoethanol, 2 mM Glutamine, 100 U/ml penicillin/streptomycin, and 10 μM SB431542). After 5 days of suspension culture, the embryoid bodies were subjected to adhesion culture in N2 medium (DMEM/F12 supplemented with N2) for 10 days using a 6-well plate coated with 25 ng/ml human fibronectin.

2.2. DNA Constructs and Lentivirus Production and Infection. The lentiviral expression constructs pUbi-MCS-3xFlag (GV358), subcloned with mouse Lin28a/b or mouse Yap1, were purchased from the GeneChem company https://www.genechem.com.cn/. For lentiviral production and infection, lentiviral plasmid (1.2 μg), including the overexpressing plasmids or shRNA plasmid (pLKO.1-Puro), together with 0.8 μg of packaging plasmids pSPAX2 (Addgene #12260) and 0.5 μg of envelope expressing plasmid (Addgene #12259) were transiently cotransfected into the 293 T cells using the Lipofectamine 2000 reagent according to the manufacturer's instructions; 48 hours after transfection, the lentivirus supernatant was collected and filtered with 0.45 μm membrane filters (Millipore). mESCs were infected in the presence of 5 μg/mL polybrene and selected with 1 μg/mL puromycin for 72 hours, the oligo sequences of mouse Lin28a shRNA were listed in Supplementary Table 1.

2.3. Reverse Transcription and Quantitative Real-Time PCR. Reverse transcription and QRT-PCR assays were conducted following the previously described [22]. Total RNAs of mESCs were extracted using TRIzol reagent (TaKaRa) following the manufacturer's instructions. 1 μg of total RNA was used as templates to perform reverse transcription with the PrimeScript RT reagent Kit (TaKaRa) according to the instructions. Real-time PCR analysis was performed using the Bio-Rad machine with the SYBR Premix Ex Taq (TaKaRa). The generated threshold cycle (CT) value for each transcript was normalized against the CT value of an internal control, like β-Actin, and subsequently normalized against the CT value of corresponding transcripts of the control sample. The oligo sequences of RT primers were listed in Supplementary Table 1.

2.4. Western Blot Analysis. mESCs or ESC-derived cells on day 5, day 10, and day15 were lysed using the protein lysis buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1% tritone X-100, 0.1 mM EDTA, 0.5 mM MgCl2, inhibitors of proteases and phosphatases). Then, we followed the methods previously described [22]. In brief, protein samples were separated by SDS-polyacrylamide gel electrophoresis and transferred to PVDF membranes (Millipore). The membranes were blocked with 5% nonfat dry milk (BD Company) in TBST+0.1% Tween-20 and incubated with primary antibody in TBST+0.1% Tween-20 overnight at 4°C. The primary/secondary antibodies and dilutions used were listed in Supplementary Table 2.

2.5. Immunofluorescence Stain. Immunofluorescence stain of ESC differentiated cells was performed as previously describe [21, 25]. Briefly, cells were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) at room temperature (RT) for 30 minutes, followed by blocking with 1% BSA in PBS for 1 hour and then primary antibody overnight at 4 degrees. The primary antibodies and dilutions used were listed in Supplementary Table 2. After washing with PBS for 3 times on the second day, the samples were incubated with the appropriate secondary antibodies, conjugated with Alexa Fluor 594 (Molecular Probes) in PBS for 1 hour at room temperature. Cells were then counterstained with 4,6-diamidino-2-phenylindole dilactate (DAKO) for 15 min at

RT following the method as described [22]. Images were captured using a Carl Zeiss confocal microscope (LSM 800).

2.6. RNA Immunoprecipitation (RIP) Assay. The EZ-Magna RIP™ RNA Binding Protein Immunoprecipitation Kit (No.17-701; Millipore) was used to perform the RIP assay. In brief, around 2×107 control and Lin28a-flag overexpressed mESCs were lysed respectively, with RIP lysis buffer provided in the kit. Anti-Flag M2 magnetic beads (No.8223; Sigma) were incubated with lysates, and the Lin28a-flag-RNA complexes were precipitated. The complexes were washed and treated with proteinase K. RNA was extracted using the phenol/chloroform method, and the retrieved RNA was subjected to quantitative real-time RT-PCR with gene-specific primers listed in Supplementary Table 1.

2.7. Data Analysis. Statistical significance was determined by the unpaired Student's *t*-test. The *P* value is indicated by asterisks in the Figures ($P < 0.05$ [*] and $P < 0.01$ [**]). Differences of $P < 0.05$ and lower were considered statistically significant.

3. Results

3.1. Lin28a/b Induced Yap1 but Inhibited Lineage-Specific Gene Expression in mESCs. To investigate the underlying regulatory mechanisms of Lin28 during mESC to neuronal and glial differentiation, Lin28a and 28b stably expressed mESC lines were established respectively, by infecting ESCs with corresponding lentiviruses. Both Lin28a and 28b overexpression (Lin28a/b-Flag OE) cells maintained typical dome-shaped colony morphology and expressed a high level of mESC surface protein alkaline phosphatase like the control ESCs (Figure 1(a)). Consistently, the overall protein level of ESC core pluripotent factors, such as Oct4, Sox2, and Nanog, was not significantly altered upon Lin28a/b overexpression (Figure 1(b)). However, the expression of lineage-specific genes, including *Cdx2*, *Nestin*, *T*, and *Gata6*, was dramatically decreased in Lin28a/b OE mESCs (Figure 1(c)). Previous network-based analyses suggested that Lin28a may be a regulatory factor of the Hippo signaling pathway [26]. Hence, we further examined the expression of key Hippo pathway kinases, such as Mst and Lats, and phosphorylated Yap1 (S127 and 397), which marks Yap1 for cytoplasm localization and degradation through western blot [27]. All these proteins did not show obvious changes upon overexpression of Lin28a/b, but the total protein level of Yap1 and its downstream target Ctgf was dramatically upregulated, suggesting that Lin28a/b promote functional Yap1 protein via other means instead of the canonical Hippo pathway (Figure 1(d)). Further examination of *Yap1* mRNA by qRT-PCR assay revealed that Lin28a/b OE did not affect Yap1 expression at the transcriptional level in mESCs (Figure 1(e)). Additionally, two shRNAs targeting two different regions of the mouse *Lin28a* gene were constructed, and both of them could efficiently reduce Lin28a at both mRNA and protein levels. Consistent with Lin28a/b OE mESCs, knockdown of *Lin28a* significantly reduced the protein level of Yap1 and its downstream target Ctgf, but did not affect *Yap1* transcript level (Figure S1b-c). However, the decrease of *Lin28a* restricted the pluripotent maker expression and quickly induced mESC differentiation, and the qRT-PCR result showed the expression of lineage-specific genes was dramatically upregulated in *Lin28a* stably knockdown mESCs (Figure S1b-c). Taken together, these results demonstrated that Lin28a/b is necessary for maintaining the self-renewal and pluripotency of mES cells, and Lin28 specifically affects Yap1 protein abundance.

3.2. Lin28a/b Inhibited the Differentiation of mESCs to Glial Lineage Cells. Accumulating studies have showed that Lin28 regulates ESC proliferation and neurogenesis *in vitro*, but the underlying mechanism is not clear [14, 15]. Here, we utilized previously described protocols to directly differentiate mouse ESCs to neurons and glias *in vitro* [23, 24]. As shown in Figures 2(a) and 2(b), mESCs were trypsinized to single cells and cultured in an ultralow attachment 96-well plate to quickly aggregate and grow uniformly sized embryoid bodies (EBs) in KSR medium. After 5 days of suspension culture, the EBs were subjected to adhesion culture in N2 medium for another 10 days. 5 days of suspension culture in serum-free medium greatly induced cell growth and neural stem cell differentiation. We observed that the neural stem cell markers, including Sox1, Sox2, and Nestin, were dramatically increased, while the pluripotent marker Oct4 was decreased (Figures 2(c) and 2(e)). More interestingly, the protein level of Lin28a/b and Yap1 was upregulated compared to undifferentiated cells (day 0 mESCs), indicating their potential roles in neural lineage induction. Further adherent culture in N2 medium greatly facilitated the neuronal and glial lineage cell differentiation (Figure 2(b)). Western blot and immunofluorescence analyses showed the expression of neuron-specific class III-tubulin (Tubb3), glial fibrillary acidic protein (GFAP), and myelin basic protein (Mbp) was markedly increased after 5 days of adherent culture (Figures 2(d) and 2(e)). Besides, different neuronal subtypes including glutamatergic, GABAergic, dopaminergic, and cholinergic neurons were characterized based on the expression of vesicular glutamate transporter 2 (Vglut2), Gad1, Th, and Chat on day 15 of culture (Figure 2(d)). Taken together, these results manifest the establishment of the mouse ESC to glia cell and functional neuron differentiation system.

To assess the role of Lin28a/b in neural lineage differentiation, we conducted neuronal and glial lineage cell differentiation assays as above with Lin28a/b OE ESCs. The expression of Yap1 and neural progenitor markers, Sox1, Sox2, and Nestin was markedly upregulated in Lin28a/b OE cells compared to the control cells at day 5 of differentiation (Figure 3(a)). At day 10 of differentiation, the expression of neuron marker Tubb3 was comparable between control ESCs and Lin28a/b OE ESCs, but the expression of glial cell markers Gfap and Mbp was dramatically lower in Lin28a/b ESCs than the control ESCs (Figure 3(b)). Further extension of the differentiation time to day 15 could not reinstall the expression of Gfap in Lin28a/b cells, while different neuron markers Th, Chat, vGlut2, and Gad1 were expressed in both the control cells and Lin28a/b OE cells (Figure 3(c)).

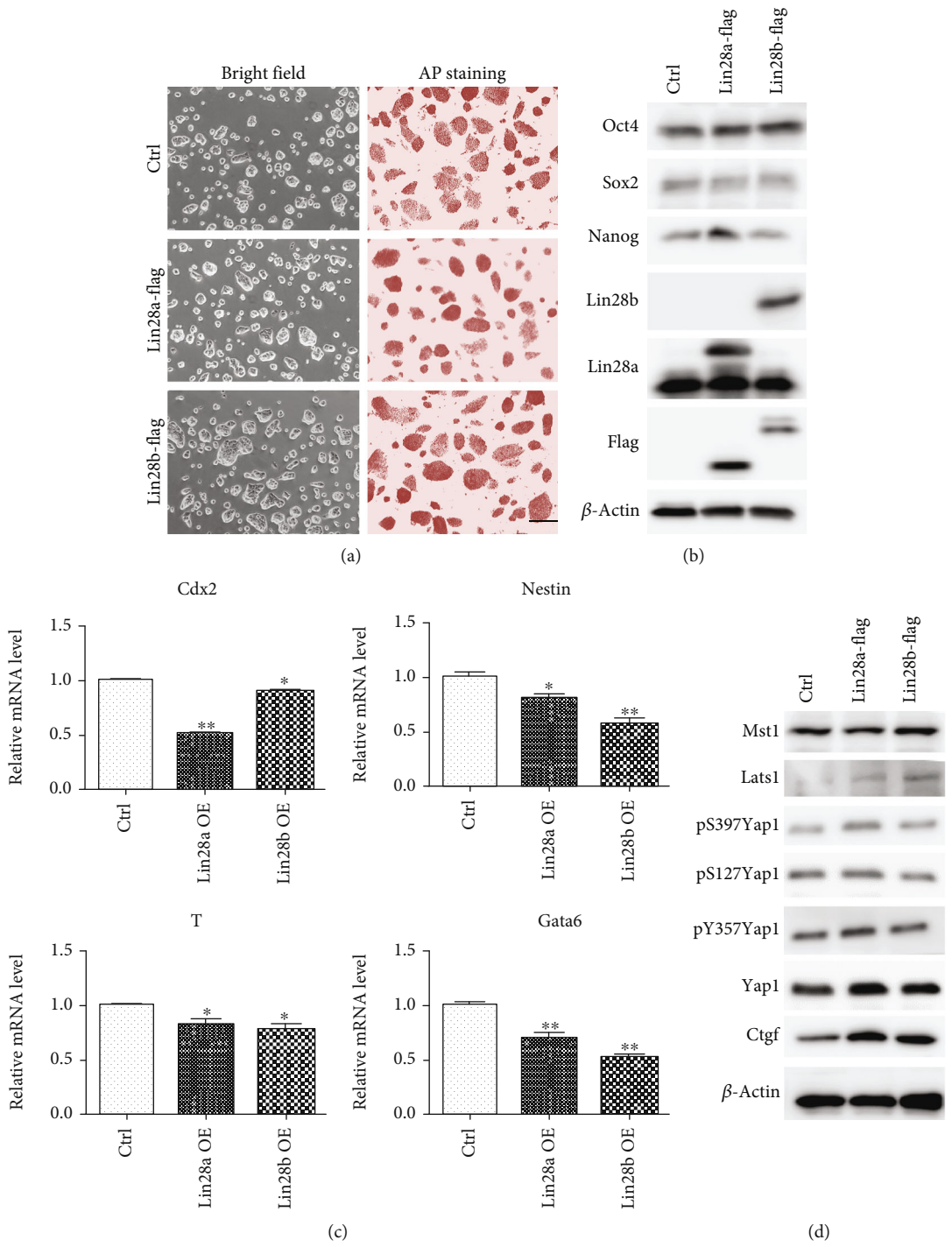


Figure 1: Continued.

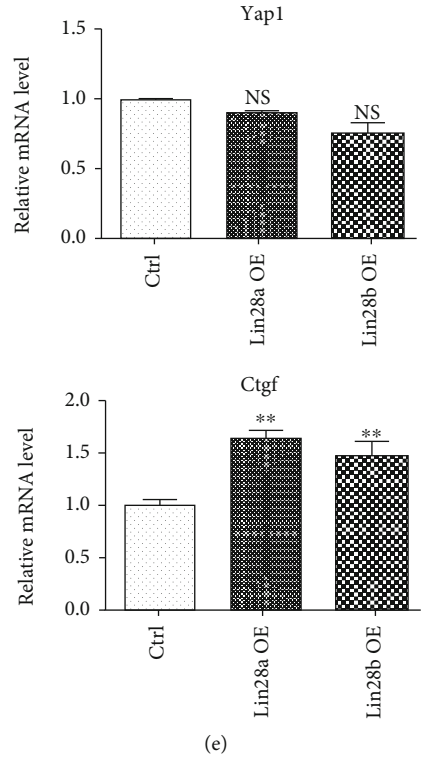


Figure 1: Constitutive expression of Lin28a/b in mouse ESC induced Yap1 expression and reduced lineage-specific gene expression. (a) Phase-contrast microscopy and AP staining of Ctrl and Lin28a/b constitutively expressed (Lin28a/b OE) mouse ESCs grown under 2i + LIF medium. Scale bar, 200 μm. (b) Western blot analyses of total proteins from Ctrl and Lin28a/b OE mouse ESCs using the indicated antibodies. Oct4, Sox2, and Nanog are pluripotent stem cell markers. (c) Quantitative real-time PCR to examine the mRNA level of lineage-specific gene expression in Ctrl and Lin28a/b OE mouse ESCs, trophectoderm gene *Cdx2*, ectoderm gene *Nestin*, mesoderm gene *T*, and endoderm gene *Gata6*. Actin was analyzed as an internal control. The data are shown as the mean ± S.D ($n = 3$). Statistically significant differences were indicated (*, $P < 0.05$ and **, $P < 0.01$). (d) Western blot analyses of total proteins from Ctrl and Lin28a/b OE mouse ESCs using the indicated antibodies. (e) Quantitative real-time PCR to examine the mRNA level of *Yap1* and its downstream target gene *Ctgf* expression in Ctrl and Lin28a/b OE mouse ESCs. The data are shown as the mean ± S.D ($n = 3$).

Immunofluorescence analyses also confirmed the expression of the neuronal markers Tubb3 and Th, but not the glial cell marker Gfap and Mbp in Lin28a/b OE ESCs, indicating a repressive role of Lin28a/b in ESC to glial cell differentiation (Figure 3(d)). More interestingly, Yap1 was always higher in Lin28a/b OE cells than the control cells throughout the *in vitro* differentiation procedure, suggesting a positive correlation between Lin28a/b and Yap1 (Figures 3(a)–3(c)).

3.3. Yap1 Overexpression mESCs Phenocopied the Glial Cell Lineage Differentiation Defect of Lin28a/b OE ESCs. To clarify whether upregulation of Yap1 in Lin28a/b OE cells was linked to the inhibition of glial lineage cells during the mouse ESC differentiation, we generated Yap1 overexpression (Yap1-Flag OE) mouse ESCs using lentivirus. These cells also maintained typical dome-shaped colony morphology and expressed a high level of mESC surface protein alkaline phosphatase similar to Lin28a/b OE cells (Figure 4(a)). In addition, Yap1 overexpression did not affect the expression of core pluripotency factors, but dramatically promoted its downstream target Ctgf (Figure 4(b)). Differentiation of Yap1 OE ESCs as above revealed that upregulation of Yap1 indeed promoted the expressions of neural stem cell markers, such as Sox1, Sox2, and Nestin, which is similar to Lin28a/b OE ESCs (Figure 4(c)). Likewise, Yap1 OE ESCs behaved like Lin28a/b OE cells in expressing a low level of glial lineage cell makers Gfap and Mbp, but comparable level of neuronal lineage marker Tubb3 and other functional neuronal markers like Th, Chat, vGlut2, and Gad1 as compared to the control cells upon induced differentiation (Figures 4(d)–4(g)), indicating that Yap1 may be a Lin28a/b effector in controlling of mouse ESC differentiation to neuronal and glial lineage cells.

3.4. Inhibition of Yap1-Tead Interaction Using Verteporfin Partially Rescued the Defect of Lin28a/b OE mESC Differentiation to Glial Cell Lineage. Yap1 has been demonstrated to act as a coactivator, and its transcriptional output is mainly dependent on the binding to Tead family transcription factors [28, 29]. To find out whether the upregulation of Yap1-mediated transcriptional output is responsible for the glial lineage differentiation defect of Lin28a/b OE mESC, we introduced the Yap1-Tead interaction inhibitor verteporfin to the *in vitro* differentiation assay of Lin28a OE mESCs.

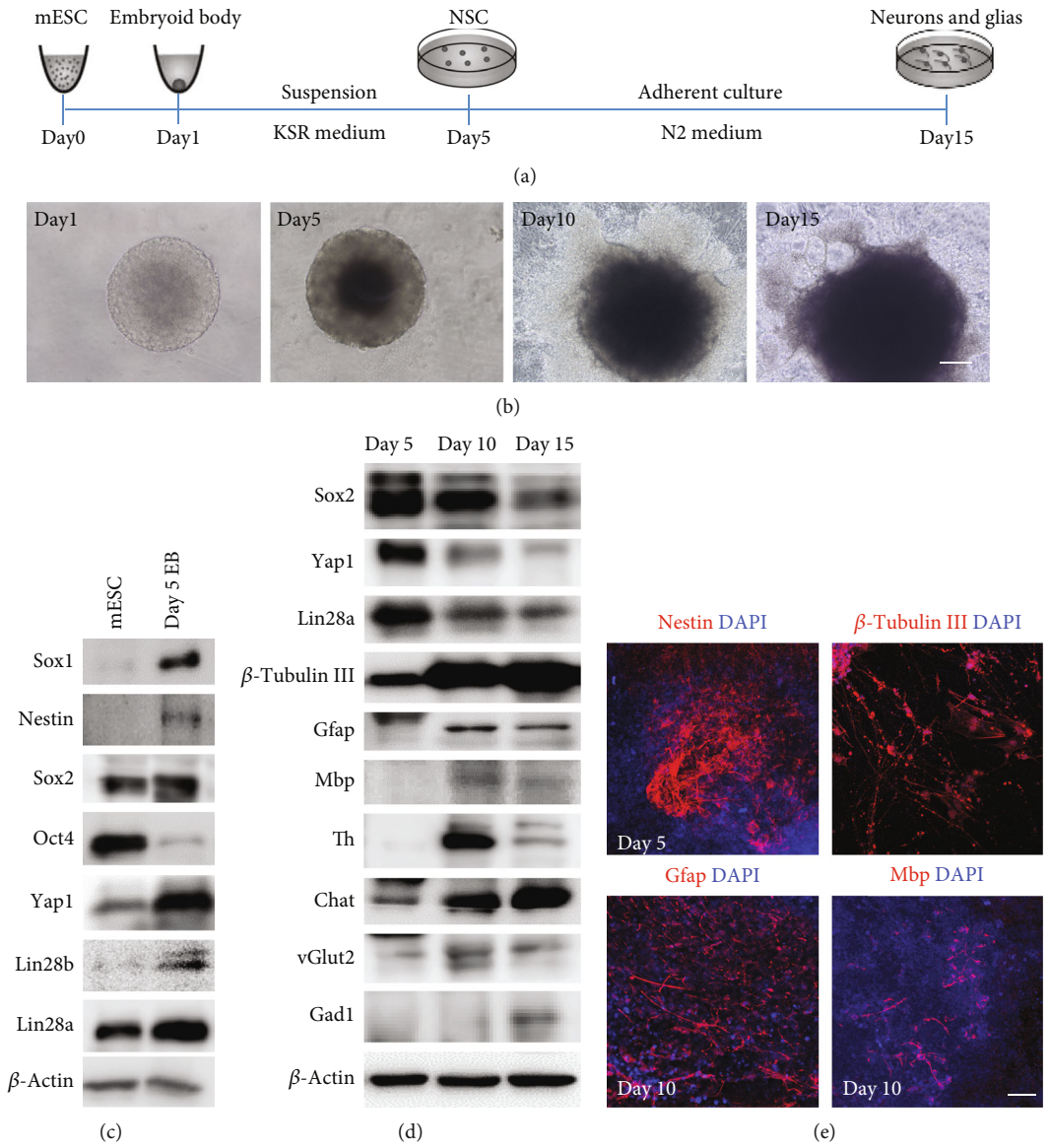


Figure 2: Establishment of the *in vitro* differentiation protocol from mouse ESCs to neuronal and glial lineage cells. (a) A schematic drawing of the direct differentiation assay from mouse ESCs to neuronal and glial lineage cells. ESCs were cultured under feeder-free condition with 2i + LIF medium for 1 day and then disassociated to single cells and quickly aggregated in differentiation medium for one day. After 5 days of suspension culture in KSR medium, the aggregates were subjected to adhesion culture for another 10 or 15 days in N2 medium. (b) Phase-contrast microscopy of mouse ESC differentiated cells on days 1, 5, 10, and 15. Scale bar, 200 μm. (c) Western blot analyses of total proteins from mESC differentiated cells on day 0 and 5 using the indicated antibodies. Mouse ESC pluripotent markers: Oct4 and Sox2, neural stem cell markers: Sox1, Nestin, and Sox2. (d) Western blot analyses of total proteins from mouse ESC differentiated cells on days 5, 10, and 15 using the indicated antibodies. Neuronal marker: β-tubulin III, glial markers: Gfap and Mbp, neuronal subtype markers: Th, vGlut2, Gad1, and Chat. (e) Immunofluorescence staining of the neural stem cell marker Nestin on day 5, neuronal marker β-tubulin III and glial markers Gfap and Mbp on day 10. Cell nuclear was stained with DAPI. Scale bar, 200 μm.

We first treated Lin28a OE ESCs with different concentrations of verteporfin for 24 hours. We found that increasing the inhibitor concentration dramatically reduced Yap1 and Ctgf protein levels, and the expression of Lin28 and core pluripotency factors Oct4 and Sox2 was reduced too (Figure 5(a)). Since 0.5 μM verteporfin noticeably reduced Yap1 function, we next treated the adherent cells differentiated from Lin28a OE mESCs on day 5 at this concentration and then analyzed the expression of neuronal and glial lineage markers by immunoblotting and immunofluorescence assays on day 10 and 15, respectively (Figure 5(b)). We found that inhibition of Yap1-Tead interaction with verteporfin partially rescued the expression of glial lineage markers, such as Gfap and Mbp, but did not affect the expression of Tubb3 and the functional neuronal markers, like Th and Chat in Lin28a OE mESC-differentiated cells (Figures 5(c)–5(e)),

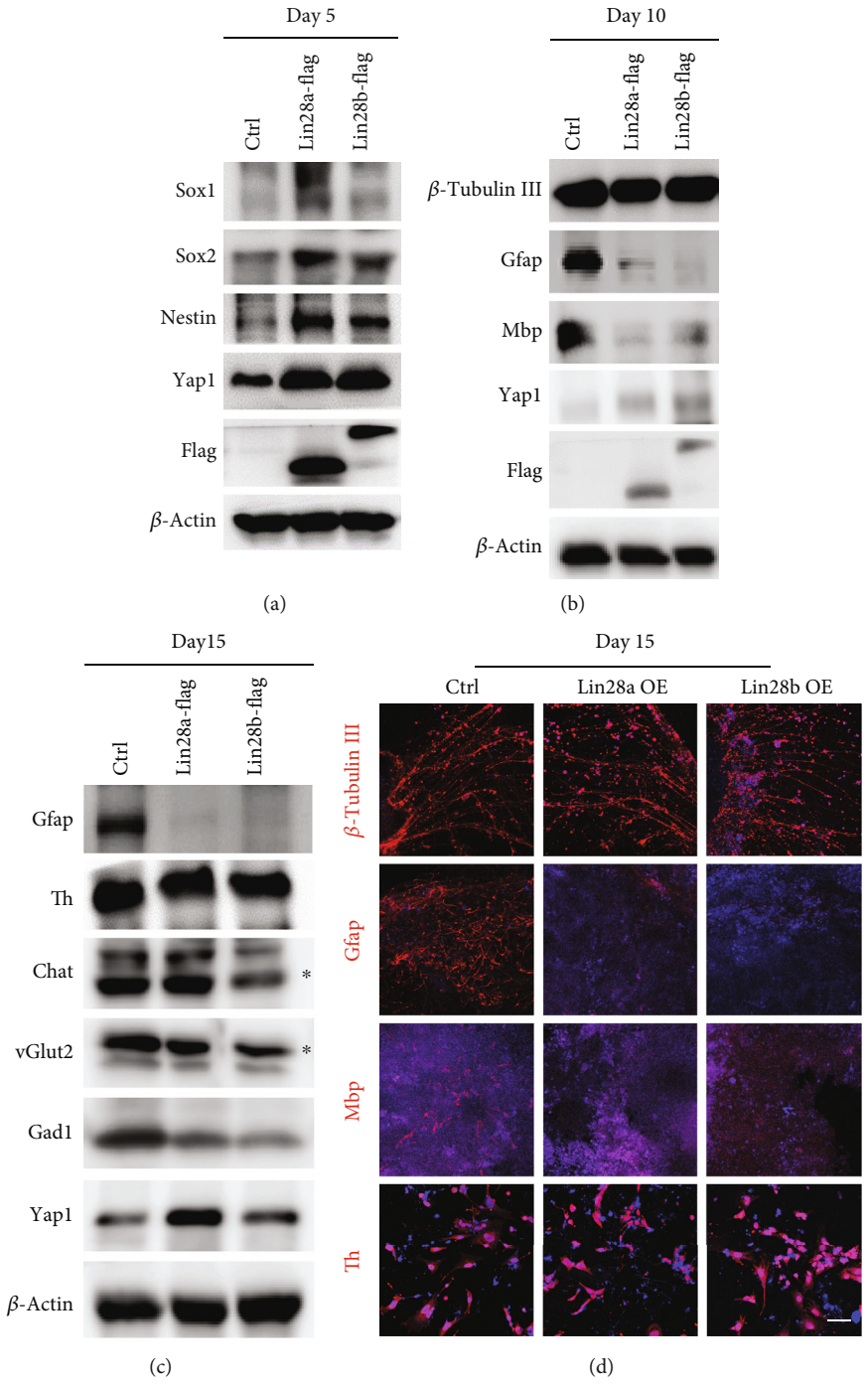


Figure 3: Constitutive overexpression of Lin28a/b inhibited the differentiation of mouse ESCs to glial lineage cells. (a) Western blot analyses of total proteins from Ctrl and Lin28a/b OE mouse ESC differentiated cells on day 5 using the indicated neural stem cell markers: Sox1, Nestin, and Sox2. (b) Western blot analyses of total proteins from Ctrl and Lin28a/b OE mouse ESC differentiated cells on day 10 using the indicated neuronal marker: β-tubulin III, glial markers: Gfap and Mbp. (c) Western blot analyses of total proteins from Ctrl and Lin28a/b OE mouse ESC differentiated cells on day 15 using the indicated neuronal subtype markers: Th, vGlut2, Gad1, and Chat. (d) Immunofluorescence staining of the Ctrl and Lin28a/b OE mouse ESC differentiated cells on day 15 using the neuronal marker β-tubulin III and glial markers Gfap and Mbp, and neuronal subtype markers Th. Cell nuclear was stained with DAPI. Scale bar, 200 μm.

indicating that the glial lineage differentiation defect of Lin28a/b OE mESCs was to some extent caused by increased Yap1/Tead-mediated transcriptional output.

3.5. Induction of Yap1 by Lin28a/b in Mouse ESCs Was Independent of Let7. Lin28 is well known to be a suppressor of let-7 miRNA biogenesis and let-7b plays a critical role

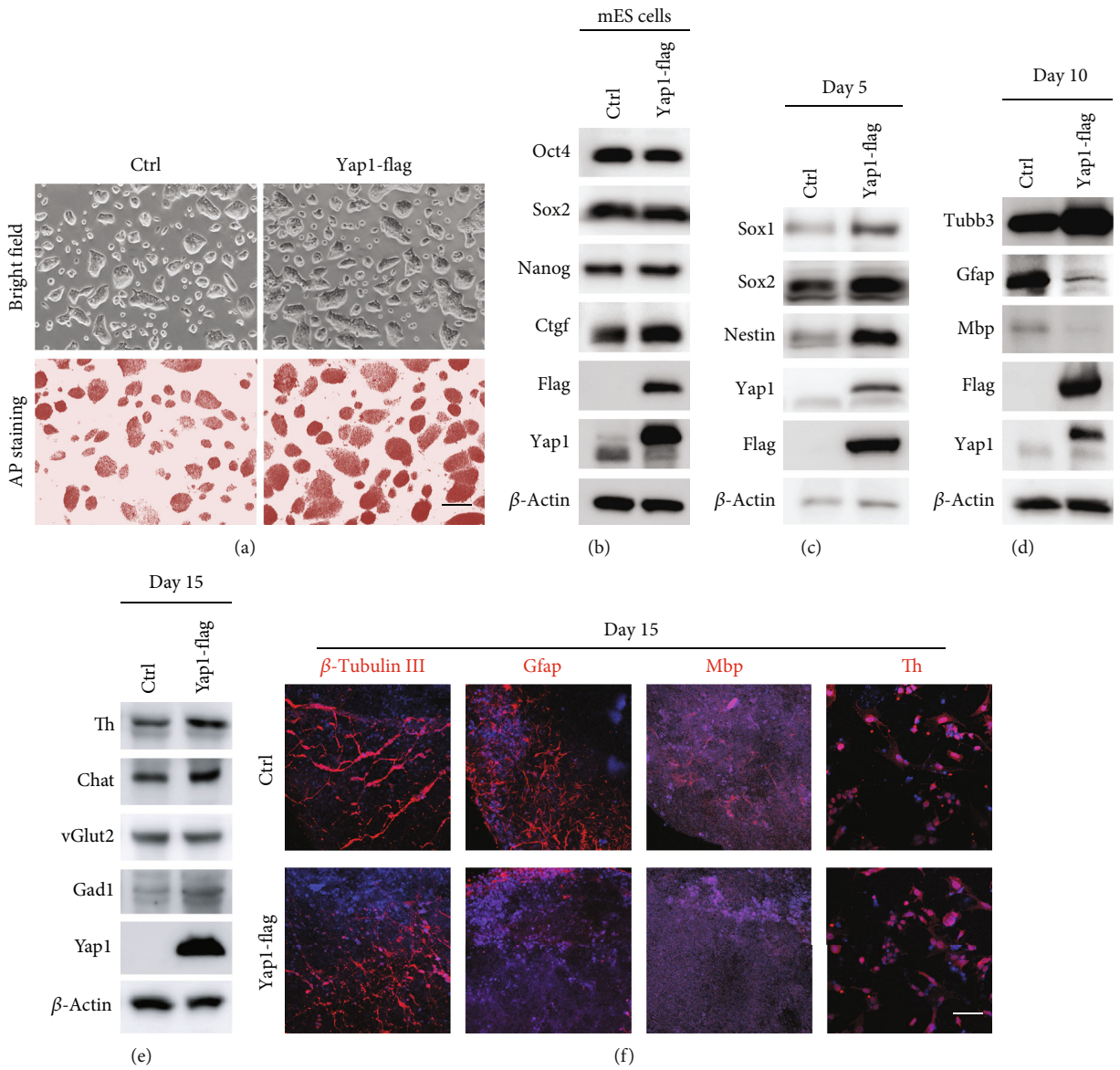


FIGURE 4: Yap1 overexpression in mESCs phenocopied the glial cell lineage defect differentiated from Lin28a/b OE cells. (a) Phase-contrast microscopy and AP staining of Ctrl and Yap1 constitutively expressed (Yap1 OE) mouse ESCs grown under 2i + LIF medium. Scale bar, 200 μm. (b) Western blot analyses of total proteins from Ctrl and Yap1 OE mouse ESCs using the indicated antibodies. Oct4, Sox2, and Nanog are pluripotent stem cell markers. (c) Western blot analyses of total proteins from Ctrl and Yap1 OE mouse ESC differentiated cells on day 5 using the indicated neural stem cell markers: Sox1, Nestin, and Sox2. (d) Western blot analyses of total proteins from Ctrl and Yap1 OE mouse ESC differentiated cells on day 10 using the indicated neuronal marker: β-tubulin III, glial markers: Gfap and Mbp. (e) Western blot analyses of total proteins from Ctrl and Yap1 OE mouse ESC differentiated cells on day 15 using the indicated neuronal subtype markers: Th, vGlut2, Gad1, and Chat. (f) Immunofluorescence staining of the Ctrl and Yap1 OE mouse ESC differentiated cells on day 15 using the neuronal marker β-tubulin III, glial markers Gfap and Mbp, and neuronal subtype markers Th. Cell nuclear was stained with DAPI. Scale bar, 200 μm.

in neural stem cell differentiation [30, 31]. Therefore, we investigated whether Lin28 regulates Yap1 via let-7 miRNA members in mouse ESCs. As Lin28 inhibitor LI71 and Lin28-let-7a antagonist 1 can inhibit Lin28a and let-7 interaction and miRNA processing in ESCs effectively [32, 33], we then treated Lin28a OE mESCs with different concentration of LI71 and Lin28-let-7a antagonist 1 for 24 hours, respectively, for protein analyses. It turned out that increasing the inhibitor concentration did not alter the expression of Yap1 in mouse ESCs (Figures 6(a) and 6(b)), indicating that induction of Yap1 by Lin28a/b overexpression in mouse ESCs was independent of the Let7 pathway. Then, we want to address whether Lin28 can directly bind *Yap1* mRNA to regulate its translation, we performed RNA immunoprecipitation using anti-Flag antibody in Lin28a-Flag overexpressed mESCs. *Tubulin* mRNA was used as a control for nonspecific RNA binding, and *H2a* and *Cyclin B* mRNAs were used as positive controls. We found both of *Yap1* and *Taz* mRNAs exhibited dramatic enrichments in the Lin28a-Flag overexpressed mESCs; these data

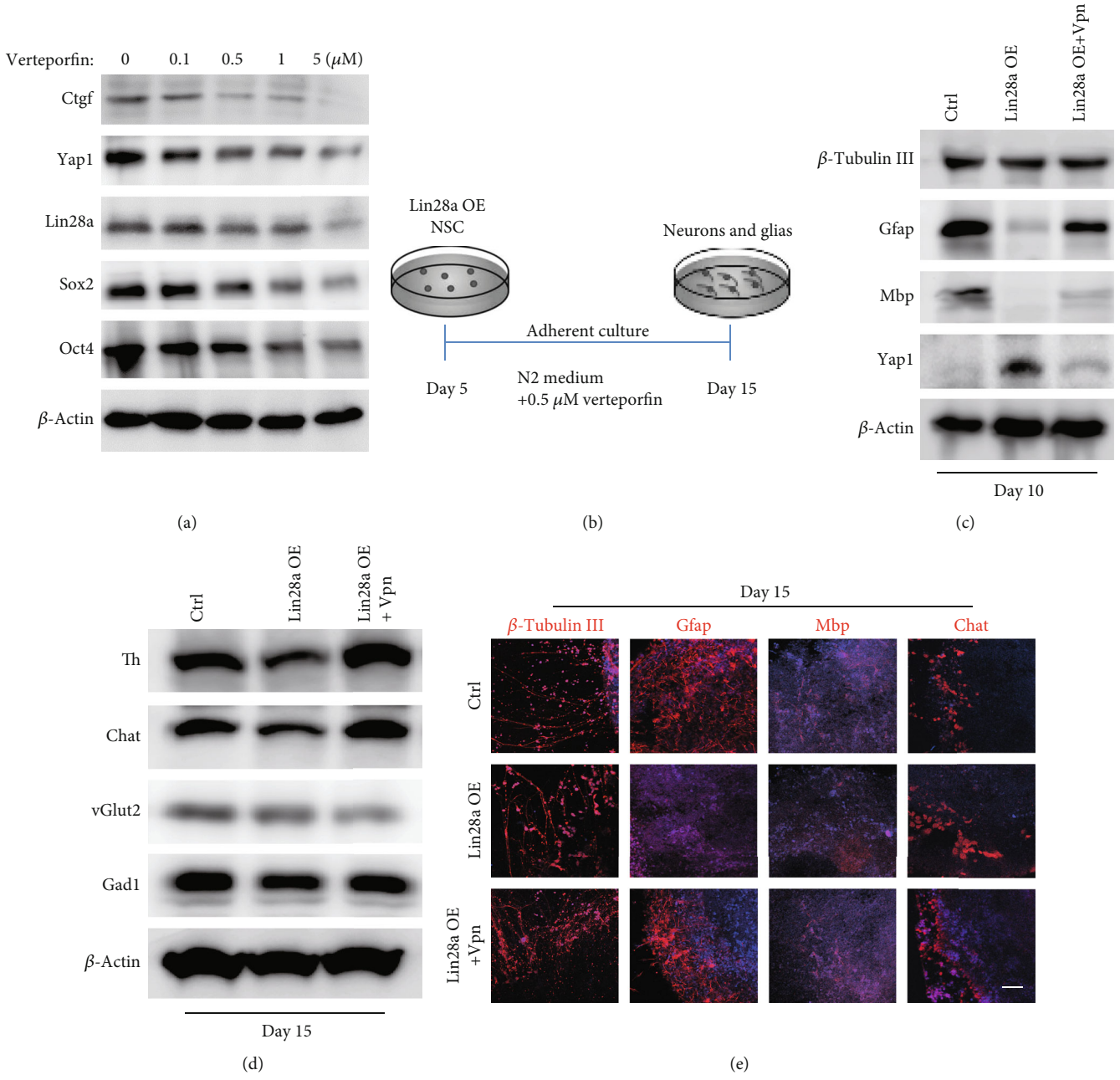


Figure 5: Inhibition of Yap1-Tead interaction using verteporfin partially rescued the glial cell lineage differentiation defect of Lin28a/b OE mouse ESCs. (a) Western blot analyses of total proteins from Ctrl mouse ESCs treated with different concentrations of Yap1-Tead interaction inhibitor (verteporfin) using the indicated antibodies. (b) A schematic drawing of the differentiation assay from neural stem cells (day 5) to neuronal and glial lineage cells (day 15) combining with verteporfin. (c) Western blot analyses of total proteins from Ctrl and Lin28a OE mouse ESC differentiated cells (treated with DMSO and Verteporfin, respectively) on day 10 using the indicated antibodies. (d) Western blot analyses of total proteins from Ctrl and Lin28a OE mouse ESC differentiated cells (treated with DMSO and verteporfin, respectively) on day 15 using the indicated antibodies. (e) Immunofluorescence staining of the Ctrl and Lin28a OE mouse ESC differentiated cells (treated with DMSO and verteporfin, respectively) on day 15 using the neuronal marker β-tubulin III, glial markers Gfap and Mbp, and neuronal subtype markers chat. Cell nuclear was stained with DAPI. Scale bar, 200 μm.

demonstrated that Lin28 can directly bind to *Yap1* mRNA (Figure S2). Collectively, we reported a novel Lin28a/b-Yap1 regulatory axis in mouse ESC to glial lineage cell differentiation (Figure 6(c)). Lin28a/b may directly bind to Yap1 mRNA to regulate its translation, and the induction of Yap1 by Lin28a in mESCs is independent of canonical Hippo pathway and let-7 family members.

4. Discussion

Previous *in vitro* and *in vivo* studies have demonstrated that Lin28a/b plays essential roles during central nervous development, and its functions are independent of the let-7 microRNA [13–15]. However, the downstream targets underlying Lin28 function in neurogliogenesis are not well understood.

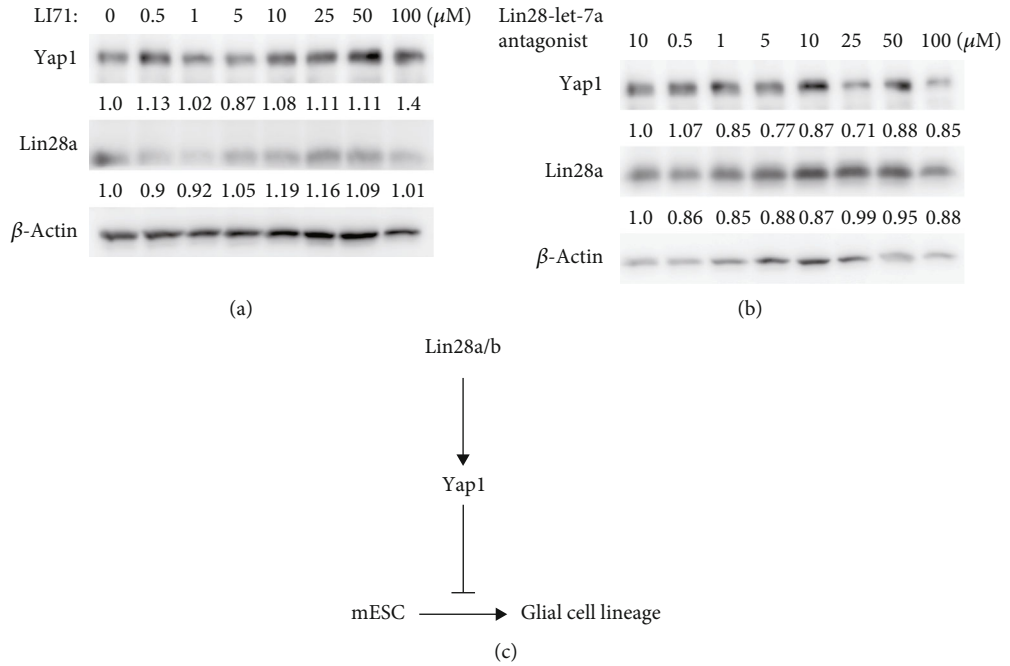


Figure 6: Induction of Yap1 by Lin28a/b in mouse ESCs was independent of Let7 pathway. (a) Western blot analyses of total proteins from Lin28a OE mouse ESCs treated with different concentrations of Lin28-let7 inhibitor (LI71) using the indicated antibodies. (b) Western blot analyses of total proteins from Lin28a OE mouse ESCs treated with different concentrations of Lin28-let7a antagonist using the indicated antibodies. (c) Schematic of the Lin28a/b-Yap1 pathway activity regulating mouse ESC differentiation to glial lineage cells.

Here, our studies have identified Yap1 as a crucial downstream effector of Lin28 and demonstrated that Yap1/Tead-mediated transcriptional output is partially responsible for Lin28a/b overexpression induced glial cell differentiation defect from mouse ESCs.

Yap1, a key transcriptional cofactor that is negatively regulated by the Hippo pathway, is essential for the development and size control of multiple organs [34, 35]. In addition, more diverse functions of the Hippo pathway have been recognized, including cell proliferation, differentiation, and migration [12, 36]. Overexpression of Yap1 in mouse ESCs inhibits ESC differentiation and is sufficient to maintain stem cell characteristics. Activation of Yap1 in mouse fibroblasts enhances reprogramming efficiencies of mouse iPSCs. All these evidences, including what we observed in this study and previously described, are consistent with the phenotypes of overexpressing Lin28a/b in mouse ESCs [37–40]. Yap1 is dramatically increased during ESC to neural stem cell differentiation, and Yap1 overexpression promotes the expression of neural progenitor cell markers, including Sox1 and Nestin, indicating its critical role in neural stem cell commitment from ESCs. However, in the late differentiation stage, from day 5-15, Yap1 is gradually decreased. This is consistent with the previous identification of Yap1 as a repressor during neurogliogenesis [38, 41]. The Sox2-Lin28 pathway has been demonstrated to govern the neural progenitor cell proliferation and neurogenesis but repressed the gliogenesis *in vitro* [15, 30]. In our *in vitro* differentiation assays, we observed that the Yap1 expression profile is extremely similar to Sox2 and Lin28, and constitutive overexpression of Yap1 in mouse ESCs restricts the cell differentiation to glial cell lineage, and further inhibition of the Yap1/Tead-mediated transcriptional output partially rescued the differentiation defect of Lin28a/b OE ESCs to glial cell lineage, confirming the critical role of Yap1 during ES to glial cell differentiation.

As the downstream effector of the Hippo pathway, Yap1 is regulated by a highly conserved kinase cascade Mst and Lats [34]. Previous network-based expression analyses have revealed that Lin28 is a possible regulatory nuclear factor of the Hippo pathway in stem cells [26]. Overexpression of Lin28a/b in mouse ESCs does not affect the expression of Yap1 upstream kinases, such as Mst and Lats, and Yap1 phosphorylation levels on the key functional sites, including S127, 397, and Y357, indicating the regulation of Yap1 by Lin28a/b are independent of the canonical Hippo pathway. In our study, inhibition of the Lin28 and Let-7 interaction in mouse ESCs does not alter the Yap1 protein level, suggesting that the induction of Yap1 by Lin28a/b is independent of the Let7 pathway, and some novel mechanisms may be adopted by Lin28 to regulate Yap1 during mESC to glial lineage cell differentiation. Previous studies have shown that Lin28 could either function as a DNA-based regulator or directly affect target mRNA translation and splicing; ChIP-seq data has demonstrated that Lin28a could directly recruit Tet1 and bind to the 5-UTR of *Yap1* in mouse ES cells [11]. However, in our study, we observed that *Yap1* mRNA was not changed, while the Yap1 protein level and its downstream target Ctgf were dramatically induced in the mESC stage and throughout neural lineage differentiation with Lin28-overexpressed cells, suggesting that Lin28a/b may participate in the posttranscriptional regulation of *Yap1* expression. In mouse and human ES cells, several studies have shown that Lin28 regulated the expression of cell cycle-related and pluripotency-associated genes, such as *Cyclin B*

and *Oct4* by directly binding to these target mRNAs and enhancing their translation [42–44]. Indeed, our RNA-IP and qPCR assays also validated that Lin28a can directly bind to *Yap1* mRNA, which is similar to the regulation of *H2a* and *Cyclin B* by Lin28a in mouse ESCs. Collectively, our study supports the hypothesis that Lin28a/b directly binds to *Yap1* mRNA and participates in its translation regulation, and further elucidating the regulatory mechanism of Yap1 by Lin28 would be necessary in the future.

Data Availability

Data can be available on request.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Juan Luo and Hailin Zou contributed equally to this work.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (Grant No. 20ykzd02), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2019A1515110285), and National Natural Science Foundation of China (Grant No. 32000679).

Supplementary Materials

Supplementary Table 1: oligo sequences of primers used in this paper. Supplementary Table 2: antibody list used in this paper. Supplementary Figure S1: (a) phase-contrast microscopy of shNC and Lin28a stable knockdown mouse ESCs grown under 2i + LIF medium. Scale bar, 200 μm. (b) Western blot analyses of total proteins from shNC and Lin28a stable knockdown mouse ESCs using the indicated antibodies. (c) qRT-PCR to examine the mRNA level of Lin28a, Yap1, Ctgf, and lineage-specific gene expression in shNC and Lin28a stable knockdown mouse ESCs. The data are shown as the mean ± S.D ($n = 3$). Statistically significant differences were indicated (*, $P < 0.05$ and **, $P < 0.01$). Supplementary Figure S2: Ctrl and Lin28-Flag overexpressed mouse ES cells were used to do Flag immunoprecipitation (Flag-IP), and RNA samples extracted from IP complexes were reverse-transcripted to generate cDNAs, followed by qPCR using the following gene primers. mRNA levels present in Lin28-Flag overexpressed mouse ES cells relative to control are shown. Each bar represents mean ± S.D ($n = 3$). *(Supplementary Materials)*

References

[1] E. G. Moss, R. C. Lee, and V. Ambros, "The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA," *Cell*, vol. 88, no. 5, pp. 637–646, 1997.

[2] E. G. Moss and L. Tang, "Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites," *Developmental Biology*, vol. 258, no. 2, pp. 432–442, 2003.

[3] N. Shyh-Chang and G. Q. Daley, "Lin28: primal regulator of growth and metabolism in stem cells," *Cell Stem Cell*, vol. 12, no. 4, pp. 395–406, 2013.

[4] Z. Ye, Z. Su, S. Xie et al., "Yap-lin28a axis targets let7-Wnt pathway to restore progenitors for initiating regeneration," *Elife*, vol. 9, 2020.

[5] H. Xiong, J. Shen, Z. Chen et al., "H19/let-7/Lin28 ceRNA network mediates autophagy inhibiting epithelial-mesenchymal transition in breast cancer," *International Journal of Oncology*, vol. 56, no. 3, pp. 794–806, 2020.

[6] T. Sato, K. Kataoka, Y. Ito et al., "Lin28a/let-7 pathway modulates the Hox code via Polycomb regulation during axial patterning in vertebrates," *Elife*, vol. 9, 2020.

[7] M. Farzaneh, F. Attari, and S. E. Khoshnam, "Concise review: LIN28/let-7 signaling, a critical double-negative feedback loop during pluripotency, reprogramming, and tumorigenicity," *Cellular Reprogramming*, vol. 19, no. 5, pp. 289–293, 2017.

[8] F. E. Loughlin, L. F. Gebert, H. Towbin, A. Brunschweiger, J. Hall, and F. H. Allain, "Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28," *Nature Structural & Molecular Biology*, vol. 19, no. 1, pp. 84–89, 2011.

[9] Y. Nam, C. Chen, R. I. Gregory, J. J. Chou, and P. Sliz, "Molecular basis for interaction of let-7 microRNAs with Lin28," *Cell*, vol. 147, no. 5, pp. 1080–1091, 2011.

[10] I. Bussing, F. J. Slack, and H. Grosshans, "*let-7* microRNAs in development, stem cells and cancer," *Trends in Molecular Medicine*, vol. 14, no. 9, pp. 400–409, 2008.

[11] Y. Zeng, B. Yao, J. Shin et al., "Lin28A binds active promoters and recruits Tet1 to regulate gene expression," *Molecular Cell*, vol. 61, no. 1, pp. 153–160, 2016.

[12] C. Ibar and K. D. Irvine, "Integration of hippo-YAP signaling with metabolism," *Developmental Cell*, vol. 54, no. 2, pp. 256–267, 2020.

[13] M. Yang, S. L. Yang, S. Herrlinger et al., "Lin28 promotes the proliferative capacity of neural progenitor cells in brain development," *Development*, vol. 142, no. 9, pp. 1616–1627, 2015.

[14] E. Balzer, C. Heine, Q. Jiang, V. M. Lee, and E. G. Moss, "LIN28 alters cell fate succession and acts independently of the let-7 microRNA during neurogliogenesis in vitro," *Development*, vol. 137, no. 6, pp. 891–900, 2010.

[15] F. Cimadamore, A. Amador-Arjona, C. Chen, C. T. Huang, and A. V. Terskikh, "SOX2-LIN28/let-7 pathway regulates proliferation and neurogenesis in neural precursors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 32, pp. E3017–E3026, 2013.

[16] S. Yamanaka, "Pluripotent stem cell-based cell therapy-promise and challenges," *Cell Stem Cell*, vol. 27, no. 4, pp. 523–531, 2020.

[17] S. J. Sharkis, R. J. Jones, C. Civin, and Y. Y. Jang, "Pluripotent stem cell-based cancer therapy: promise and challenges," *Sci Transl Med*, vol. 4, no. 127, 2012.

[18] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *Cell*, vol. 126, no. 4, pp. 663–676, 2006.

[19] K. Takahashi, K. Tanabe, M. Ohnuki et al., "Induction of pluripotent stem cells from adult human fibroblasts by defined factors," *Cell*, vol. 131, no. 5, pp. 861–872, 2007.

[20] J. A. Thomson, J. Itskovitz-Eldor, S. S. Shapiro et al., "Embryonic stem cell lines derived from human blastocysts," *Science*, vol. 282, no. 5391, pp. 1145–1147, 1998.

[21] P. Li, Y. Chen, X. Meng et al., "Suppression of malignancy by Smad3 in mouse embryonic stem cell formed teratoma," *Stem Cell Reviews and Reports*, vol. 9, no. 5, pp. 709–720, 2013.

[22] P. Li, Y. Chen, K. K. Mak, C. K. Wong, C. C. Wang, and P. Yuan, "Functional role of Mst1/Mst2 in embryonic stem cell differentiation," *PLoS One*, vol. 8, no. 11, article e79867, 2013.

[23] I. Verma, Z. Rashid, S. K. Sikdar, and P. B. Seshagiri, "Efficient neural differentiation of mouse pluripotent stem cells in a serum-free medium and development of a novel strategy for enrichment of neural cells," *International Journal of Developmental Neuroscience*, vol. 61, no. 1, pp. 112–124, 2017.

[24] M. Eiraku, K. Watanabe, M. Matsuo-Takasaki et al., "Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals," *Cell Stem Cell*, vol. 3, no. 5, pp. 519–532, 2008.

[25] P. Li, X. Ma, I. R. Adams, and P. Yuan, "A tight control of *Rif1* by Oct4 and Smad3 is critical for mouse embryonic stem cell stability," *Cell Death & Disease*, vol. 6, no. 1, article e1588, 2015.

[26] F. Dehghanian, Z. Hojati, F. Esmaeili, and A. Masoudi-Nejad, "Network-based expression analyses and experimental validations revealed high co-expression between Yap1 and stem cell markers compared to differentiated cells," *Genomics*, vol. 111, no. 4, pp. 831–839, 2019.

[27] P. Li, M. R. Silvis, Y. Honaker, W. H. Lien, S. T. Arron, and V. Vasioukhin, "αE-catenin inhibits a Src-YAP1 oncogenic module that couples tyrosine kinases and the effector of Hippo signaling pathway," *Genes & Development*, vol. 30, no. 7, pp. 798–811, 2016.

[28] B. Zhao, X. Ye, J. Yu et al., "TEAD mediates YAP-dependent gene induction and growth control," *Genes & Development*, vol. 22, no. 14, pp. 1962–1971, 2008.

[29] L. Chen, P. G. Loh, and H. Song, "Structural and functional insights into the TEAD-YAP complex in the Hippo signaling pathway," *Protein & Cell*, vol. 1, no. 12, pp. 1073–1083, 2010.

[30] A. L. Morgado, C. M. P. Rodrigues, and S. Solá, "MicroRNA-145 regulates neural stem cell differentiation through the Sox2-Lin28/let-7 signaling pathway," *Stem Cells*, vol. 34, no. 5, pp. 1386–1395, 2016.

[31] C. Zhao, G. Sun, S. Li et al., "MicroRNA let-7b regulates neural stem cell proliferation and differentiation by targeting nuclear receptor TLX signaling," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 5, pp. 1876–1881, 2010.

[32] D. Lim, W. G. Byun, J. Y. Koo, H. Park, and S. B. Park, "Discovery of a small-molecule inhibitor of protein-microRNA interaction using binding assay with a site-specifically labeled Lin28," *Journal of the American Chemical Society*, vol. 138, no. 41, pp. 13630–13638, 2016.

[33] L. Wang, R. G. Rowe, A. Jaimes et al., "Small-molecule inhibitors disrupt let-7 oligouridylation and release the selective blockade of let-7 processing by LIN28," *Cell Reports*, vol. 23, no. 10, pp. 3091–3101, 2018.

[34] F. X. Yu and K. L. Guan, "The Hippo pathway: regulators and regulations," *Genes & Development*, vol. 27, no. 4, pp. 355–371, 2013.

[35] F. Reggiani, G. Gobbi, A. Ciarrocchi, and V. Sancisi, "YAP and TAZ are not identical twins," *Trends in Biochemical Sciences*, vol. 46, no. 2, pp. 154–168, 2021.

[36] Z. Wu and K. L. Guan, "Hippo signaling in embryogenesis and development," *Trends in Biochemical Sciences*, vol. 46, no. 1, pp. 51–63, 2021.

[37] J. Yu, M. A. Vodyanik, K. Smuga-Otto et al., "Induced pluripotent stem cell lines derived from human somatic cells," *Science*, vol. 318, no. 5858, pp. 1917–1920, 2007.

[38] I. Lian, J. Kim, H. Okazawa et al., "The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation," *Genes & Development*, vol. 24, no. 11, pp. 1106–1118, 2010.

[39] H. Chung, B. K. Lee, N. Uprety, W. Shen, J. Lee, and J. Kim, "Yap1 is dispensable for self-renewal but required for proper differentiation of mouse embryonic stem (ES) cells," *EMBO Reports*, vol. 17, no. 4, pp. 519–529, 2016.

[40] L. LeBlanc, B. K. Lee, A. C. Yu et al., "Yap1 safeguards mouse embryonic stem cells from excessive apoptosis during differentiation," *Elife*, vol. 7, 2018.

[41] N. Robledinos-Anton, M. Escoll, K. L. Guan, and A. Cuadrado, "TAZ represses the neuronal commitment of neural stem cells," *Cells*, vol. 9, no. 10, p. 2230, 2020.

[42] C. Qiu, Y. Ma, J. Wang, S. Peng, and Y. Huang, "Lin28-mediated post-transcriptional regulation of Oct4 expression in human embryonic stem cells," *Nucleic Acids Research*, vol. 38, no. 4, pp. 1240–1248, 2010.

[43] B. Xu and Y. Huang, "Histone H2a mRNA interacts with Lin28 and contains a Lin28-dependent posttranscriptional regulatory element," *Nucleic Acids Research*, vol. 37, no. 13, pp. 4256–4263, 2009.

[44] B. Xu, K. Zhang, and Y. Huang, "Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells," *RNA*, vol. 15, no. 3, pp. 357–361, 2009.

Hindawi
Stem Cells International
Volume 2021, Article ID 6657597, 10 pages
https://doi.org/10.1155/2021/6657597

Research Article

Discovery of a Novel Long Noncoding RNA *Lx8-SINE B2* as a Marker of Pluripotency

Fuquan Chen, Miao Zhang, Xiao Feng, Xiaomin Li, Haotian Sun, and Xinyi Lu

State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy, Nankai University, Tianjin 300307, China

Correspondence should be addressed to Xinyi Lu; luxy@nankai.edu.cn

Received 17 October 2020; Revised 31 December 2020; Accepted 22 January 2021; Published 8 February 2021

Academic Editor: Qiang Wu

Pluripotency and self-renewal of embryonic stem cells (ESCs) are marked by core transcription regulators such as Oct4, Sox2, and Nanog. Another important marker of pluripotency is the long noncoding RNA (lncRNA). Here, we ind that a novel long noncoding RNA (lncRNA) *Lx8-SINE B2* is a marker of pluripotency. LncRNA *Lx8-SINE B2* is enriched in ESCs and downregulated during ESC differentiation. By rapid amplification of cDNA ends, we identified the full-length sequence of lncRNA *Lx8-SINE B2*. We further showed that transposable elements at upstream of lncRNA *Lx8-SINE B2* could drive the expression of lncRNA *Lx8-SINE B2*. Furthermore, ESC-specific expression of lncRNA *Lx8-SINE B2* was driven by Oct4 and Sox2. In summary, we identified a novel marker lncRNA of ESCs, which is driven by core pluripotency regulators.

1. Introduction

Most of the mammalian genome is composed of noncoding sequences. Among them, transposable elements (TEs) contribute to ~40% of the genome [1]. The majority of TEs are silenced, however, a small percentage of TEs are expressed during development and in diseases [2]. They play multiple roles in these processes, including function as enhancers, promoters, and long noncoding RNAs (lncRNAs) [3–6]. In vertebrates, 70% lncRNAs are composed of TEs [7]. TEs also confer tissue-specific expression on lncRNAs through the recruitment of transcription factors [3, 4, 6]. TE-derived lncRNAs actively participate in development. TE-derived lncRNA ROR functions as a sponge to miRNA and also works with hnRNPA1 to promote c-Myc expression during reprogramming [8–10]. Endogenous retrovirus HERVH-derived lncRNAs maintain pluripotency of human embryonic stem cells [3, 11–13]. Asymmetrical expression of ERV1 and ERVK-derived lncRNA LincGET in two- to four-cell mouse embryos biases cell fate toward inner cell mass [14]. These findings all suggest an important role of TE-derived lncRNA in development. Most of these findings are based on human cell lines. We are still lack of understanding of TE-derived lncRNAs in mouse embryonic stem cells (ESCs). In this study, we investigated the expression and regulation of one representative lncRNA *Lx8-SINE B2* in ESCs.

2. Methods

2.1. Cell Culture. Mouse ESCs (E14) were cultured on plates coated with 0.2% gelatin (G1890, Sigma) in medium with 15% fetal bovine serum (FBS, SH30070.03, Hyclone), 2 mM L-glutamine (Gibco), 1% penicillin-streptomycin (P1400, Solarbio), 0.1 mM nonessential amino acids (Gibco), 0.1 mM β-mercaptoethanol (M3148–250, Sigma), and 10 ng/ml leukemia inhibitory factor (LIF; Z03077, GenScript). Mouse embryonic fibroblasts (MEFs) and 3 -T3 cells were maintained on plates (703001, NEST Biotechnology) in Dulbecco's modified Eagle's medium (DMEM, Hyclone) supplemented with 10% FBS, 2 mM L-glutamine, and 1% penicillin-streptomycin. Cells were cultured at 37°C in CO_2 incubator.

2i culture medium contain 50% DMEM/F12 (BasalMedia), 50% Neurobasal media (Gibco), 1% N2 supplement, 1% B27 (Gibco), 0.1 mM nonessential amino acids (Gibco), 2 mM L-glutamine(Gibco), 1% penicillin-streptomycin (P1400, Solarbio), 0.1 mM β-mercaptoethanol (M3148–250, Sigma), 1 μM MEK inhibitor PD0325901 (T6189, TargetMol), and 3 μM GSK3 inhibitor CHIR99021 (2520691, BioGems). 10 ng/ml

Table 1: Primer sequences for qPCR analysis.

Gene	Forward	Reverse
lncRNA Lx8-SINE B2	GCTGTTATGACTTGTTTCCTGGT	CTCTTCCTTGCAGGCTTAGAAC
Oct4	GTGGAAAGCAACTCAGAGG	GGTTCCACCTTCTCCAACT
Sox2	GCGGAGTGGAAACTTTTGTCC	CGGGAAGCGTGTACTTATCCTT
Nanog	TTGCTTACAAGGGTCTGCTACT	ACTGGTAGAAGAATCAGGGCT
Esrrb	GCACCTGGGCTCTAGTTGC	TACAGTCCTCGTAGCTCTTGC
Prdm14	CTCTTGATGCTTTTCGGATGACT	GTGACAATTTGTACCAGGGCA
Lysmd3	ACGGTTTCCCTCCCAGGAAT	CATCAAGTCTATCTCTCGATGCG
Adgrv1	CAGCCCTGAATCACTCTTCGT	CCCATCCAGGTCCGAGTCTA
LINE1	GGACCAGAAAAGAAATTCCTCCCG	CTCTTCTGGCTTTCATAGTCTCTGG
SINE B2	GAGTAAGAGCACCCGACTGC	AGAAGAGGGAGTCAGATCTCGT

leukemia inhibitory factor (LIF; Z03077, GenScript) was added for 2i/LIF condition.

2.2. RNA Extraction, Reverse Transcription, and Quantitative PCR (qPCR). Total RNA was extracted with RNAiso Reagent (B9109, Takara) as described [15] and treated with DNase I to remove genomic DNA in DEPC water (B501005, Sangon Biotech). The cDNA synthesis was carried out in RNase-free tubes (401001, NEST Biotechnology) with the Transcriptor First Strand cDNA Synthesis Kit (4897030001, Roche), according to the manufacturer's instructions. Quantitative PCR (qPCR) reactions were performed using the Hieff qPCR SYBR Green Master Mix (H97410, Yeasen) in a QuantStudio 6 Real-Time PCR System (Life Technologies). Primer sequences for qPCR analysis are listed in Table 1.

2.3. Depletion of Gene Expression with shRNAs. For gene knockdown, short hairpin RNAs (shRNAs) for luciferase (control) or target genes were designed by an online tool (http://sirna.wi.mit.edu/) and synthesized by GENEWIZ corporation. The shRNA plasmids were constructed using the pSuper-puro system and purified with a kit (1211-01, Biomiga). mESCs were transfected with DNA using Polyjet (SL100688, SignaGen), according to the manufacturer's protocol. Transfected ESCs were selected with 1 μg/ml puromycin from 24 h after transfection. After four days of puromycin selection, transfected cells were harvested. The sequences of shRNAs are listed in Table 2.

2.4. 5′ and 3′ Rapid Amplification of cDNA Ends (RACE) Analysis. For 3′ RACE, first-strand cDNA synthesis is initiated at the poly(A) tail of total RNA using the anneal oligo(dT)-containing RT Adapter Primer (AP) to mRNA. Gene-specific primer pF1 was designed based on the known sequence. 3′ fragment was amplified by primer pF1 and general primer gR1, the RACE PCR products were separated on a 1.5% agarose gel.

For 5′ RACE, the first-strand cDNA was synthesized from total RNA using a gene-specific primer (RT GSP1), which was designed according to the 3′ known sequence. A homopolymer tail was subsequently added to the 3′-end of the cDNA using terminal deoxynucleotidyl transferase kit (2230A, Takara), according to the manufacturer's instruction. First-round PCR was performed based on poly(C) tail designed dG adaptor primer to synthesize double-stranded cDNA. Then, general primer gP1 and gene-specific primer pR2 were used for second-round PCR to amplify the cDNA 5′ end sequence. The RACE PCR products were separated on a 1.5% agarose gel and cloned into pEASY-T1 (TransGen Biotech) for Sanger sequencing. The gene-specific RACE primers used for mapping each end were from Sangon Biotech and were listed in Table 3.

Table 2: Targeting sequences of shRNAs.

Gene	shRNA target sequence
Oct4 shRNA	GTGGAAAGCAACTCAGAGG
Sox2 shRNA	GCGGAGTGGAAACTTTTGTCC
Nanog shRNA	TTGCTTACAAGGGTCTGCTACT

Table 3: RACE primers.

RACE primer name	Sequence
RT-adaptor primer	GCGAGCACAGAATTAATACGACTC ACTATAGG(T)18VN
gR1	GCGAGCACAGAATTAATACGAC
pF1	ATACCTTCCTAAAACTAATGTGGACT
RT GSP1	TGAAGAACTTTTAGCACAGCAGC
dG-adaptor primer	GACTCGAGTCGACATCGAGGGGGG GGGGGGGGGGG
gP1	GACTCGAGTCGACATCG
pR2	CAACTGTTCTAAACGCTTCTTAG

2.5. Dual-Luciferase Reporter Gene Assay. Mouse ESCs were seeded at a density of 8×10^4 cells per well in a 24-well plate. Luciferase assay was performed as previously described [16]. The total amount of 200 ng of the various promoters of lncRNA *Lx8-SINE B2* or pGL4.23 empty vector was transfected into each well of E14 ESC on a 24-well plate together with 10 ng of pCMV-Renilla. The medium was changed 12 h after transfection. After transfection of 36 h, cells were collected and lysed in 1x passive lysis buffer. The luciferase activity was determined by Dual-Luciferase Reporter Assay

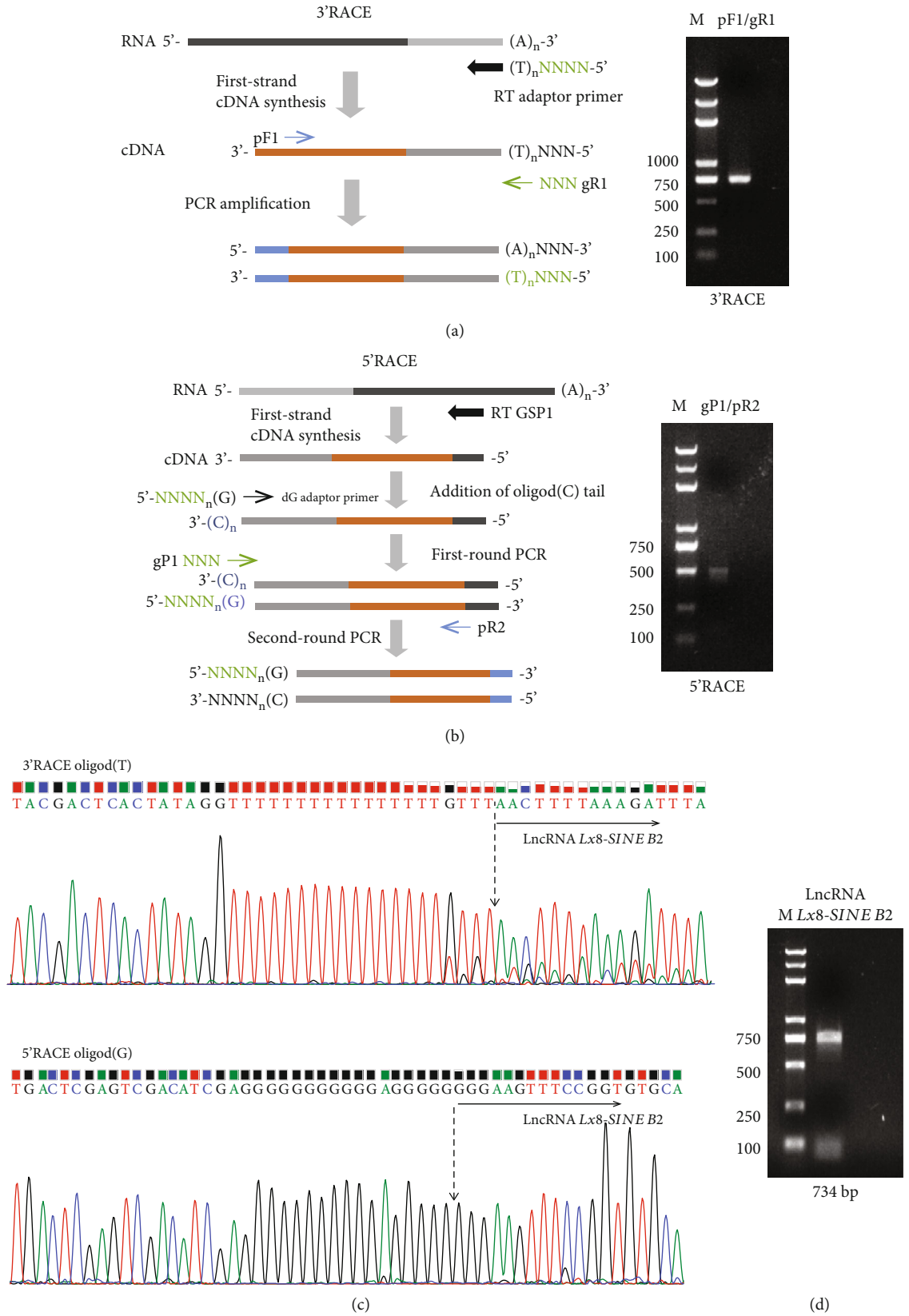


Figure 1: Mapping the full-length sequence of lncRNA *Lx8-SINE B2*. (a) Schematic of the 3′-rapid amplification of cDNA ends (RACE) (left) and 3′ RACE result for lncRNA *Lx8-SINE B2* (right). (b) Schematic of the 5′ RACE and its result for lncRNA *Lx8-SINE B2*. (c) DNA sequencing of RACE using a universal primer in pEASY-T1 vector. (d) Validation of lncRNA *Lx8-SINE B2* transcript size by PCR from cDNA. M, DNA marker.

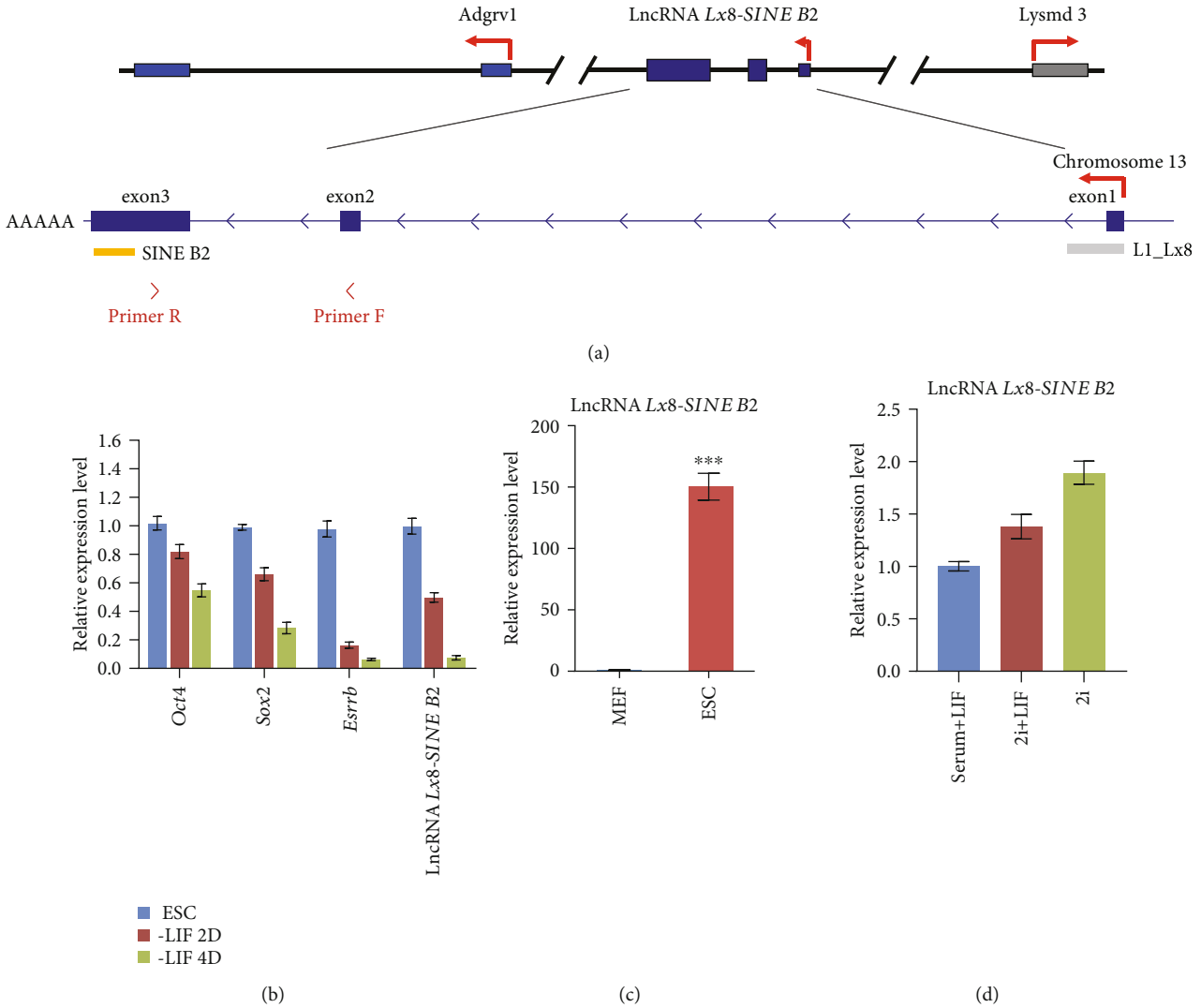


Figure 2: Genomic location and expression of lncRNA *Lx8-SINE B2*. (a) Schematic of the mouse lncRNA *Lx8-SINE B2* locus on chromosome 13. lncRNA *Lx8-SINE B2* is between *Adgrv1* and *Lysmd3*. There are three exons and some retrotransposon fragments of LINE or SINE in lncRNA *Lx8-SINE B2*. RT–qPCR primers were indicated below. (b) The expression level of lncRNA *Lx8-SINE B2*, *Oct4 Sox2*, and *Esrrb* in mESCs in the presence or absence of LIF for 2-4 days, as measured by RT–qPCR and normalized to *Gapdh* levels. Biological-triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. (c) qPCR analysis of the expression level of lncRNA *Lx8-SINE B2* in mouse ESCs and MEF cells. $^{***}p < 0.001$ according to two-sided Student's *t*-test (d) Expression analysis of lncRNA *Lx8-SINE B2* in ESCs cultured under serum/LIF, 2i/LIF or 2i condition. Biological-triplicate data ($n = 3$ extracts) are presented as mean ± s.e.m.

System (#E1910, Promega) according to the manufacturer's instructions.

2.6. Statistical Analysis. Data were analyzed with Student's *t* -test (two-tailed). Significant differences were defined as ns for nonsignificant, $^{**}p < 0.01$, and $^{***}p < 0.001$.

3. Results

3.1. Mapping the Full-Length Sequence of lncRNA Lx8-SINE B2. Through mining the previous publication [17], it was shown that lincRNA-1282 was expressed in ESCs and its depletion leads to downregulation of c-Myc [17], which is an important reprogramming factor. Therefore, we set out to perform RACE to identify the full-length of lincRNA-1282 [17], which is a partial sequence of lncRNA *Lx8-SINE B2*. To identify the full length of *Lx8-SINE B2*, we performed 3′ RACE and 5′ RACE with primers as designed (Figures 1(a) and 1(b)). Our amplicons for both 5′ and 3′ RACE were visible as a single DNA band without multiple or unspecific bands (Figures 1(a) and 1(b)). Next, we sequenced the amplicons and identified the sequences of lncRNA *Lx8-SINE B2* (Figure 1(c)). With the 5′ and 3′ ends of lncRNA *Lx8-SINE B2* found, we designed primers to amplify the full length of lncRNA *Lx8-SINE B2* and subcloned the lncRNA into TA cloning vector (Figure 1(d)). The lncRNA *Lx8-SINE B2* was revealed to be a 734 bp lncRNA.

3.2. Expression Pattern of lncRNA Lx8-SINE B2. We searched the sequences of lncRNA *Lx8-SINE B2* against the mouse

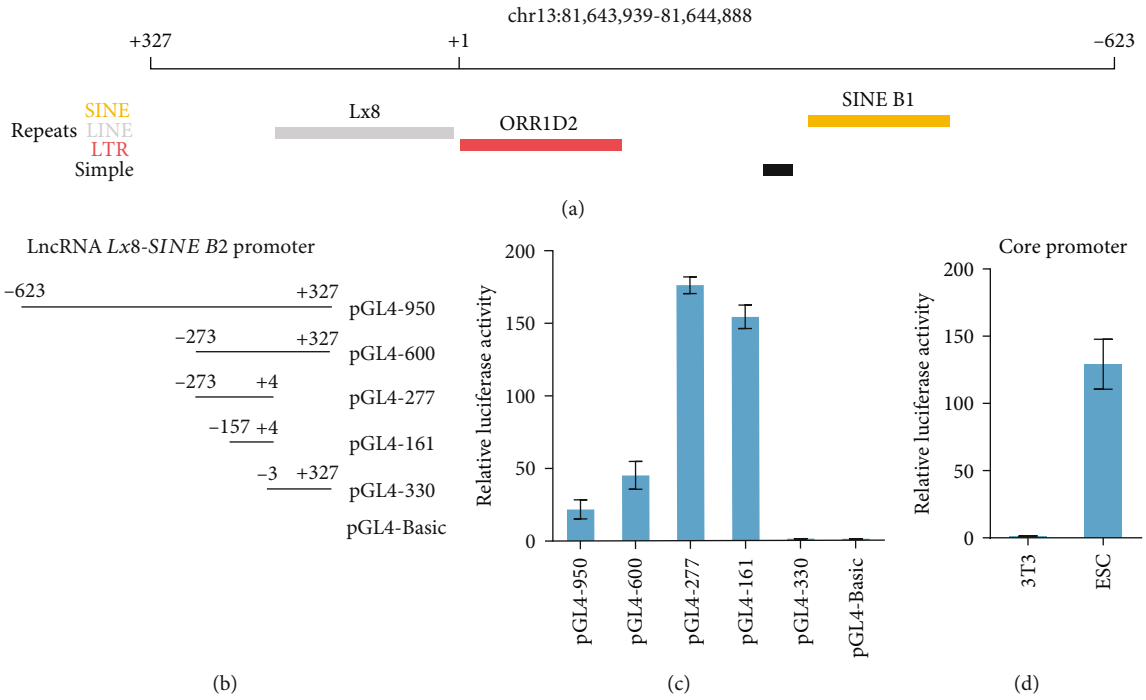


FIGURE 3: Dissection of lncRNA *Lx8-SINE B2* promoter region. (a) Schematic of the positions of TEs on according to mouse mm10 genome the promoter region of lncRNA *Lx8-SINE B2*. (b) Schematic of various length fragments of lncRNA *Lx8-SINE B2* promoter constructs. (c) Activities of various length fragments of lncRNA *Lx8-SINE B2* promoter constructs were determined by *luciferase* reporter gene assays in E14 ESCs. Biological-triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. (d) Luciferase assay analysis of core promoter activity of lncRNA *Lx8-SINE B2* in ESCs and 3T3 cells. Biological-triplicate data (n =3 extracts) are presented as mean ± s.e.m.

genome (mm10) and discovered that lncRNA *Lx8-SINE B2* contained 3 exons, which are located between *Adgrv1* and *Lysmd3* gene (Figure 2(a)). Exon 1 of lncRNA *Lx8-SINE B2* overlapped with LINE1 family Lx8 and its third exon overlapped with SINE B2 (Figure 2(a)); therefore, we named this lncRNA as *Lx8-SINE B2*. We designed primers on the nonrepeat region of exon 2 and 3 to detect the expression of lncRNA *Lx8-SINE B2*. Interestingly, it is noticed that lncRNA *Lx8-SINE B2*was downregulated during ESC differentiation, similar to the pluripotency gene *Oct4*, *Sox2*, and *Esrrb*, according to qPCR results (Figure 2(b)). We also found that lncRNA *Lx8-SINE B2* was also expressed in ESCs instead of differentiated cells such as MEF (Figure 2(c)). Furthermore, we demonstrated that the expression of lncRNA *Lx8-SINE B2*was not affected by the alternation of ESC culture condition. Its expression was slightly upregulated in the presence of 2i/LIF or 2i condition in contrast to the serum/LIF culture condition (Figure 2(d)). These suggest lncRNA *Lx8-SINE B2* as a marker of ESC.

3.3. Promoter Structure of lncRNA Lx8-SINE B2. After that, we examined how the specific expression of lncRNA *Lx8-SINE B2* was achieved. The upstream 1 kb promoter region of lncRNA *Lx8-SINE B2* contains ORR1D2 and SINE B1 (Figure 3(a)). To study how *Lx8-SINE B2* is regulated in ESCs, we cloned -623 bp to +327 bp of lncRNA *Lx8-SINE B2* gene into luciferase reporter (Figures 3(a) and 3(b)). We also created various truncation versions of this region to identify the core promoter of *Lx8-SINE B2* (Figure 3(b)). The region corresponding to ERV, origin-region repeat 1 type D2 (ORR1D2, -157 bp to +3 bp) carried the strongest promoter activity in contrast to those of other truncations (Figure 3(c)). The promoter activity of ORR1D2 was specific to ESCs but inactivated in 3T3 fibroblasts (Figure 3(d)). These results support that lncRNA *Lx8-SINE B2* is driven by ERV ORR1D2, implicating that TEs not only contribute to the exons of lncRNAs but also the promoter of lncRNAs.

3.4. Transcriptional Regulation of lncRNA Lx8-SINE B2 by Oct4 and Sox2. To identify which transcription factor activates lncRNA *Lx8-SINE B2*, we depleted three core pluripotency transcription factors (Oct4, Sox2, and Nanog) (Figures 4(a)–4(c)). Depletion of *Oct4* or *Sox2*, but not *Nanog*, strongly suppressed lncRNA *Lx8-SINE B2* expression (Figures 4(a)–4(c)). We also examined the expression of lncRNA *Lx8-SINE B2* after the depletion of *Oct4*, *Sox2*, and *Nanog* (Figures 4(a)–4(c)). However, depletion of either *Sox2* or *Oct4*, but not *Nanog*, affected the promoter activity of ORR1D2 (Figures 4(d)–4(f)). *Sox2* depletion imposed stronger inhibition on ORR1D2 than Oct4 and Nanog (Figures 4(d)–4(f)). Furthermore, we examined the binding of Oct4, Sox2, and Nanog on the promoter of lncRNA *Lx8-SINE B2*. Consistent with results from luciferase assay, only Oct4 and Sox2 bound to the promoter according to our analysis of published ChIP-seq data (Figure 4(g)). These results suggest that Sox2 and Oct4 directly bind to ORR1D2 to activate *Lx8-SINE B2* in ESCs (Figure 4(g)).

To exclude the possibility that Oct4 and Sox2 activate neighboring genes of lncRNA *Lx8-SINE B2* together with it, we examined the expression of *Lysmd3* and *Adgrv1* during

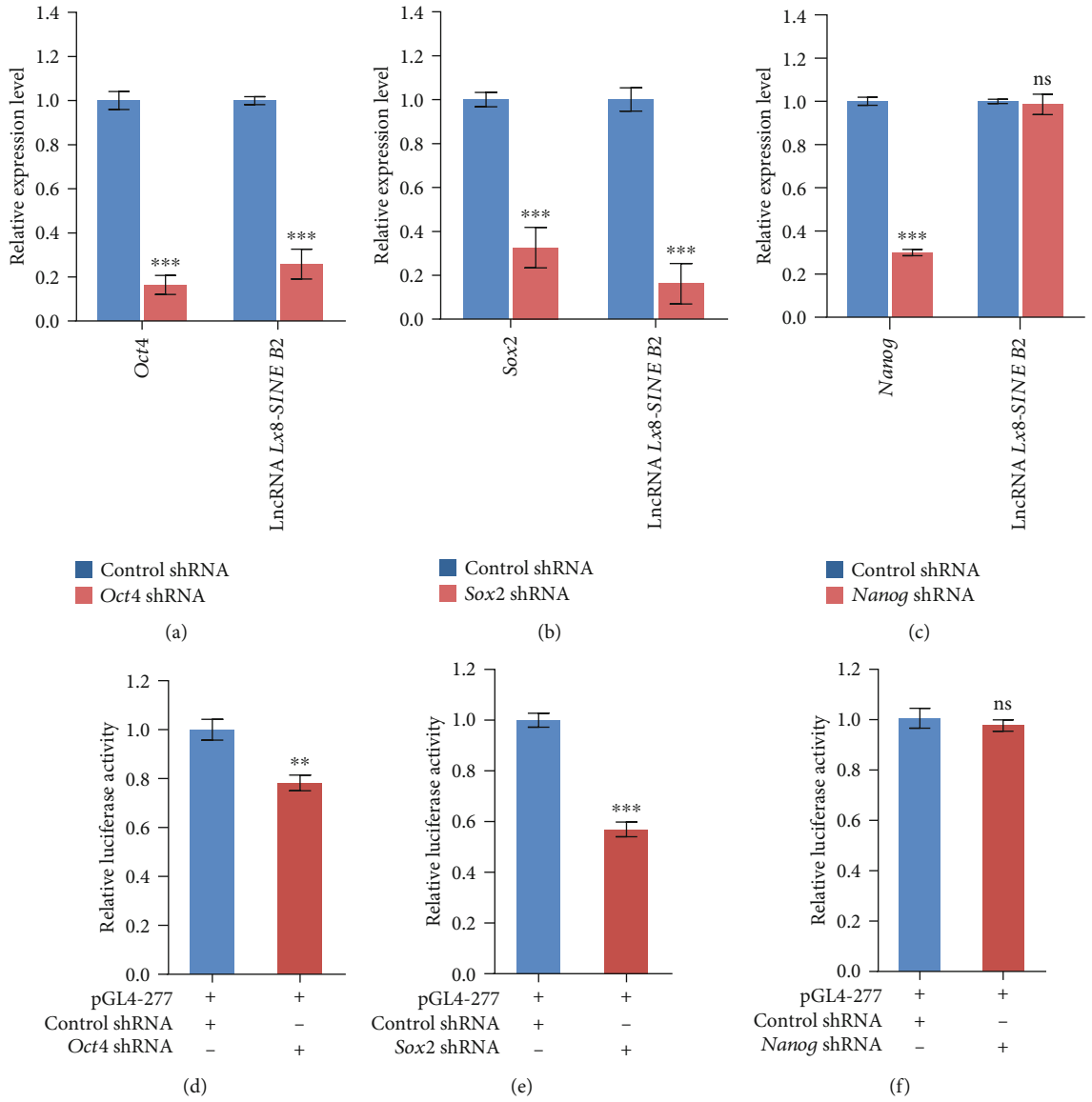


Figure 4: Continued.

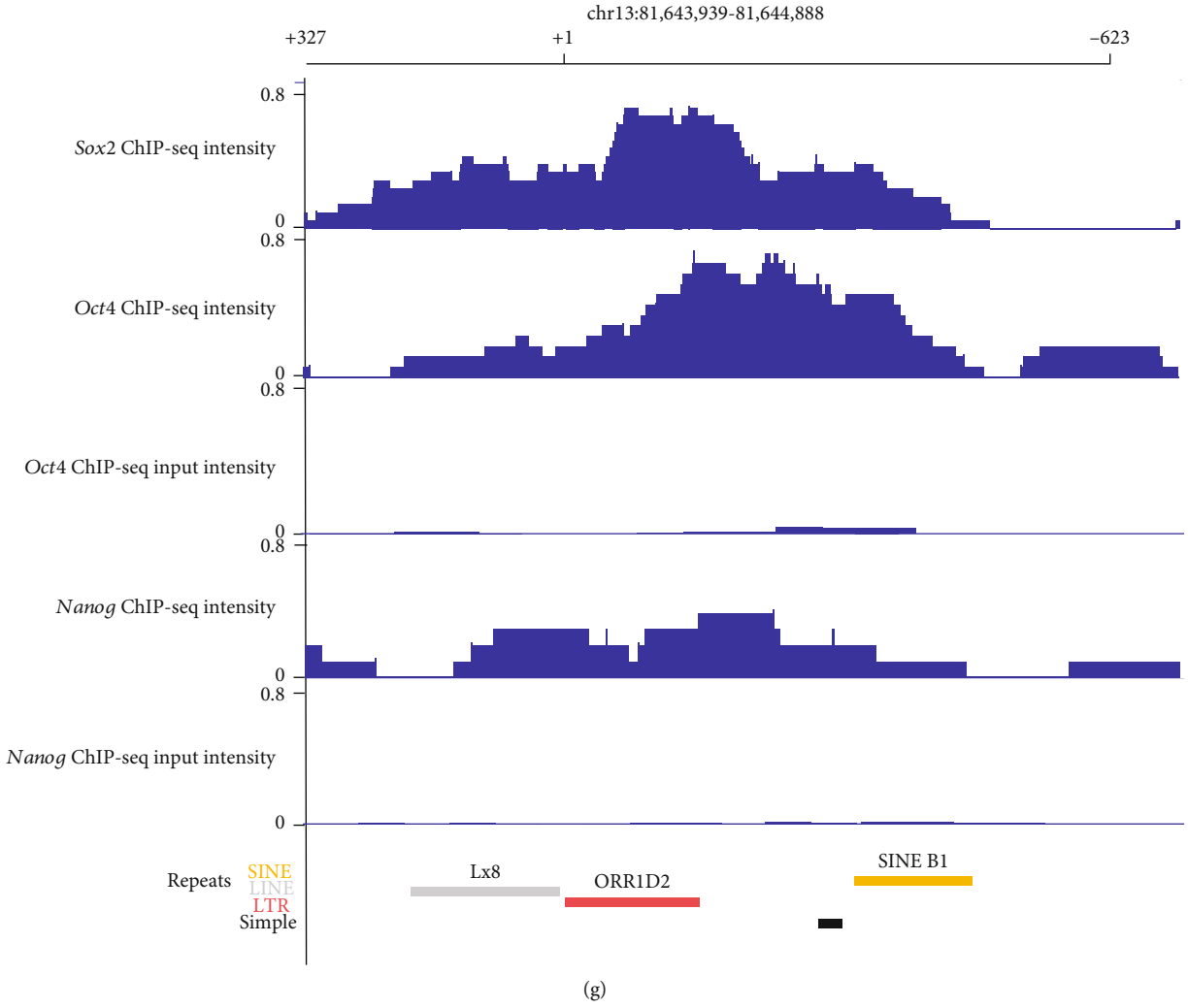


Figure 4: Oct4 and Sox2 drive the expression of lncRNA *Lx8-SINE B2*. (a–c) qPCR analysis of lncRNA *Lx8-SINE B2* expression after depletion of core transcription factors *Oct4* (a), *Sox2* (b), and *Nanog* (c) in ESCs. The data are represented as mean ± s.e.m. from three biological replicates. ns, non-significant; $^{**}p < 0.01$; $^{***}p < 0.001$ according to two-sided Student's *t*-test. Biological-triplicate data ($n = 3$ dishes). (d–f) Luciferase assay analysis of core promoter activity of lncRNA *Lx8-SINE B2* after depletion of core transcription factors *Oct4* (d), *Sox2* (e), and *Nanog* (f) in ESCs. Biological-triplicate data ($n = 3$ extracts) are presented as mean ± s.e.m. (g) Binding profile of Sox2, Oct4, and Nanog on the promoter region of lncRNA *Lx8-SINE B2* according to published data as described in methods. Input was included as a control.

ESC differentiation. Different from lncRNA *Lx8-SINE B2*, both *Lysmd3* and *Adgrv1* were unaffected by LIF withdrawal (Figure 5(a)). Furthermore, the expression of *Lysmd3* and *Adgrv1* were activated by depletion of *Oct4* or *Sox2*, suggesting they are regulated differently from *Lx8-SINE B2* (Figures 5(b) and 5(c)). Moreover, the expression of LINE1 and SINE B2 were not affected by *Oct4* or *Sox2* depletion (Figure 5(d)), confirming the specificity of Oct4 and Sox2 in activating the expression of lncRNA *Lx8-SINE B2*.

4. Discussion

In summary, we identified a novel pluripotency marker lncRNA *Lx8-SINE B2*, whose expression is driven by the binding of Oct4 and Sox2 on ORR1D2. Oct4 and Sox2 are the core pluripotency regulators in ESCs [18, 19]. Oct4 and Sox2 can drive the expression of lncRNAs in cancer cells and ESCs [20–22]. The binding profiles of OCT4 are different in human and mouse ESCs [23], which can be explained by its binding differences on species-specific TEs [23]. Here, we found that Oct4 and Sox2 targeted mouse TE ORR1D2 to drive ESC-specific lncRNA expression (Figure 4), further supporting the important role of TEs in driving the expression of species-specific lncRNAs. There are many pluripotency markers; however, we provide *Lx8-SINE B2* as an additional novel marker of pluripotency. It lies at the downstream of key pluripotency genes *Oct4* and *Sox2* (Figure 4). It composes of TEs and is distinct from traditional markers of pluripotency. In comparison to other ESC markers, *Lx8-SINE B2* is unique as an ORR1D2-driven pluripotency marker, which demonstrates that transposable elements can function as cell type-specific lncRNA and promoter, similarly

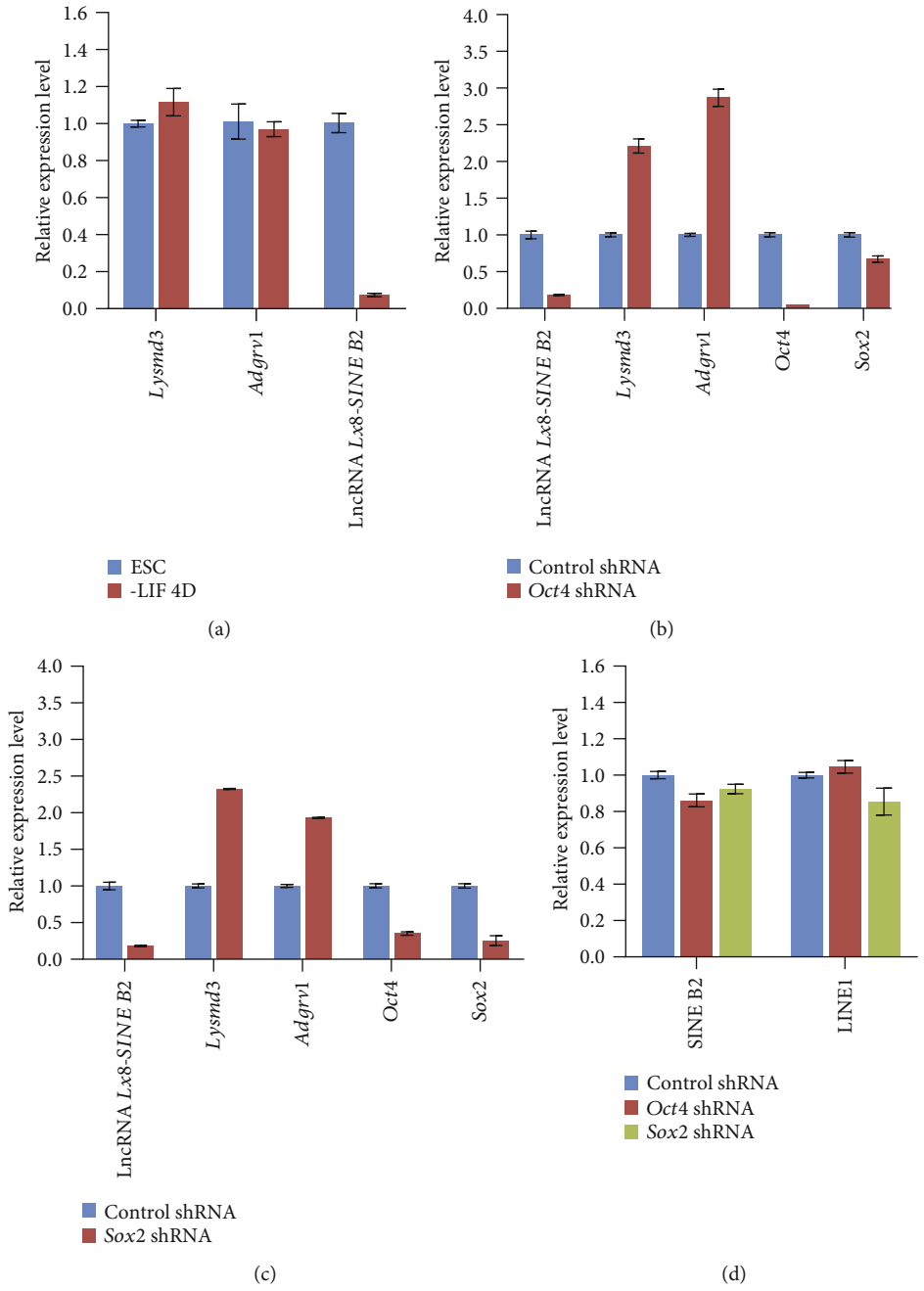


Figure 5: Expression of neighboring genes of *Lx8-SINE B2*. (a) qPCR analysis of neighboring genes (*Adgrv1* and *Lysmd3*) of lncRNA *Lx8-SINE B2* in ESCs cultured with or without LIF; (b, c) RT-qPCR analysis of lncRNA *Lx8-SINE B2*, *Adgrv1*, *Lysmd3* and pluripotent genes (*Oct4* and *Sox2*) expression after depletion of *Oct4* (b) or *Sox2* (c) in E14 ESCs. The data are represented as mean ± s.e.m. from three biological replicates. (d) Expression of LINE1 and SINE B2 after depletion of *Oct4* or *Sox2*.

to protein-coding genes. Finally, its depletion is associated with the downregulation of *Myc* in ESCs [17]; therefore, *Lx8-SINE B2* expression also reflects *Myc* expression status of ESCs. Myc represses primitive endoderm differentiation [24]. Myc also maintains ESC pluripotency and self-renewal [25]. Therefore, we speculate that the depletion of lncRNA *Lx8-SINE B2* may cause a phenotype similar to that of Myc downregulation.

Our study demonstrates that different types of TEs combine to form lncRNA and drive lncRNA expression (Figures 2 and 3), implicating TEs as important components of lncRNA. TEs in lncRNAs work as an important RNA domain [26, 27]. TEs within lncRNAs regulate the tissue-specific expression of lncRNAs [4, 28]. In human, lncRNAs containing HERVH are specifically expressed in human ESCs [3, 4, 7]. TEs within lncRNAs also contribute to their functions. For example, SINE B2 in antisense lncRNA of *Uchl1* interacts with *Uchl1* mRNA and promotes the translation of Uchl1 through enhancing the association of mRNA with polysome [29]. These studies demonstrate that TEs are

critical to the expression and function of lncRNAs. Given that lncRNA *Lx8-SINE B2* is composed of TE Lx8 and SINE B2, it will be interesting to investigate whether ORR1D2 drive the expression of other lncRNAs and the function of Lx8 and SINE B2 within lncRNAs in the future study.

5. Conclusion

In conclusion, we mapped the full-length sequence of lncRNA *Lx8-SINE B2* and found it as an ESC-specific lncRNA. We also found that it was driven by ORR1D2 which was bound by Sox2 and Oct4 to drive its transcription. These findings support TEs as important compositions and promoter of lncRNA.

Data Availability

Published ChIP data analyzed by Cistrome [30] in this study are GSE54103 for Sox2 [31], GSE78073 for Oct4 [32], and GSE56312 for Nanog [33].

Conflicts of Interest

We declare that there is no conflict of interest present for this study.

Authors' Contributions

X. Lu conceived and designed the study. F.C., M.Z., and X. Li performed most experiments. X.F. and H.S. did bioinformatics analysis. F.C., M.Z., and X. Lu wrote the manuscript.

Acknowledgments

This work is supported by grants from the National Natural Science Foundation of China (31871488, 32070858), the National Key Research and Development Program of China (2018YFA0107000), and Fundamental Research Funds for the Central Universities (63151113, 63161150). We thank Ruiqing Chen for his technical assistance.

References

[1] R. A. Weiss, "Exchange of genetic sequences between viruses and hosts," *Current Topics in Microbiology and Immunology*, vol. 407, pp. 1–29, 2017.

[2] K. A. O'Donnell and K. H. Burns, "Mobilizing diversity: transposable element insertions in genetic variation and disease," *Mobile DNA*, vol. 1, no. 1, p. 21, 2010.

[3] X. Lu, F. Sachs, L. A. Ramsay et al., "The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity," *Nature Structural & Molecular Biology*, vol. 21, no. 4, pp. 423–425, 2014.

[4] D. Kelley and J. Rinn, "Transposable elements reveal a stem cell-specific class of long noncoding RNAs," *Genome Biology*, vol. 13, no. 11, p. R107, 2012.

[5] C. D. Todd, Ö. Deniz, D. Taylor, and M. R. Branco, "Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells," *eLife*, vol. 8, 2019.

[6] B. Miao, S. Fu, C. Lyu, P. Gontarz, T. Wang, and B. Zhang, "Tissue-specific usage of transposable element-derived promoters in mouse development," *Genome Biology*, vol. 21, no. 1, p. 255, 2020.

[7] A. Kapusta, Z. Kronenberg, V. J. Lynch et al., "Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs," *PLoS Genetics*, vol. 9, no. 4, article e1003470, 2013.

[8] S. Loewer, M. N. Cabili, M. Guttman et al., "Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells," *Nature Genetics*, vol. 42, no. 12, pp. 1113–1117, 2010.

[9] J. Huang, A. Zhang, T. T. Ho et al., "Linc-RoR promotes c-Myc expression through hnRNP I and AUF1," *Nucleic Acids Research*, vol. 44, no. 7, pp. 3059–3069, 2016.

[10] Y. Wang, Z. Xu, J. Jiang et al., "Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal," *Developmental Cell*, vol. 25, no. 1, pp. 69–80, 2013.

[11] J. Durruthy-Durruthy, V. Sebastiano, M. Wossidlo et al., "The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming," *Nature Genetics*, vol. 48, no. 1, pp. 44–52, 2016.

[12] J. He, X. Fu, M. Zhang et al., "Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells," *Nature Communications*, vol. 10, no. 1, p. 34, 2019.

[13] J. Wang, G. Xie, M. Singh et al., "Primate-specific endogenous retrovirus-driven transcription defines naive- like stem cells," *Nature*, vol. 516, no. 7531, pp. 405–409, 2014.

[14] J. Wang, L. Wang, G. Feng et al., "Asymmetric expression of LincGET biases cell fate in two-cell mouse embryos," *Cell*, vol. 175, no. 7, pp. 1887–1901.e18, 2018, e18.

[15] E. Fu, J. Shen, Z. Dong et al., "Histone demethylase Kdm2a regulates germ cell genes and endogenous retroviruses in embryonic stem cells," *Epigenomics*, vol. 11, no. 7, pp. 751–766, 2019.

[16] F. Chen, W. Zhang, D. Xie, T. Gao, Z. Dong, and X. Lu, "Histone chaperone FACT represses retrotransposon MERVL and MERVL-derived cryptic promoters," *Nucleic Acids Research*, vol. 48, no. 18, pp. 10211–10225, 2020.

[17] M. Guttman, J. Donaghey, B. W. Carey et al., "lincRNAs act in the circuitry controlling pluripotency and differentiation," *Nature*, vol. 477, no. 7364, pp. 295–300, 2011.

[18] L. A. Boyer, T. I. Lee, M. F. Cole et al., "Core transcriptional regulatory circuitry in human embryonic stem cells," *Cell*, vol. 122, no. 6, pp. 947–956, 2005.

[19] X. Chen, H. Xu, P. Yuan et al., "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells," *Cell*, vol. 133, no. 6, pp. 1106–1117, 2008.

[20] J. Jen, Y. A. Tang, Y. H. Lu, C. C. Lin, W. W. Lai, and Y. C. Wang, "Oct4 transcriptionally regulates the expression of long non-coding RNAs NEAT1 and MALAT1 to promote lung cancer progression," *Molecular Cancer*, vol. 16, no. 1, p. 104, 2017.

[21] J. Sheik Mohamed, P. M. Gaughwin, B. Lim, P. Robson, and L. Lipovich, "Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells," *RNA*, vol. 16, no. 2, pp. 324–337, 2010.

[22] W. Gao, C. Q. Qi, M. G. Feng, P. Yang, L. Liu, and S. H. Sun, "SOX2-induced upregulation of lncRNA LINC01561 promotes non-small-cell lung carcinoma progression by sponging miR-760 to modulate SHCBP1 expression," *Journal of Cellular Physiology*, vol. 235, no. 10, pp. 6684–6696, 2020.

[23] G. Kunarso, N. Y. Chia, J. Jeyakani et al., "Transposable elements have rewired the core regulatory network of human embryonic stem cells," *Nature Genetics*, vol. 42, no. 7, pp. 631–634, 2010.

[24] K. N. Smith, A. M. Singh, and S. Dalton, "Myc represses primitive endoderm differentiation in pluripotent stem cells," *Cell Stem Cell*, vol. 7, no. 3, pp. 343–354, 2010.

[25] N. V. Varlakhanova, R. F. Cotterman, W. N. deVries et al., "Myc maintains embryonic stem cell pluripotency and self-renewal," *Differentiation; Research in Biological Diversity*, vol. 80, no. 1, pp. 9–19, 2010.

[26] V. Fort, G. Khelifi, and S. M. I. Hussein, "Long non-coding RNAs and transposable elements: A functional relationship," *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, vol. 1868, no. 1, article 118837, 2021.

[27] A. P. Hutchins and D. Pei, "Transposable elements at the center of the crossroads between embryogenesis, embryonic stem cells, reprogramming, and long non-coding RNAs," *Science Bulletin*, vol. 60, no. 20, pp. 1722–1733, 2015.

[28] T. Chishima, J. Iwakiri, and M. Hamada, "Identification of transposable elements contributing to tissue-specific expression of long non-coding RNAs," *Genes*, vol. 9, no. 1, 2018.

[29] C. Carrieri, L. Cimatti, M. Biagioli et al., "Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat," *Nature*, vol. 491, no. 7424, pp. 454–457, 2012.

[30] T. Liu, J. A. Ortiz, L. Taing et al., "Cistrome: an integrative platform for transcriptional regulation studies," *Genome Biology*, vol. 12, no. 8, p. R83, 2011.

[31] J. Chen, Z. Zhang, L. Li et al., "Single-molecule dynamics of enhanceosome assembly in embryonic stem cells," *Cell*, vol. 156, no. 6, pp. 1274–1285, 2014.

[32] J. Shin, T. W. Kim, H. Kim et al., "Aurkb/PP1-mediated resetting of Oct4 during the cell cycle determines the identity of embryonic stem cells," *eLife*, vol. 5, article e10877, 2016.

[33] C. Galonska, M. J. Ziller, R. Karnik, and A. Meissner, "Ground state conditions induce rapid reorganization of core pluripotency factor binding before global epigenetic reprogramming," *Cell Stem Cell*, vol. 17, no. 4, pp. 462–470, 2015.

Hindawi
Stem Cells International
Volume 2021, Article ID 6681337, 13 pages
https://doi.org/10.1155/2021/6681337

Review Article

Advance in the Role of Epigenetic Reprogramming in Somatic Cell Nuclear Transfer-Mediated Embryonic Development

Xiaolei Zhang,[1,2] **Shaorong Gao**,[1,2,3] **and Xiaoyu Liu**[1]

[1]*Institute for Regenerative Medicine, Shanghai East Hospital, Shanghai Key Laboratory of Signaling and Disease Research, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China*
[2]*Clinical and Translational Research Center of Shanghai First Maternity and Infant Hospital, Shanghai Key Laboratory of Signaling and Disease Research, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China*
[3]*Tsingtao Advanced Research Institute, Tongji University, Qingdao 266071, China*

Correspondence should be addressed to Shaorong Gao; gaoshaorong@tongji.edu.cn and Xiaoyu Liu; liuxiaoyu@tongji.edu.cn

Received 30 October 2020; Revised 6 January 2021; Accepted 25 January 2021; Published 5 February 2021

Academic Editor: Qiang Wu

Somatic cell nuclear transfer (SCNT) enables terminally differentiated somatic cells to gain totipotency. Many species are successfully cloned up to date, including nonhuman primate. With this technology, not only the protection of endangered animals but also human therapeutics is going to be a reality. However, the low efficiency of the SCNT-mediated reprogramming and the defects of extraembryonic tissues as well as abnormalities of cloned individuals limit the application of reproductive cloning on animals. Also, due to the scarcity of human oocytes, low efficiency of blastocyst development and embryonic stem cell line derivation from nuclear transfer embryo (ntESCs), it is far away from the application of this technology on human therapeutics to date. In recent years, multiple epigenetic barriers are reported, which gives us clues to improve reprogramming efficiency. Here, we reviewed the reprogramming process and reprogramming defects of several important epigenetic marks and highlighted epigenetic barriers that may lead to the aberrant reprogramming. Finally, we give our insights into improving the efficiency and quality of SCNT-mediated reprogramming.

1. Introduction

Somatic cell nuclear transfer (SCNT), first demonstrated by Gurdon in 1962 [1], is a technology to form reconstructed embryos by injecting donor nucleus into enucleated oocytes and generate cloned animals. The success of SCNT makes the transition from terminally differentiated cells to totipotent cells a reality [2]. It has been about two decades that the first cloned mammal, "Dolly," the sheep, was born [3]. Since then, investigations on SCNT and cloned animals boomed, and different species were successfully cloned by various donor cell types [4–6]. In 2018, the first nonhuman primate species has been cloned by using fetal fibroblasts as donor cells [7]. Besides animal cloning, SCNT technology is widely used to acquire nuclear transfer embryonic stem cells (ntESCs), which is called therapeutic cloning [8–10]. The derivation of human ntESCs, which was first achieved at 2013 [11] and further improved in the following years [12–14], implies SCNT technology holds great application prospects in human therapeutics.

Although successful, low efficiency (Table 1) as well as defects in extraembryonic tissues and cloned individuals in many species impedes the application of SCNT technology, which has been fully reviewed [4, 15, 16]. SCNT embryos are often arrested at the early stages of preimplantation development. For the most used animal model, mouse, SCNT embryos are usually arrested at 2-cell and 4-cell stages [17, 18]. Even if the embryos develop to blastocyst stage, postimplantation defects and abnormal placentas, like enlarged placenta, were still observed [19]. Only about 1-2% of reconstructed embryos enable to develop to term [4, 20]. For other species, the highest cloning efficiency was demonstrated in bovine, which is about 5-20%, still much lower than that of IVF (about 40-60%) [20]. Even after born,

Table 1: Cloning efficiency of inner-species SCNT-mediated reprogramming.

Species	Donor cell type	Total oocytes	Reconstructed oocytes	Cleaved embryo/rate	Blastocyst number/rate of cleaved embryo	Transferred embryo number	Birth pups	Birth rate of cleaved embryo	Birth rate of transferred embryo	References
Sheep	Adult mammary epithelium	—	277	247	29	29	1	0.40%	3.45%	[3]
Cow	Fetal fibroblasts (transgenic)	—	276	—	33/-	28	4	—	14.29%	[113]
Cow	Oviductal cells	150	88	77	20/25.97%	4	3	3.90%	75%	[114]
Cow	Adult cumulus	99	37	31	18/58.06%	6	5	16.13%	83.33%	[114]
Mouse	Adult cumulus cells	—	136	45	—	45 (transferred with 2-cell embryos)	7 (two died at day 6-7)	15.56%	15.56%	[115]
Goat	Fetal fibroblasts (transgenic)	—	138	48	—	47 (transferred with cleaved embryos)	1	2.13%	2.08%	[116]
Pig	Fetal fibroblasts	210	188	110	—	110 (2- and 8-cell stage embryos were transferred)	1	0.91%	0.91%	[117]
Pig	Granulosa cells	245	74	—	—	72	5	—	6.94%	[118]
Rabbit	Adult transgenic cumulus cells	—	775	—	—	371 (transferred with 4-cell stage embryos)	6	—	1.62%	[119]
Cat (*Felis domesticus*)	Adult cumulus cells	—	—	—	—	3	1	—	33.3%	[120]
Mule	Fetal fibroblasts	120	113	—	—	113 (transferred at different days)	1	—	0.88%	[121]
Horse	Adult skin fibroblasts	—	841	753	22/2.92%	22	1	0.13%	4.55%	[122]
Rat	Fetal fibroblasts	—	—	129	—	129 (transferred with 2-cell stage embryos)	2	1.55%	1.55%	[123]
Dog	Adult skin fibroblasts	—	—	1095	—	1095 (transferred with cleaved embryos)	2	0.18%	0.18%	[124]
Ferret	Adult cumulus cells	—	487	—	—	375 (transferred immediately after activation)	2	—	0.53%	[125]
Buffalo	Fetal fibroblasts & adult granulosa cells	—	—	—	42/11.04-31.39%	42	5 (one died 20 min after birth & 1 died on day 14 after birth)	—	11.9%	[126]
Camel	Adult cumulus cells	75	58	—	-/(63.88 ± 8.66)	26	1	—	3.85%	[127]
Cynomolgus monkey	Fetal fibroblast	127	109	79	—	79 (transferred with 2-cell stage embryos)	2	2.53%	2.53%	[7]

abnormalities may still exist, for example, large offspring syndrome, failure of the immune system, and respiratory disorders [19, 20]. Although abnormal phenotypes exist, cloned animals are mostly fertile and the offspring show normal phenotypes [21–23]. Therefore, the abnormalities are largely caused by epigenetic reprogramming defects rather than genetic mutations. Indeed, it has been reported that aberrant reprogramming and epigenetic memories inherited from donor cells are barriers that impede reprogramming [17, 18, 24–27]. Therefore, understanding of epigenetic reprogramming process is essential for prompting the improvement of SCNT technology.

Up to date, great efforts have been made to improve cloning efficiency. However, due to the limitation of methodology and the scarcity of the required materials, especially 1-cell and 2-cell stage embryos, the progress went slowly. While with the development and improvement of low input high throughput sequencing technology, higher resolution of genome-wide epigenetic modification landscapes in SCNT embryos were detected, and our understanding of epigenetic reprogramming becomes clearer [18, 24, 26, 28, 29].

In this review, we will summarize our current knowledge on epigenetic reprogramming, mainly on DNA methylation, histone modifications, histone variants, X chromosome inactivation (XCI), chromatin accessibility, and 3D chromatin structures during SCNT embryo development and recent progress on elevating cloning efficiency and quality. Focusing on how to overcome reprogramming barriers to facilitate SCNT reprogramming and further improve reproductive as well as therapeutic cloning.

2. DNA Methylation

DNA methylation (5-methylcytosine, 5mC) is an epigenetic mark that occurs at cytosine residues in the CpG dinucleotide, generally regarded as associated with transcriptional silencing [30]. About 60-80% of the CpG sites in the mammalian genome are modified by 5mC [31]. DNMT3A and DNMT3B are two methyltransferases essential for *de novo* DNA methylation, and DNMT1 is responsible for its maintenance during embryogenesis [32–34]. DNA demethylation is triggered by ten-eleven translocation (TET) protein-mediated oxidation from 5mC to 5-hydroxymethylcytosine (5hmC) followed by thymine DNA glycosylase- (TDG-) mediated base excision repair [35–37]. In mouse, both maternal and paternal alleles undergo demethylation through active and/or passive manner after fertilization and finally reached the lowest level at the blastocyst stage [38, 39]. It has been reported that knockout of *Dnmt3a* and *Dnmt3b* leads to mouse infertility [32, 40], and deletion of *Tet3* causes an increased frequency of developmental failure in embryos [41], suggesting that optimized DNA methylation pattern is essential for normal development. Thus, a DNA methylation pattern that resembled that of fertilized embryos may be a permissive state for SCNT embryo development.

2.1. DNA Methylation Is Globally Reprogrammed during SCNT Embryo Development. Given that somatic donor cells usually possess high DNA methylation levels [31], SCNT embryos must undergo global demethylation to reprogram the DNA methylation pattern of somatic cells to that of fertilized embryos. After activation, oocyte-stored TET3 immediately incorporated into pseudopronucleus (PPN) of the reconstructed embryo to catalyze conversion from 5mC to 5hmC, which implies active demethylation during SCNT embryo development [42], bearing resemblance with normal embryo development [42, 43]. Whole-genome bisulfite sequencing (WGBS) of mouse SCNT blastocysts revealed a very low DNA methylation level (15.6%) similar to that of IVF blastocysts (19.1%) [25]. Considering the high methylation level of the donor mouse embryonic fibroblast (MEF) cells (78%) used in the study, the result indicates successful global reprogramming of DNA methylation state. But this demethylation has not completed when the mouse SCNT embryos developed to the late 1 cell stage [28]. Our lab analyzed DNA methylation levels of SCNT embryos by using an embryo biopsy system along with single-cell reduced representation bisulfite sequencing (RRBS), and the results showed that at 2- and 4-cell stage, the SCNT samples possessed generally higher methylation level than the corresponding fertilized embryos [18], suggesting global demethylation in SCNT embryos may require several rounds of replication delusion.

2.2. Aberrant DNA Methylation Reprogramming in SCNT Embryos. Although successful global demethylation in blastocyst, aberrant DNA methylation patterns can be detected in SCNT embryos, even after implantation [5, 44, 45]. In mouse 4-cell stage SCNT embryos, especially arrested samples, the averaged methylation levels on gene body regions were significantly increased, resembling the trend of donor cells [18]. Similarly, cloned, but not fertilized, bovine morula possesses highly methylated nuclei in all blastomeres that resembled those of the fibroblast donor cells [44].

RRBS on 1 cell stage mouse SCNT embryos uncovered more than 20 genes, along with long interspersed elements (LINEs) and long terminal repeats (LTRs) defined as demethylation-resistant regions [28]. Nevertheless, by using ultralow-input WGBS, Gao et al. found that the persistently methylated differentially methylated regions (pDMRs) were moderately similar in arrest and normally developed NT embryos and were more frequently inherited from cleaved embryos to blastocyst stage, which reflects their functional irrelevance in the arrest of SCNT reprogramming [24]. Furthermore, they identified wide-spread regions that were aberrantly remethylated in SCNT embryos compared to the IVF counterparts, called remethylated differentially methylated regions (rDMRs), which are twice in arrested samples as many as in normally developed NT embryos (Figure 1). These rDMRs lead to misexpression of genes and retrotransposons important for zygotic genome activation (ZGA). Reduction of inappropriate DNA methylation rescued the developmental arrest at cleavage stages and facilitated proceeding to blastocyst development, increasing the blastocyst rate to 48.2% (compared to control of 39.5%) [24]. In conclusion, excessive DNA remethylation is a potent barrier that limits the full-term development of SCNT embryos, but the

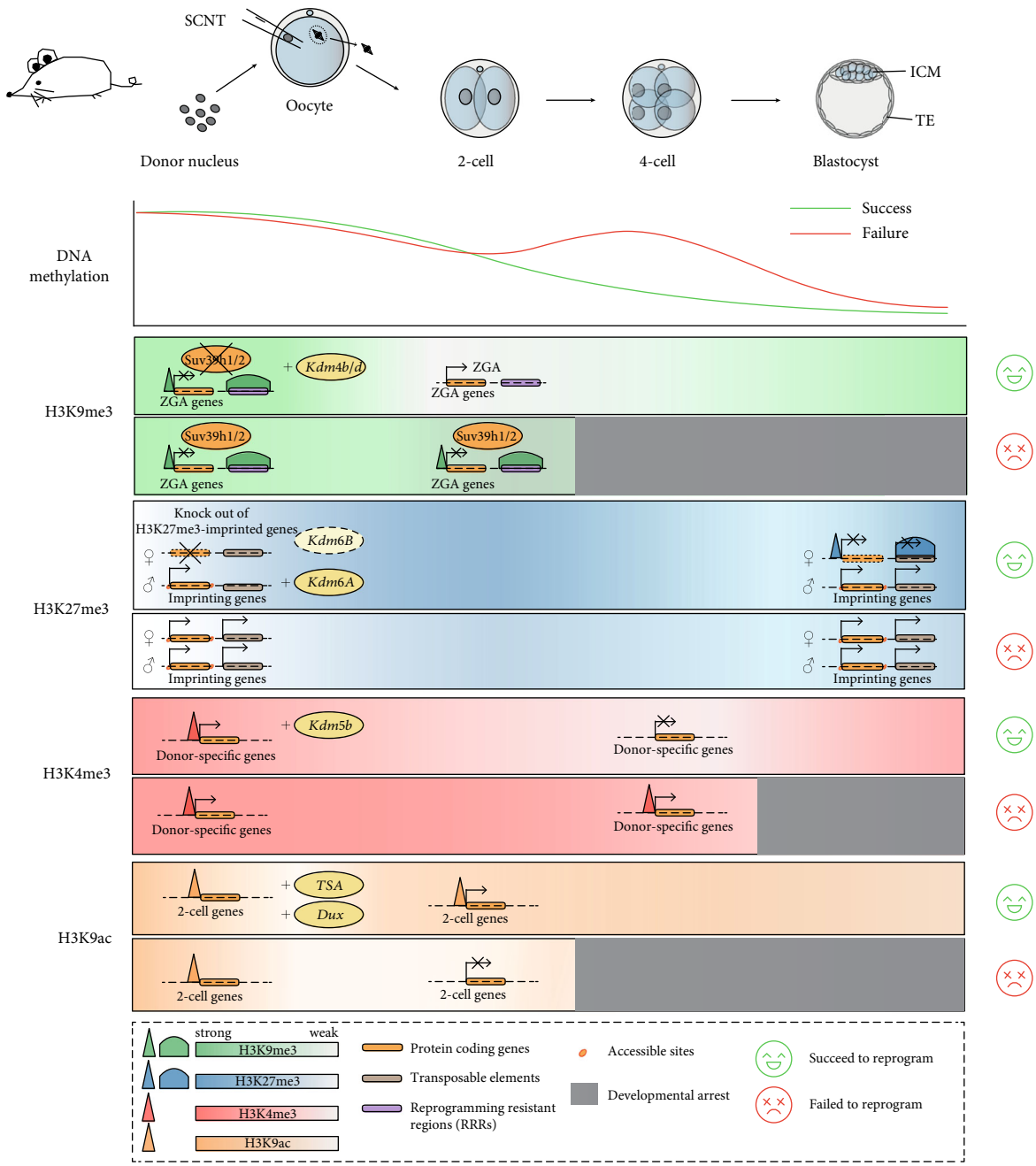


Figure 1: Epigenetic reprogramming of DNA methylation and histone modifications during mouse SCNT embryo development. DNA methylation: somatic donor cells usually possess high DNA methylation levels. After activation, the reconstructed embryos undergo global DNA demethylation although this demethylation has not been completed at the late 1 cell stage of SCNT embryos and requires several rounds of replication delusion. However, there is an aberrant remethylation in arrested 4 cell stage SCNT embryos and reduction of the inappropriate DNA methylation rescued the developmental arrest. *H3K9me3*: during SCNT embryo development, some zygotic genome activation (ZGA) genes and reprogramming resistant regions (RRRs) harbor donor cell-inherited H3K9me3 mark, which may be the cause of reprogramming failure. Removal of donor-inherited H3K9me3 either by ectopic expressing *Kdm4b/d* (H3K9me3-specific demethylases) or knockdown of *Suv39h1/2* (H3K9 methyltransferases) can help the embryo overcoming the reprogramming defects. *H3K27me3*: loss of H3K27me3-mediated imprinting leads to defects of extraembryonic tissues of SCNT embryos, such as large placenta phenotype. Although overexpression of H3K27me3-specific demethylase KDM6A elevated blastocyst developmental rate but not full-term development, both knock out of H3K27me3-imprinted genes and knockdown KDM6B can help SCNT embryos undergo successful reprogramming. *H3K4me3*: donor-inherited H3K4me3 is defined as an epigenetic barrier of SCNT reprogramming. H3K4me3 demethylation by *Kdm5b* overexpression is an important step to overcome reprogramming failure. *H3K9ac*: during SCNT development, aberrant H3K9ac regions impair ZGA. TSA treatment and Dux overexpression can correct the aberrant H3K9ac signal and help the embryos achieve successful reprogramming.

role of somatic-inherited DNA methylation still needs further proven, after all, an optimized DNA methylation pattern that resembled that of fertilized embryos is essential for SCNT reprogramming.

3. Histone Modifications

In eukaryotic cells, the basic functional unit of chromatin is the nucleosome, containing ~147 bp genomic DNA wrapped around a core histone octamer. Covalent histone modifications, such as acetylation, methylation, ubiquitination, and phosphorylation, are major epigenetic marks that regulate transcription [46–48]. Successful reprogramming of SCNT embryos should include reprogramming of histone modification patterns from somatic donor cells to those of normal embryos. Here, we will discuss the roles of several major histone modifications, including trimethylation at the 9th lysine residue of histone H3 (H3K9me3), trimethylation at the 27th lysine residue of histone H3 (H3K27me3), trimethylation at the 4th lysine residue of histone H3 (H3K4me3), and histone acetylation on SCNT reprogramming.

3.1. Aberrant H3K9me3 Reprogramming Impairs Preimplantation Development. H3K9me3 has been shown to play important roles in heterochromatin formation and repression of gene expression in various types of cells, including preimplantation embryos [7, 49]. In 2014, Matoba and colleagues identified 222 reprogramming resistant regions (RRRs) that failed to be activated in SCNT 2-cell embryos compared to IVF 2-cell embryos. Interestingly, these RRRs are enriched for H3K9me3 in somatic cells [17]. Removal of this epigenetic mark either through ectopic expression of *Kdm4d* (an H3K9me3-specific demethylase) in oocytes or knockdown of *Suv39h1* and *Suv39h2* (two H3K9 methyltransferases) in donor MEF cells not only attenuated the ZGA defect but also improved the reprogramming efficiency of SCNT embryos [17]. Further investigations by Liu et al. identified 7248 genes resisted donor-liked H3K9me3 signal at promoters in 2-cell stage SCNT embryos. Removal of the H3K9me3 mark inherited from donor cells by injecting *Kdm4b* helped the SCNT embryos go over 2-cell arrest and finally significantly elevated the potential of ntESC derivation, blastocyst rate, and even birth rate [18] (Figure 1). In bovine, KDM4D and KDM4E function as regulators that help SCNT embryos to break through H3K9me3 barriers [50]. Moreover, the expression of H3K9me3 demethylase Kdm4d/4a could reduce H3K9me3 level and significantly improve the efficiency of human SCNT blastocyst and ntESC cell line formation [13]. And the use of *Kdm4d* combined with histone deacetylase inhibitor (HDACi) trichostatin A (TSA) treatment successfully generated cloned cynomolgus (*Macaca fascicularis*) monkeys by using adult cumulus cells as donor cells [7], although the positive effect of TSA treatment might be functionally linked to H3K9me3 removal in mouse due to unchanged development potential by TSA treatment with *Kdm4d*-mRNA-injected mouse SCNT embryos [17]. The results above imply a conserved barrier of H3K9me3 inherited from donor cells during SCNT reprogramming in mammalian species.

Although the use of *Kdm4d* in SCNT results in an implantation rate comparable with that of IVF, only less than 15% of the implanted SCNT embryos develop to term, and abnormal large placentae are still observed in *Kdm4d*-injected SCNT embryos [17]. Additionally, Kdm4A addition was not able to enhance the in vivo long-term development capacity of porcine SCNT embryo [51], indicating H3K9me3 may mainly impede preimplantation development of SCNT embryos and other barriers may affect postimplantation development.

3.2. H3K27me3 Reprogramming Defects Are Obstacles in Pre- and Postimplantation SCNT Embryos. H3K27me3 is an epigenetic regulator widely known as a transcription repressor [52, 53]. During mouse preimplantation development, H3K27me3 is rapidly lost at both maternal and paternal alleles followed by dynamic especially when lineage specification of inner cell mass (ICM) and trophectoderm (TE) [54, 55]. Lots of studies have elucidated the critical role of H3K27me3 during both pre- and postimplantation embryo development [54, 56–59].

Aberrant H3K27me3 reprogramming may be a barrier of SCNT embryo development in various species [25, 60, 61]. Okae et al. identified three DNA methylation-independent imprinted genes *Gab1*, *Sfmbt2*, and *Slc38a4* showed loss of imprinting in all cloned mouse embryos [62], which might be involved in placentomegaly of cloned mouse when considering their important roles in placental development [63, 64]. Further studies found 76 genes with paternal allele-specific DNase I hypersensitive sites (DHSs) that are devoid of DNA methylation but harbor maternal allele-specific H3K27me3 [65]. Interestingly, all the three genes above are included in the 76 genes, which rise the suspect that the defect of H3K27me3 mediated imprinting may cause the abnormality of SCNT placentae. Indeed, many groups proved that loss of H3K27me3-imprinting in SCNT embryos disrupts mouse postimplantation development, and this defect can be detected earliest in blastocyst stage embryos up to now [25, 66, 67]. However, whether this defect exists more earlier in SCNT embryos requires further exploration [68]. A recent study found that the majority of H3K27me3-mediated imprinting regions are located to solo ERVK LTR repeats, which act as imprinted transcription initiation sites for noncoding RNAs and chimeric mRNA in extraembryonic tissues [69]. It is possible that the defects of H3K27me3 reprogramming are relevant to aberrant expression of transposable element during SCNT embryo development. Although restore the normal paternal expression of H3K27me3-imprinting genes (*Sfmbt2*, *Gab1*, and *Slc38a4*) in SCNT placentae by maternal knockout unchanged the enlarged placentae state [66], both correcting the expression of clustered miRNAs within the *Sfmbt2* gene and quadruple monoallelic deletion of *Sfmbt2*, *Jade1*, *Gab1*, and *Smoc1* ameliorates the placental phenotype, especially *Sfmbt2* [66, 67].

Apart from the impact of loss of H3K27me3-imprinting on SCNT postimplantation, another group demonstrated H3K27me3 as an obstacle of SCNT preimplantation development. Overexpression of the H3K27me3-specific demethylase KDM6A significantly increased the SCNT blastocyst

formation rate but did not improve the rate of full-term development, implies lack of KDM6A may be not the reason for loss of H3K27me3-dependent imprinting, at least in mouse. Contrastingly, knockdown of KDM6B not only facilitated ZGA and improved the blastocyst formation rate but also increased birth rate and ntESC establishment efficiency [68] (Figure 1). Collectively, both deposition on specific regions (like H3K27me3-imprinting genes) and appropriate removal of H3K27me3 are important for successful SCNT reprogramming although underlaid mechanisms are still unknown.

3.3. Somatic Inherited H3K4me3 Is a Potent Barrier of SCNT-Mediated Reprogramming. H3K4me3 is usually associated with transcriptional activation. Many groups have depicted the pattern of H3K4me3 during preimplantation in mouse [54, 70, 71]. Both appropriate removal of noncanonical H3K4me3 by *Kdm5b* in oocyte and establishment of canonical and broad H3K4me3 in preimplantation embryos are essential for normal mouse development [54, 70]. Unlike the well-described H3K4me3 pattern in normal mouse preimplantation embryos, studies about the whole H3K4me3 pattern during SCNT reprogramming have not been reported until now.

In 2016, we found that *Kdm5b* failed to be activated in 4-cell-arrest SCNT embryos. Injection of si-*Kdm5b* in MII oocytes largely reduced the rate of high-quality blastocyst development, and overexpression of *Kdm5b* helped the SCNT embryos to pass 4-cell arrest and significantly increased blastocyst formation rate and quality. What is more, the gene expression levels of NT 4-cell embryos were largely rescued by the overexpression of *Kdm5b* [18]. Considering the role of *Kdm5b* as H3K4me3 demethylase and the function of H3K4me3 on transcription initiation, it is possible that H3K4me3 mark with donor-specific signature may be a barrier of SCNT reprogramming (Figure 1). This point has been proved in *Xenopus*, human, and bovine SCNT embryos that donor-inherited H3K4me3 acts as an epigenetic barrier impacts SCNT reprogramming [72, 73]. H3K4me3 demethylation by *Kdm5b* overexpression not only attenuated ON-memory genes (genes highly expressed in donor cells and SCNT embryos but not IVF embryos) but also improved cloning efficiency. The results indicate that removal of the donor-specific H3K4me3 mark may efficiently reprogram the SCNT embryos but much more further investigations about roles of H3K4me3 during SCNT-mediated reprogramming need to be performed.

3.4. Aberrant Histone Acetylation Impairs the SCNT Efficiency. Histone acetylation usually occurs on the lysine residues of core histones and marks both promoters and enhancers. Acetylation has the potential to loosen nucleosome configuration and increase chromatin accessibility for transcription factors [74]. During ZGA, the persistent accessible enhancers are marked by H3K27ac and characterized by distal H3K4me3 deposition in human early embryos, while the poised enhancers are likely to be activated in later development by remarked H3K27ac in a tissue-specific manner [75]. In early zebrafish embryos, widespread H3K27ac deposition is found to be required for gene activation [76]. This indicates that histone acetylation reprogramming is another critical step for early embryo development.

When somatic cell nuclei are injected into the enucleated MII oocytes, the acetylated lysine residues are quickly deacetylated and then reacetylated after activation. The reestablishment of histone acetylation is essential for zygotic gene activation in cloned embryos [77]. However, several acetylation marks on histones, such as H4K8ac and H4K12ac, are persisted in the genome during SCNT, which may be responsible for the low cloning efficiency. On the other hand, histone deacetylase inhibitors (HDACi), which can improve histone acetylation and the success rate of cloning significantly, have been widely used during SCNT [78]. Recently, our group generated the genome-wide H3K9ac map during SCNT development and found the aberrant acetylated regions impair the zygotic gene activation. TSA treatment and Dux overexpression can correct the aberrant H3K9ac signal [79] (Figure 1). These suggest the reestablishment of histone acetylation is also a necessary part of epigenetic reprogramming. It should be noted that HDACi treatment can also improve nascent mRNA production [80] and gene expression [81] during SCNT embryo development, so the mechanism of HDACi treatment improves cloning efficiency still deserve further investigation.

4. Histone Variants

Aside from the canonical histones, histone variants endow chromatin critical functions, and their roles in oocyte-mediated reprogramming have been reviewed elsewhere [82–85]. The mammalian sperm genome is packaged into highly condensed chromatin consisting primarily of protamine but 5-15% residual histones. After fertilization, the paternal genome undergoes dramatic chromatin remodeling, and maternally stored histones, such as H3.3 (coded by *H3f3a* and *H3f3b*), are incorporated into the sperm nucleus as early as 1 h after fertilization [85]. And the incorporation is essential for the activation of the paternal genome and preimplantation development during embryogenesis [86].

Although the somatic cell genome is packaged by histones rather than protamine, global chromatin remodeling was still observed [85, 87]. After activation, donor cell-derived histone H3 variants H3.1, H3.2, and H3.3, as well as H2A, H2A.Z, and microH2A, were rapidly eliminated from the chromatin [87, 88]. All the three oocyte-stored H3 variants, H2A.X, and oocyte-specific H1 variant, H1FOO, were incorporated into the donor genome within minutes of nuclear transfer [87, 89, 90]. Knockdown of histone variant H3.3 in mouse oocytes results in compromised reprogramming and downregulation of key pluripotent genes, and this compromised reprogramming was rescued by injecting exogenous H3.3 mRNA, but not H3.2 mRNA into oocytes [85], revealing a critical role of optimized chromatin variants incorporation in normal SCNT reprogramming.

5. X Chromosome Inactivation (XCI)

XCI is a remarkable event during normal embryogenesis [62, 91]. X chromosome is inactivated during spermatogenesis. During mouse embryogenesis, the paternal X chromosome is reactivated at the 2-cell stage. After that, the paternal X chromosome will be silenced again through an imprinted manner and persisted in extraembryonic lineages. In contrast, the paternal X chromosome is reactivated in the epiblast in the late blastocyst, then, the X chromosome from maternal or paternal genomes is randomly inactivated during embryo development [92–94]. The precise regulation of dynamic activity of the X chromosome is crucial for the epigenetic reprogramming during early embryo development [95].

XCI ensures a similar dosage of X-linked genes between male and female cells. However, this event in SCNT embryos is largely abnormal among species [27, 51]. In mouse SCNT embryos, X-linked genes were largely downregulated, which is caused by ectopic expression of *Xist* from the active X chromosome regardless of sex, leading to abnormal inactivation of both X chromosomes [27]. Similarly, *Xist* is also known to be aberrantly expressed in bovine and pig SCNT embryos and proven to be associated with prenatal death [96, 97], suggesting excessive *Xist* expression may be a barrier of SCNT-mediated reprogramming. Deletion of *XIST* on the active X chromosome rescued global gene expression and resulted in about an 8- to 9-fold increase in cloning efficiency [27]. Concordantly, prior injection of *Xist*-siRNA into reconstructed oocytes normalized global gene expression of mouse SCNT embryos at the morula stage and further improved cloning efficiency 10-folds higher than control [98]. Moreover, correction of the abnormal XCI has a synergistic effect with TSA but ectopic activation of *Xist* is reprogramming barrier independent of H3K9me3 inherited from donor cells [25, 98]. Differently in pig, abnormal XCI seems linked with H3K9me3 for that increased quality of *XIST*-deficient SCNT embryos was associated with the global H3K9me3 reduction and vice versa; *Kdm4a* addition also induced *XIST* derepression in the active X chromosome [51]. This discrepancy may be a result of different XCI processes among different species, and the underlaid mechanisms require further understanding.

6. Chromatin Accessibility

Chromatin accessibility is a good indicator of transcriptional regulatory elements and can serve as a predictor of gene transcription activity. In recent years, with the development and improvement of low-input DNase I hypersensitive sequencing (liDNase-seq) and assay for transposase-accessible chromatin using sequencing (ATAC-seq), accessible chromatin sites of mouse and human preimplantation embryos enabled to be profiled [29, 65, 99, 100]. By using liDNase-seq, Lu et al. uncovered that DNase I-hypersensitive site (DHS) landscape is progressively established with a drastic increase at the 8-cell stage of mouse preimplantation embryos [29]. The global chromatin de- and recondensation is likely promoted by cis-regulating of LINE-1 transcriptional activity [101]. Transcription factors Nfya and Oct4 were responsible for DHS formation at 2- and 8-cell stage embryos, respectively [29].

Full-pattern of chromatin accessibility during mouse and human SCNT embryo development has not been elucidated, but a recent research profiled DHSs in donor cells and late-1-cell stage mouse SCNT embryos. They found SCNT-mediated reprogramming of chromatin accessibility is largely completed by 12 h after activation because DHSs of the donor cells are drastically changed to recapitulate that of the IVF zygotes within 12 h. Surprisingly, this change is DNA replication-independent, which is conserved in *Xenopus* SCNT embryos [102], and the switch from donor-specific TF network to that of zygotic may be the critical factor responsible for the DHS profile reprogramming [103].

Despite global reprogramming, some regions are resistant to reprogram [103]. Failure to close accessible somatic promoters or to open distal regulatory regions required for differentiation program may be the major reprogramming barriers. It is interesting that these regions are enriched for H3K9me3, a robust reprogramming barrier discussed above, in both donor cells and 2-cell SCNT embryos [103]. Considering the change of the TF network which accompanies with this reprogramming, failure of specific somatic cell TFs to dissociate from chromatin can also be a barrier in SCNT reprogramming. ATAC-seq on *Xenopus* SCNT embryos revealed great loss of chromatin accessible sites before first cleavage compared to that of donor cells, which is concordant with the pattern in mouse. The researchers found genes that are silenced but have preexisting open transcription start sites (TSSs) in donor cells are prone to be activated after SCNT, while genes resistant to reprogramming are associated with closed chromatin configurations [102]. It is possible that preexisted open accessibility of donor-specific genes and closed accessibility of zygotic-essential genes inherited from donor cells may be barriers during SCNT reprogramming, but it needs further proven.

7. Higher-Order Chromatin Structure

Chromatin in the nucleus of eukaryotic cells is packaged in a hierarchical structure, which is associated with many biological processes [104, 105]. The role of the 3D genome organization during mammalian embryogenesis has been investigated benefit from the advance of the low-input Hi-C (genome-wide chromosome conformation capture) technology in recent years [106–108], which reveals the removal and reestablishment of chromatin higher-order structure are essential for both mouse [106, 107] and human [108] embryogenesis.

A recent study of our group profiles the spatiotemporal dynamic of 3D chromatin structure in SCNT early embryos and reveals 3D chromatin structure can be rapidly reorganized to an embryo-like state after nuclear transfer. However, the aberrant TADs and compartment A/B organization can be observed and remain throughout preimplantation SCNT embryo development. Overexpression of KDM4B, a H3K9me3 demethylase, can partially improve the abnormal 3D chromatin structures [26] (Figure 2). This

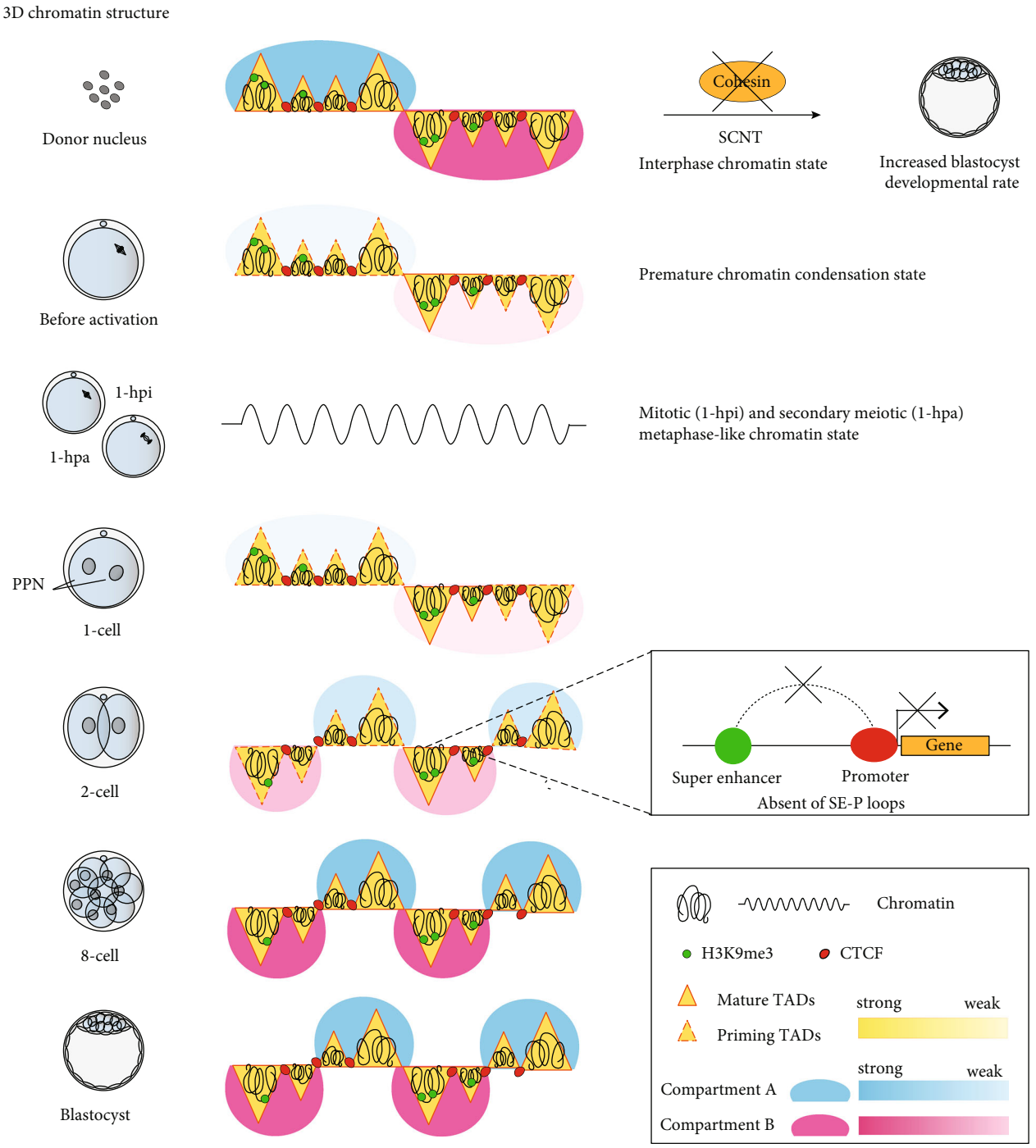


Figure 2: The higher-order chromatin organization in mouse SCNT embryos. Somatic donor cells exhibit interphase-state chromatin characterized by mature compartments and topologically associating domains (TADs). Before activated, the transferred nucleus first enters a mitotic-like state (premature chromatin condensation) followed by exhibiting mitotic and secondary meiotic metaphase-like chromatin states lacking compartments and TADs 1 hour postinjection (1-hpi) and 1-hour postactivation (1-hpa), respectively. TADs are stronger in SCNT 1-cell stage embryos and then become weaker at the 2-cell stage and gradually consolidating. Super enhancer-promoter (SE-P) loops that exist in fertilized 2-cell embryos are absent in SCNT 2-cell embryos, which is correlated with aberrant H3K9me3 and TAD persistence. Compartments A/B are markedly weak in 1-cell SCNT embryos and become increasingly strengthened afterward. By the 8-cell stage, somatic chromatin architecture is largely reset to embryonic patterns until the blastocyst stage. Predepleting cohesin in donor cells increases SCNT reprogramming efficiency.

indicates a correlation between the organization of 3D chromatin structure and histone modifications during epigenetic reprogramming.

8. Removal of Multiple Barriers Is a Promising Approach to Improve SCNT Reprogramming

It has been over two decades that the first mammalian species has been successfully cloned, but low efficiency was still observed until recently. Numerous efforts have been made to increase reprogramming efficiency by removing epigenetic barriers. Matoba et al. found H3K9me3 inherited from donor cells act as a barrier that impede mouse SCNT-mediated preimplantation development. Removal of H3K9me3 in donor cells by injecting *Kdm4d* mRNA into reconstructed embryos 5 hours postactivation (hpa) significantly increased the blastocyst rate up to 81.2% (% blastocyst of cleaved embryos) regardless of donor cell types and elevated birth rate from only 1% up to 8.7%. Besides, the rate of ntESC line derivation was increased from 10.1% to 50% after *Kdm4d* injection. Moreover, to prevent the establishment of H3K9me3 in donor cells, they knock down *Suv39h1/2* (H3K9me3 transferases) in donor cells prior SCNT and improved blastocyst rate from 6.7% to 49.9% [17]. Accordantly, our lab found another H3K9me3 demethylase, *Kdm4b*, efficiently removed the H3K9me3 barrier to increase blastocyst rate from about 30% to over 80%. Simultaneously, we found H3K4me3 may be a candidate epigenetic barrier that impedes SCNT-mediated reprogramming. Injection of *Kdm5b* mRNA into enucleated oocyte significantly improves mouse blastocyst rate from about 30% to over 50%. It is worth noting that coinjection of *Kdm4b* and *Kdm5b* successfully elevated blastocyst rate over 95% and led to over 11% of cloned embryos developing to live animals, moreover, 60% ntESC derivation efficiency based on the total number of MII oocytes rise the possibility that removing multibarriers may be a more efficient way to improve cloning efficiency [18].

We found excessive remethylation is a potent epigenetic barrier in another study. Optimized DNA methylation level by injecting siRNAs of *Dnmt3a* and *Dnmt3b* into enucleated oocytes, 48.2% blastocysts were generated from cleaved embryos (39.5% blastocyst rate of control group). Furthermore, of enucleated oocytes that subjected to Kdm4b+5b mRNA and siDnmt3a+3b co-injection, 92.3% cleaved embryos developed to blastocyst stage [24]. Another study of Matoba et al. demonstrated that using a combination of Xist knockout donor cells and overexpression of Kdm4b, more than 20% birth rate of mouse SCNT embryos were achieved [25], which is coincided with the idea that removing multiple epigenetic barriers is a more efficient method for SCNT reprogramming.

9. Concluding Remarks

SCNT provides the only way to reprogram somatic cells into totipotent embryos and generate viable animals [9, 109, 110]. After injected into enucleated oocytes, the donor nucleus quickly undergoes nuclear membrane breakdown followed by premature chromosome condensation (PCC), which is triggered by the M-phase-prompting factors (MPFs) stored at ooplasm [111]. After activation, the nuclear membrane is reformed to envelop PPN and incorporates amounts of maternal factors [112]. Then, the reconstructed embryos undergo SCNT-mediated embryogenesis. However, only few of reconstructed embryos can develop to the blastocyst stage, let alone develop to term. ZGA failure and disrupted transcriptome were detected in SCNT embryos very often, and this is largely affected by aberrant epigenetic reprogramming [4].

In this review, we concluded our understanding on epigenetic barriers of SCNT-mediated reprogramming and methods to overcome these epigenetic berries. Given that removal of multiple barriers that impede SCNT-mediated reprogramming gives a blastocyst rate over 95% of cleaved embryos and ntESC derivation efficiency of 60% based on the total number of MII oocytes [18], and over 20% birth rate of mouse [25], we demonstrate removing multiple barriers may be a more efficient approach to achieve complete reprogramming compared to single barrier removal. However, low birth rate compared with IVF counterpart and large placentae were still observed. Therefore, further studies need to focus on exploring more about reprogramming barriers and emphasizing on removing multiple barriers to achieve nearly complete SCNT reprogramming.

Abbreviations

SCNT: Somatic cell nuclear transfer
ntESCs: Nuclear transfer embryonic stem cells
XCI: X chromosome inactivation
5mC: 5-Methylcytosine
DNMT3A: DNA methyltransferase 3A
DNMT3B: DNA methyltransferase 3B
DNMT1: DNA methyltransferase 1
TET: Ten-eleven translocation
5hmC: 5-Hydroxymethylcytosine
TDG: Thymine DNA glycosylase
PPN: Pseudopronucleus
WGBS: Whole-genome bisulfite sequencing
MEF: Mouse embryonic fibroblast
RRBS: Reduced representation bisulfite sequencing
LINEs: Long interspersed elements
LTRs: Long terminal repeats
pDMRs: Persistently methylated differentially methylated regions
rDMRs: Remethylated differentially methylated regions
ZGA: Zygotic genome activation
H3K9me3: Trimethylation at the 9^{th} lysine residue of histone H3
H3K27me3: Trimethylation at the 27^{th} lysine residue of histone H3
H3K4me3: Trimethylation at the 4^{th} lysine residue of histone H3
HDACi: Histone deacetylase inhibitor
TSA: Trichostatin A
PRC2: Polycomb repressive complex 2
liDNase-seq: DNase I hypersensitive sequencing

ATAC-seq: Transposase-accessible chromatin using sequencing
DHS: DNase I-hypersensitive site
TSSs: Transcription start sites
hpa: Hour postactivation
PCC: Premature chromosome condensation
MPFs: M-phase-prompting factors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China (2018YFC1004000), the National Natural Science Foundation of China (32070802, 31721003, 31820103009, 31701262, and 81630035), and the Shanghai Science and Technology Development Funds (20QA1409700).

References

[1] J. B. Gurdon, "The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles," *Journal of Embryology and Experimental Morphology*, vol. 10, pp. 622–640, 1962.

[2] F. Lu and Y. Zhang, "Cell totipotency: molecular features, induction, and maintenance," *National Science Review*, vol. 2, no. 2, pp. 217–225, 2015.

[3] I. Wilmut, A. E. Schnieke, J. McWhir, A. J. Kind, and K. H. Campbell, "Viable offspring derived from fetal and adult mammalian cells," *Nature*, vol. 385, no. 6619, pp. 810–813, 1997.

[4] S. Matoba and Y. Zhang, "Somatic cell nuclear transfer reprogramming: mechanisms and applications," *Cell Stem Cell*, vol. 23, no. 4, pp. 471–485, 2018.

[5] M. Teperek and K. Miyamoto, "Nuclear reprogramming of sperm and somatic nuclei in eggs and oocytes," *Reproductive medicine and biology*, vol. 12, no. 4, pp. 133–149, 2013.

[6] X. Wang, J. Qu, J. Li, H. He, Z. Liu, and Y. Huan, "Epigenetic reprogramming during somatic cell nuclear transfer: recent progress and future directions," *Frontiers in Genetics*, vol. 11, p. 205, 2020.

[7] Z. Liu, Y. Cai, Y. Wang et al., "Cloning of macaque monkeys by somatic cell nuclear transfer," *Cell*, vol. 172, no. 4, pp. 881–887.e7, 2018.

[8] M. J. Munsie, A. E. Michalska, C. M. O'Brien, A. O. Trounson, M. F. Pera, and P. S. Mountford, "Isolation of pluripotent embryonic stem cells from reprogrammed adult mouse somatic cell nuclei," *Current Biology*, vol. 10, no. 16, pp. 989–992, 2000.

[9] T. Wakayama, V. Tabar, I. Rodriguez, A. C. Perry, L. Studer, and P. Mombaerts, "Differentiation of embryonic stem cell lines generated from adult somatic cells by nuclear transfer," *Science*, vol. 292, no. 5517, pp. 740–743, 2001.

[10] J. A. Byrne, D. A. Pedersen, L. L. Clepper et al., "Producing primate embryonic stem cells by somatic cell nuclear transfer," *Nature*, vol. 450, no. 7169, pp. 497–502, 2007.

[11] M. Tachibana, P. Amato, M. Sparman et al., "Human embryonic stem cells derived by somatic cell nuclear transfer," *Cell*, vol. 153, no. 6, pp. 1228–1238, 2013.

[12] Y. G. Chung, J. H. Eum, J. E. Lee et al., "Human somatic cell nuclear transfer using adult cells," *Cell Stem Cell*, vol. 14, no. 6, pp. 777–780, 2014.

[13] Y. G. Chung, S. Matoba, Y. Liu et al., "Histone demethylase expression enhances human somatic cell nuclear transfer efficiency and promotes derivation of pluripotent stem cells," *Cell Stem Cell*, vol. 17, no. 6, pp. 758–766, 2015.

[14] M. Yamada, B. Johannesson, I. Sagi et al., "Human oocytes reprogram adult somatic nuclei of a type 1 diabetic to diploid pluripotent stem cells," *Nature*, vol. 510, no. 7506, pp. 533–536, 2014.

[15] X. Yang, S. L. Smith, X. C. Tian, H. A. Lewin, J. P. Renard, and T. Wakayama, "Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning," *Nature Genetics*, vol. 39, no. 3, pp. 295–302, 2007.

[16] R. Krishnakumar and R. H. Blelloch, "Epigenetics of cellular reprogramming," *Current Opinion in Genetics & Development*, vol. 23, no. 5, pp. 548–555, 2013.

[17] S. Matoba, Y. Liu, F. Lu et al., "Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation," *Cell*, vol. 159, no. 4, pp. 884–895, 2014.

[18] W. Liu, X. Liu, C. Wang et al., "Identification of key factors conquering developmental arrest of somatic cell cloned embryos by combining embryo biopsy and single-cell sequencing," *Cell discovery*, vol. 2, no. 1, 2016.

[19] P. Loi, D. Iuso, M. Czernik, and A. Ogura, "A new, dynamic era for somatic cell nuclear transfer?," *Trends in Biotechnology*, vol. 34, no. 10, pp. 791–797, 2016.

[20] A. Ogura, K. Inoue, and T. Wakayama, "Recent advancements in cloning by somatic cell nuclear transfer," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 368, 2013.

[21] K. L. Tamashiro, T. Wakayama, H. Akutsu et al., "Cloned mice have an obese phenotype not transmitted to their offspring," *Nature Medicine*, vol. 8, no. 3, pp. 262–267, 2002.

[22] J. Fulka Jr., N. Miyashita, T. Nagai, and A. Ogura, "Do cloned mammals skip a reprogramming step?," *Nature Biotechnology*, vol. 22, no. 1, pp. 25-26, 2004.

[23] S. Wakayama, T. Kohda, H. Obokata et al., "Successful serial recloning in the mouse over multiple generations," *Cell Stem Cell*, vol. 12, no. 3, pp. 293–297, 2013.

[24] R. Gao, C. Wang, Y. Gao et al., "Inhibition of aberrant DNA re-methylation improves post-implantation development of somatic cell nuclear transfer embryos," *Cell stem cell*, vol. 23, no. 3, pp. 426–435.e5, 2018.

[25] S. Matoba, H. Wang, L. Jiang et al., "Loss of H3K27me3 imprinting in somatic cell nuclear transfer embryos disrupts post-implantation development," *Cell stem cell*, vol. 23, no. 3, pp. 343–354.e5, 2018.

[26] M. Chen, Q. Zhu, C. Li et al., "Chromatin architecture reorganization in murine somatic cell nuclear transfer embryos," *Nature Communications*, vol. 11, no. 1, p. 1813, 2020.

[27] K. Inoue, T. Kohda, M. Sugimoto et al., "Impeding Xist expression from the active X chromosome improves mouse somatic cell nuclear transfer," *Science*, vol. 330, no. 6003, pp. 496–499, 2010.

[28] M. M. Chan, Z. D. Smith, D. Egli, A. Regev, and A. Meissner, "Mouse ooplasm confers context-specific reprogramming capacity," *Nature Genetics*, vol. 44, no. 9, pp. 978–980, 2012.

[29] F. Lu, Y. Liu, A. Inoue, T. Suzuki, K. Zhao, and Y. Zhang, "Establishing chromatin regulatory landscape during mouse

preimplantation development," *Cell*, vol. 165, no. 6, pp. 1375–1388, 2016.

[30] A. Bird, "DNA methylation patterns and epigenetic memory," *Genes & Development*, vol. 16, no. 1, pp. 6–21, 2002.

[31] Z. D. Smith and A. Meissner, "DNA methylation: roles in mammalian development," *Nature Reviews. Genetics*, vol. 14, no. 3, pp. 204–220, 2013.

[32] M. Okano, D. W. Bell, D. A. Haber, and E. Li, "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development," *Cell*, vol. 99, no. 3, pp. 247–257, 1999.

[33] H. Wu and Y. Zhang, "Reversing DNA methylation: mechanisms, genomics, and biological functions," *Cell*, vol. 156, no. 1-2, pp. 45–68, 2014.

[34] A. Hermann, R. Goyal, and A. Jeltsch, "The Dnmt1 DNA-(cytosine-C5)-methyltransferase Methylates DNA Processively with High Preference for Hemimethylated Target Sites," *The Journal of Biological Chemistry*, vol. 279, no. 46, pp. 48350–48359, 2004.

[35] X. Wu and Y. Zhang, "TET-mediated active DNA demethylation: mechanism, function and beyond," *Nature Reviews Genetics*, vol. 18, no. 9, pp. 517–534, 2017.

[36] N. Verma, H. Pan, L. C. Doré et al., "TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells," *Nature Genetics*, vol. 50, no. 1, pp. 83–95, 2018.

[37] R. M. Kohli and Y. Zhang, "TET enzymes, TDG and the dynamics of DNA demethylation," *Nature*, vol. 502, no. 7472, pp. 472–479, 2013.

[38] F. Guo, X. Li, D. Liang et al., "Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote," *Cell Stem Cell*, vol. 15, no. 4, pp. 447–459, 2014.

[39] L. Shen, A. Inoue, J. He, Y. Liu, F. Lu, and Y. Zhang, "Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes," *Cell Stem Cell*, vol. 15, no. 4, pp. 459–471, 2014.

[40] E. Li, T. H. Bestor, and R. Jaenisch, "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality," *Cell*, vol. 69, no. 6, pp. 915–926, 1992.

[41] J. Kang, M. Lienhard, W. A. Pastor et al., "Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 31, pp. E4236–E4245, 2015.

[42] T.-P. Gu, F. Guo, H. Yang et al., "The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes," *Nature*, vol. 477, no. 7366, pp. 606–610, 2011.

[43] M. Wossidlo, T. Nakamura, K. Lepikhov et al., "5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming," *Nature Communications*, vol. 2, no. 1, 2011.

[44] W. Dean, F. Santos, M. Stojkovic et al., "Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 24, pp. 13734–13738, 2001.

[45] J. R. Peat and W. Reik, "Incomplete methylation reprogramming in SCNT embryos," *Nature Genetics*, vol. 44, no. 9, pp. 965-966, 2012.

[46] C. Martin and Y. Zhang, "The diverse functions of histone lysine methylation," *Nature Reviews Molecular Cell Biology*, vol. 6, no. 11, pp. 838–849, 2005.

[47] M. Grunstein, "Histone acetylation in chromatin structure and transcription," *Nature*, vol. 389, no. 6649, pp. 349–352, 1997.

[48] P. Tessarz and T. Kouzarides, "Histone core modifications regulating nucleosome structure and dynamics," *Nature Reviews. Molecular Cell Biology*, vol. 15, no. 11, pp. 703–708, 2014.

[49] J. S. Becker, D. Nicetto, and K. S. Zaret, "H3K9me3-dependent heterochromatin: barrier to cell fate changes," *Trends in Genetics*, vol. 32, no. 1, pp. 29–41, 2016.

[50] X. Liu, Y. Wang, Y. Gao et al., "H3K9 demethylase KDM4E is an epigenetic regulator for bovine embryonic development and a defective factor for nuclear reprogramming," *Development*, vol. 145, no. 4, p. dev158261, 2018.

[51] D. Ruan, J. Peng, X. Wang et al., "XIST derepression in active X chromosome hinders pig somatic cell nuclear transfer," *Stem Cell Reports*, vol. 10, no. 2, pp. 494–508, 2018.

[52] L. Di Croce and K. Helin, "Transcriptional regulation by polycomb group proteins," *Nature Structural & Molecular Biology*, vol. 20, no. 10, pp. 1147–1155, 2013.

[53] J. A. Simon and R. E. Kingston, "Mechanisms of polycomb gene silencing: knowns and unknowns," *Nature Reviews Molecular Cell Biology*, vol. 10, no. 10, pp. 697–708, 2009.

[54] X. Liu, C. Wang, W. Liu et al., "Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos," *Nature*, vol. 537, no. 7621, pp. 558–562, 2016.

[55] H. Zheng, B. Huang, B. Zhang et al., "Resetting epigenetic memory by reprogramming of histone modifications in mammals," *Molecular Cell*, vol. 63, no. 6, pp. 1066–1079, 2016.

[56] A. Inoue, Z. Chen, Q. Yin, and Y. Zhang, "Maternal Eed knockout causes loss of H3K27me3 imprinting and random X inactivation in the extraembryonic cells," *Genes & Development*, vol. 32, no. 23-24, pp. 1525–1536, 2018.

[57] W. Zhang, Z. Chen, Q. Yin, D. Zhang, C. Racowsky, and Y. Zhang, "Maternal-biased H3K27me3 correlates with paternal-specific gene expression in the human morula," *Genes & Development*, vol. 33, no. 7-8, pp. 382–387, 2019.

[58] Q. Xu and W. Xie, "Epigenome in early mammalian development: inheritance, reprogramming and establishment," *Trends in Cell Biology*, vol. 28, no. 3, pp. 237–253, 2018.

[59] R. Xu, C. Li, X. Liu, and S. Gao, "Insights into epigenetic patterns in mammalian early embryos," *Protein & cell*, vol. 12, no. 1, pp. 7–28, 2020.

[60] B. Xie, H. Zhang, R. Wei et al., "Histone H3 lysine 27 trimethylation acts as an epigenetic barrier in porcine nuclear reprogramming," *Reproduction*, vol. 151, pp. 9–16, 2016.

[61] C. Zhou, Y. Wang, J. Zhang et al., "H3K27me3 is an epigenetic barrier while KDM6A overexpression improves nuclear reprogramming efficiency," *FASEB journal*, vol. 33, pp. 4638–4652, 2018.

[62] H. Okae, S. Matoba, T. Nagashima et al., "RNA sequencing-based identification of aberrant imprinting in cloned mice," *Human Molecular Genetics*, vol. 23, no. 4, pp. 992–1001, 2014.

[63] M. Itoh, Y. Yoshida, K. Nishida, M. Narimatsu, M. Hibi, and T. Hirano, "Role of Gab1 in heart, placenta, and skin development and growth factor- and cytokine-induced extracellular

signal-regulated kinase mitogen-activated protein kinase activation," *Molecular and Cellular Biology*, vol. 20, no. 10, pp. 3695–3704, 2000.

[64] K. Miri, K. Latham, B. Panning, Z. Zhong, A. Andersen, and S. Varmuza, "The imprinted polycomb group gene Sfmbt2 is required for trophoblast maintenance and placenta development," *Development*, vol. 140, no. 22, pp. 4480–4489, 2013.

[65] A. Inoue, L. Jiang, F. Lu, T. Suzuki, and Y. Zhang, "Maternal H3K27me3 controls DNA methylation-independent imprinting," *Nature*, vol. 547, no. 7664, pp. 419–424, 2017.

[66] K. Inoue, N. Ogonuki, S. Kamimura et al., "Loss of H3K27me3 imprinting in the Sfmbt2 miRNA cluster causes enlargement of cloned mouse placentas," *Nature Communications*, vol. 11, no. 1, p. 2150, 2020.

[67] L. Y. Wang, Z. K. Li, L. B. Wang et al., "Overcoming intrinsic H3K27me3 imprinting barriers improves post-implantation development after somatic cell nuclear transfer," *Cell stem cell*, vol. 27, no. 2, pp. 315–325.e5, 2020.

[68] L. Yang, L. Song, X. Liu, L. Bai, and G. Li, "KDM6A and KDM6B play contrasting roles in nuclear transfer embryos revealed by MERVL reporter system," *EMBO reports*, vol. 19, no. 12, 2018.

[69] C. W. Hanna, R. Pérez-Palacios, L. Gahurova et al., "Endogenous retroviral insertions drive non-canonical imprinting in extra-embryonic tissues," *Genome Biology*, vol. 20, no. 1, p. 225, 2019.

[70] B. Zhang, H. Zheng, B. Huang et al., "Allelic reprogramming of the histone modification H3K4me3 in early mammalian development," *Nature*, vol. 537, no. 7621, pp. 553–557, 2016.

[71] J. A. Dahl, I. Jung, H. Aanes et al., "Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition," *Nature*, vol. 537, no. 7621, pp. 548–552, 2016.

[72] E. Hörmanseder, A. Simeone, G. E. Allen et al., "H3K4 methylation-dependent memory of somatic cell identity inhibits reprogramming and development of nuclear transfer embryos," *Cell stem cell*, vol. 21, no. 1, pp. 135–143.e6, 2017.

[73] C. Zhou, J. Zhang, M. Zhang et al., "Transcriptional memory inherited from donor cells is a developmental defect of bovine cloned embryos," *FASEB journal*, vol. 34, no. 1, pp. 1637–1651, 2020.

[74] S. K. Kurdistani, S. Tavazoie, and M. Grunstein, "Mapping global histone acetylation patterns to gene expression," *Cell*, vol. 117, no. 6, pp. 721–733, 2004.

[75] J. Wu, J. Xu, B. Liu et al., "Chromatin analysis in human early development reveals epigenetic transition during ZGA," *Nature*, vol. 557, no. 7704, pp. 256–260, 2018.

[76] B. Zhang, X. Wu, W. Zhang et al., "Widespread enhancer dememorization and promoter priming during parental-to-zygotic transition," *Molecular cell*, vol. 72, 2018.

[77] F. Wang, Z. Kou, Y. Zhang, and S. Gao, "Dynamic reprogramming of histone acetylation and methylation in the first cell cycle of cloned mouse embryos," *Biology of Reproduction*, vol. 77, no. 6, pp. 1007–1016, 2007.

[78] B. P. Enright, C. Kubota, X. Yang, and X. C. Tian, "Epigenetic characteristics and development of embryos cloned from donor cells treated by trichostatin A or 5-aza-2'-deoxycytidine," *Biology of Reproduction*, vol. 69, no. 3, pp. 896–901, 2003.

[79] G. Yang, L. Zhang, W. Liu et al., "Dux-mediated correactions of aberrant H3K9ac during 2-cell genome activation optimize efficiency of somatic cell nuclear transfer," *Cell stem cell*, vol. 28, no. 1, pp. 150–163. e5, 2020.

[80] N. Van Thuan, H. T. Bui, J. H. Kim et al., "The histone deacetylase inhibitor scriptaid enhances nascent mRNA production and rescues full-term development in cloned inbred mice," *Reproduction*, vol. 138, no. 2, pp. 309–317, 2009.

[81] Y. Tsuji, Y. Kato, and Y. Tsunoda, "The developmental potential of mouse somatic cell nuclear-transferred oocytes treated with trichostatin A and 5-aza-2'-deoxycytidine," *Zygote*, vol. 17, no. 2, pp. 109–115, 2009.

[82] P. Yang, W. Wu, and T. S. Macfarlan, "Maternal histone variants and their chaperones promote paternal genome activation and boost somatic cell reprogramming," *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, vol. 37, no. 1, pp. 52–59, 2015.

[83] P. B. Talbert and S. Henikoff, "Histone variants on the move: substrates for chromatin dynamics," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 2, pp. 115–126, 2017.

[84] D. Filipescu, E. Szenker, and G. Almouzni, "Developmental roles of histone H3 variants and their chaperones," *Trends in Genetics*, vol. 29, no. 11, pp. 630–640, 2013.

[85] D. Wen, L. A. Banaszynski, Z. Rosenwaks, C. D. Allis, and S. Rafii, "H3.3 replacement facilitates epigenetic reprogramming of donor nuclei in somatic cell nuclear transfer embryos," *Nucleus*, vol. 5, no. 5, pp. 369–375, 2014.

[86] Q. Kong, L. A. Banaszynski, F. Geng et al., "Histone variant H3.3-mediated chromatin remodeling is essential for paternal genome activation in mouse preimplantation embryos," *The Journal of Biological Chemistry*, vol. 293, no. 10, pp. 3829–3838, 2018.

[87] B. Nashun, T. Akiyama, M. G. Suzuki, and F. Aoki, "Dramatic replacement of histone variants during genome remodeling in nuclear-transferred embryos," *Epigenetics*, vol. 6, pp. 1489–1497, 2011.

[88] C. C. Chang, S. Gao, L. Y. Sung et al., "Rapid elimination of the histone variant MacroH2A from somatic cell heterochromatin after nuclear transfer," *Cellular Reprogramming*, vol. 12, no. 1, pp. 43–53, 2010.

[89] S. Gao, Y. G. Chung, M. H. Parseghian, G. J. King, E. Y. Adashi, and K. E. Latham, "Rapid H1 linker histone transitions following fertilization or somatic cell nuclear transfer: evidence for a uniform developmental program in mice," *Developmental Biology*, vol. 266, no. 1, pp. 62–75, 2004.

[90] T. Teranishi, M. Tanaka, S. Kimoto et al., "Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer," *Developmental Biology*, vol. 266, no. 1, pp. 76–86, 2004.

[91] A. Sahakyan, Y. Yang, and K. Plath, "The role of Xist in X-chromosome dosage compensation," *Trends in Cell Biology*, vol. 28, no. 12, pp. 999–1013, 2018.

[92] K. Plath, J. Fang, S. K. Mlynarczyk-Evans et al., "Role of histone H3 lysine 27 methylation in X inactivation," *Science*, vol. 300, pp. 131–135, 2003.

[93] R. Cao, L. Wang, H. Wang et al., "Role of histone H3 lysine 27 methylation in polycomb-group silencing," *Science*, vol. 298, no. 5595, pp. 1039–1043, 2002.

[94] J. T. Lee and M. S. Bartolomei, "X-inactivation, imprinting, and long noncoding RNAs in health and disease," *Cell*, vol. 152, no. 6, pp. 1308–1323, 2013.

[95] T. Ohhata and A. Wutz, "Reactivation of the inactive X chromosome in development and reprogramming," *Cellular and*

Molecular Life Sciences: CMLS, vol. 70, no. 14, pp. 2443–2461, 2013.

[96] F. Xue, X. C. Tian, F. Du et al., "Aberrant patterns of X chromosome inactivation in bovine clones," *Nature Genetics*, vol. 31, no. 2, pp. 216–220, 2002.

[97] L. Jiang, L. Lai, M. Samuel, R. S. Prather, X. Yang, and X. C. Tian, "Expression of X-linked genes in deceased neonates and surviving cloned female piglets," *Molecular Reproduction and Development*, vol. 75, no. 2, pp. 265–273, 2008.

[98] S. Matoba, K. Inoue, T. Kohda et al., "RNAi-mediated knockdown of Xist can rescue the impaired postimplantation development of cloned mouse embryos," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 51, pp. 20621–20626, 2011.

[99] L. Gao, K. Wu, Z. Liu et al., "Chromatin accessibility landscape in human early embryos and its association with evolution," *Cell*, vol. 173, no. 1, pp. 248–259.e15, 2018.

[100] J. Wu, B. Huang, H. Chen et al., "The landscape of accessible chromatin in mammalian preimplantation embryos," *Nature*, vol. 534, no. 7609, pp. 652–657, 2016.

[101] J. W. Jachowicz, X. Bing, J. Pontabry, A. Bošković, O. J. Rando, and M. E. Torres-Padilla, "LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo," *Nature Genetics*, vol. 49, no. 10, pp. 1502–1510, 2017.

[102] K. Miyamoto, K. T. Nguyen, G. E. Allen et al., "Chromatin accessibility impacts transcriptional reprogramming in oocytes," *Cell Reports*, vol. 24, no. 2, pp. 304–311, 2018.

[103] M. N. Djekidel, A. Inoue, S. Matoba et al., "Reprogramming of chromatin accessibility in somatic cell nuclear transfer is DNA replication independent," *Cell Reports*, vol. 23, no. 7, pp. 1939–1947, 2018.

[104] M. J. Fullwood, M. H. Liu, Y. F. Pan et al., "An oestrogen-receptor-α-bound human chromatin interactome," *Nature*, vol. 462, no. 7269, pp. 58–64, 2009.

[105] Y. Atlasi and H. G. Stunnenberg, "The interplay of epigenetic marks during stem cell differentiation and development," *Nature Reviews. Genetics*, vol. 18, no. 11, pp. 643–658, 2017.

[106] Y. Ke, Y. Xu, X. Chen et al., "3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis," *Cell*, vol. 170, no. 2, pp. 367–381.e20, 2017.

[107] Z. Du, H. Zheng, B. Huang et al., "Allelic reprogramming of 3D chromatin architecture during early mammalian development," *Nature*, vol. 547, no. 7662, pp. 232–235, 2017.

[108] X. Chen, Y. Ke, K. Wu et al., "Key role for CTCF in establishing chromatin structure in human embryos," *Nature*, vol. 576, no. 7786, pp. 306–310, 2019.

[109] T. Brambrink, K. Hochedlinger, G. Bell, and R. Jaenisch, "ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 4, pp. 933–938, 2006.

[110] D. Egli, J. Rosains, G. Birkhoff, and K. Eggan, "Developmental reprogramming after chromosome transfer into mitotic mouse zygotes," *Nature*, vol. 447, no. 7145, pp. 679–685, 2007.

[111] K. H. Campbell, P. Loi, P. J. Otaegui, and I. Wilmut, "Cell cycle co-ordination in embryo cloning by nuclear transfer," *Reviews of Reproduction*, vol. 1, no. 1, pp. 40–46, 1996.

[112] R. S. Prather, B. Kühholzer, L. Lai, and K. W. Park, "Changes in the structure of nuclei after transfer to oocytes," *Cloning*, vol. 2, no. 3, pp. 117–122, 2000.

[113] J. B. Cibelli, S. L. Stice, P. J. Golueke et al., "Cloned transgenic calves produced from nonquiescent fetal fibroblasts," *Science*, vol. 280, no. 5367, pp. 1256–1258, 1998.

[114] Y. Kato, T. Tani, Y. Sotomaru et al., "Eight calves cloned from somatic cells of a single adult," *Science*, vol. 282, no. 5396, pp. 2095–2098, 1998.

[115] T. Wakayama, A. C. Perry, M. Zuccotti, K. R. Johnson, and R. Yanagimachi, "Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei," *Nature*, vol. 394, no. 6691, pp. 369–374, 1998.

[116] A. Baguisi, E. Behboodi, D. T. Melican et al., "Production of goats by somatic cell nuclear transfer," *Nature Biotechnology*, vol. 17, no. 5, pp. 456–461, 1999.

[117] A. Onishi, M. Iwamoto, T. Akita et al., "Pig cloning by microinjection of fetal fibroblast nuclei," *Science*, vol. 289, no. 5482, pp. 1188–1190, 2000.

[118] I. A. Polejaeva, S. H. Chen, T. D. Vaught et al., "Cloned pigs produced by nuclear transfer from adult somatic cells," *Nature*, vol. 407, no. 6800, pp. 86–90, 2000.

[119] P. Chesné, P. G. Adenot, C. Viglietta, M. Baratte, L. Boulanger, and J. P. Renard, "Cloned rabbits produced by nuclear transfer from adult somatic cells," *Nature Biotechnology*, vol. 20, no. 4, pp. 366–369, 2002.

[120] T. Shin, D. Kraemer, J. Pryor et al., "A cat cloned by nuclear transplantation," *Nature*, vol. 415, no. 6874, p. 859, 2002.

[121] G. L. Woods, K. L. White, D. K. Vanderwall et al., "A mule cloned from fetal cells by nuclear transfer," *Science*, vol. 301, no. 5636, p. 1063, 2003.

[122] C. Galli, I. Lagutina, G. Crotti et al., "A cloned horse born to its dam twin," *Nature*, vol. 424, no. 6949, p. 635, 2003.

[123] Q. Zhou, J. P. Renard, G. Le Friec et al., "Generation of fertile cloned rats by regulating oocyte activation," *Science*, vol. 302, no. 5648, p. 1179, 2003.

[124] B. C. Lee, M. K. Kim, G. Jang et al., "Dogs cloned from adult somatic cells," *Nature*, vol. 436, p. 641, 2005.

[125] Z. Li, X. Sun, J. Chen et al., "Cloned ferrets produced by somatic cell nuclear transfer," *Developmental Biology*, vol. 293, no. 2, pp. 439–448, 2006.

[126] D. Shi, F. Lu, Y. Wei et al., "Buffalos (Bubalus bubalis) cloned by nuclear transfer of somatic cells," *Biology of Reproduction*, vol. 77, no. 2, pp. 285–291, 2007.

[127] N. A. Wani, U. Wernery, F. A. Hassan, R. Wernery, and J. A. Skidmore, "Production of the first cloned camel by somatic cell nuclear Transfer1," *Biology of Reproduction*, vol. 82, no. 2, pp. 373–379, 2010.

Hindawi
Stem Cells International
Volume 2021, Article ID 8817581, 12 pages
https://doi.org/10.1155/2021/8817581

Review Article

Chromatin Regulation in Development: Current Understanding and Approaches

Zi Hao Zheng,[1,2] Tsz Wing Sam,[1,2] YingYing Zeng,[1,3] Justin Jang Hann Chu,[4,5,6,7] and Yuin-Han Loh[1,2,7,8]

[1]*Laboratory for Epigenetics, Stem Cells & Cell Therapy, Programme in Stem Cell, Regenerative Medicine and Aging, ASTAR Institute of Molecular and Cell Biology, Singapore 138673*
[2]*Department of Physiology, NUS Yong Loo Lin School of Medicine, 2 Medical Drive, MD9, Singapore 117593*
[3]*School of Biological Sciences, Nanyang Technological University, Singapore 637551*
[4]*Laboratory of Molecular RNA Virology and Antiviral Strategies, Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University Health System, National University of Singapore, MD4 Level 5, 5 Science Drive 2, Singapore 117597*
[5]*Infectious Diseases Translational Research Program, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Drive, Singapore 117597*
[6]*Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (ASTAR), 61 Biopolis Drive, Proteos #06-05, Singapore 138673*
[7]*NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, 28 Medical Drive, Singapore 117456*
[8]*Department of Biological Sciences, National University of Singapore, Singapore 117543*

Correspondence should be addressed to Yuin-Han Loh; yhloh@imcb.a-star.edu.sg

Received 28 September 2020; Revised 29 December 2020; Accepted 21 January 2021; Published 2 February 2021

Academic Editor: Christian Morsczeck

The regulation of mammalian stem cell fate during differentiation is complex and can be delineated across many levels. At the chromatin level, the replacement of histone variants by chromatin-modifying proteins, enrichment of specific active and repressive histone modifications, long-range gene interactions, and topological changes all play crucial roles in the determination of cell fate. These processes control regulatory elements of critical transcriptional factors, thereby establishing the networks unique to different cell fates and initiate waves of distinctive transcription events. Due to the technical challenges posed by previous methods, it was difficult to decipher the mechanism of cell fate determination at early embryogenesis through chromatin regulation. Recently, single-cell approaches have revolutionised the field of developmental biology, allowing unprecedented insights into chromatin structure and interactions in early lineage segregation events during differentiation. Here, we review the recent technological advancements and how they have furthered our understanding of chromatin regulation during early differentiation events.

1. Introduction

During natural development, embryonic stem cells progressively lose their pluripotency and upregulate cell fate specification markers, thereby producing hundreds of different cell types. The ability of a single cell to differentiate and give rise to the whole organism has fascinated biologists for decades. Epigenetic regulation, including histone modifications, histone variant substitutions, maternal factors, DNA methylation, and imprinting, plays a crucial role in the specification and determination of cell fate. Epigenetic factors can change chromosome conformation and the weak interacting forces [1], leading to differential gene expression across cell types. Molecular biology techniques such as fluorescence microscopy and RNA interference have only answered particular aspects of the underlying mechanisms. However, more

delicate approaches are required to solve increasingly sophisticated questions in the field. The discoveries of a totipotent subpopulation within mouse embryonic stem cell (mESCs) culture [2], expanded potential stem cells (EPSC) [3, 4], and induced pluripotent stem cells with higher potency [5] have reignited the interest in developing media that are capable of maintaining cells with increased differentiation potential. Studies suggest that such potential is linked to the bivalent chromatin [6, 7] and depletion of inhibitory markers that stabilise the cell fate [8]. The mESCs and primed human ESC (hESCs) are capable of differentiating into the trophoblast lineage upon manipulation [9, 10]. However, it remains unknown whether the transdifferentiation into the trophoblast lineage happens after the transition to the totipotent state [11] or induced directly from the alternate pluripotent state [12]. Recent developments in single-cell technology have allowed us to look deeper into cellular networks involving chromatin state and epigenetic regulators in early embryogenesis [13–15]. These proof of concept studies have showcased the potential of single-cell technology in meeting the needs of the field.

2. Single-Cell and Low-Input Techniques

Cellular heterogeneity primes cells towards different lineages and is difficult to study in the context of the embryogenesis. Traditional methods employing the expression of fluorescent proteins and observational studies by perturbing critical factors that are known to be involved in the formation of embryos are both time consuming and inefficient. Additionally, certain cell types with smaller population sizes are easily masked in the bulk analysis. Ever since the advent of single-cell technology in 2009 [16], which permitted the analysis of the mouse embryonic transcriptome, the field has quickly adapted this concept to questions highly relevant to epigenetic regulation. However, these methods remain technically challenging, especially during the process of amplifying signals from each cell while suppressing unspecific noises. Epigenetic studies often involve a bulk analysis of materials pooled together using millions of cells to derive the most accurate map, which is not practical in studies involving early embryos. To this end, various groups have employed different methods, such as multiple rounds of bar coding and specialised beads to improve capturing and accuracy of amplification of the epigenome [14, 17, 18] (Figure 1).

Chromatin accessibility reflects, to some degree, the expression status of genes by controlling the exposure of genomic regions to transcription factors (TFs) and other DNA-binding elements. There are currently four approaches to analyse chromatin accessibility in a single cell. Three of them quantify enrichment of DNA fragments after enzymatic DNA cleavage of accessible regions. The assay for transposase-accessible chromatin using sequencing (ATAC-seq) employs the hyperactive transposase Tn5 which simultaneously cleaves and inserts itself to the accessible regions and ligates sequencing indexes containing adaptors to these regions in each cell (Figure 1). The resultant DNA fragments are amplified *via* polymerase chain reaction (PCR), and short fragments are selected to remove partially digested fragments that are longer in length [19–21]. A second approach employs the so-called DNase I hypersensitive site sequencing (DNase-seq), whereby DNase-sensitive chromatin is cleaved and further processed with either type II restriction enzyme digestion or size selection to obtain fragments with appropriate sizes for sequencing [22, 23]. A third approach is labelled micrococcal nuclease sequencing (MNase-seq), whereby the DNA nuclease digests naked DNA and leaves DNA that binds to the nucleosomes intact, which allows profiling of the inaccessible chromatin in the cell [24]. Lastly, a fourth approach is the "nucleosome occupancy and methylome sequencing" (NOMe-seq), in which a GpC methyltransferase is used to mark accessible regions with GpC methylation (Figure 1). This is followed by bisulfite sequencing of non-methylated cytosine to obtain information on regions that are not protected by the nucleosomes [25, 26]. Recent advancements in single-cell chromatin accessibility assays involve combinations of multiple readouts to maximize information extracted from the same cell [27–29]. Each method comes with its own bias in enrichment or loss of signals. In addition, these approaches are costly, hence demanding careful consideration before embarking on the experiment.

Chromatin-immunoprecipitation-sequencing (ChIP-seq) is a commonly used technique to examine the interactions between protein and genomic DNA. Incorporating advancements of single-cell technologies, droplet-based single-cell ChIP-seq (DROP-ChIP/scChIP-seq) has since undergone rapid development and has been applied in many studies for understanding the heterogeneity within such cell populations comprehensively [18, 30] (Figure 1). Furthermore, multiple techniques such as the microfluidic-oscillatory-washing-based ChIP-seq (MOWChIP-seq), ultra-low-input native ChIP-seq (ULI-NChIP), and micro-ChIP (μChIP) have since been developed to overcome challenges that arise from low-input cell numbers and the scarcity of some tissue samples, allowing for high throughput evaluations of cell chromatin status [31–34] (Figure 1). A unique method that fuses an antibody to Tn5, termed CUT&RUN [35] or CUT&Tag [36] (Figure 1), has also opened new avenues in profiling the effects of chromatin remodelling complexes coupled with histone modifications, RNA polymerase II, and TFs in single cells [17].

Chromosome conformation capture or Hi-C is a method that enables the analysis of chromatin interactions (Figure 1). In Hi-C, interacting DNA fragments are ligated and sequenced to detect genome-wide long-range DNA interactions, which provides information on spatial arrangement and proximity of genes and their enhancers. Chromatins are partitioned into self-interacting active and silent topological associated domains (TADs), suggesting a relationship between gene activities and genome folding [37]. However, resolution remains a major issue for single-cell Hi-C over low-input Hi-C [38, 39].

3. Roles of Histone Variants on Chromatin Remodelling during Differentiation

Extensive rewiring in chromatin regulation, including histone modifications, expression and binding of TFs, and

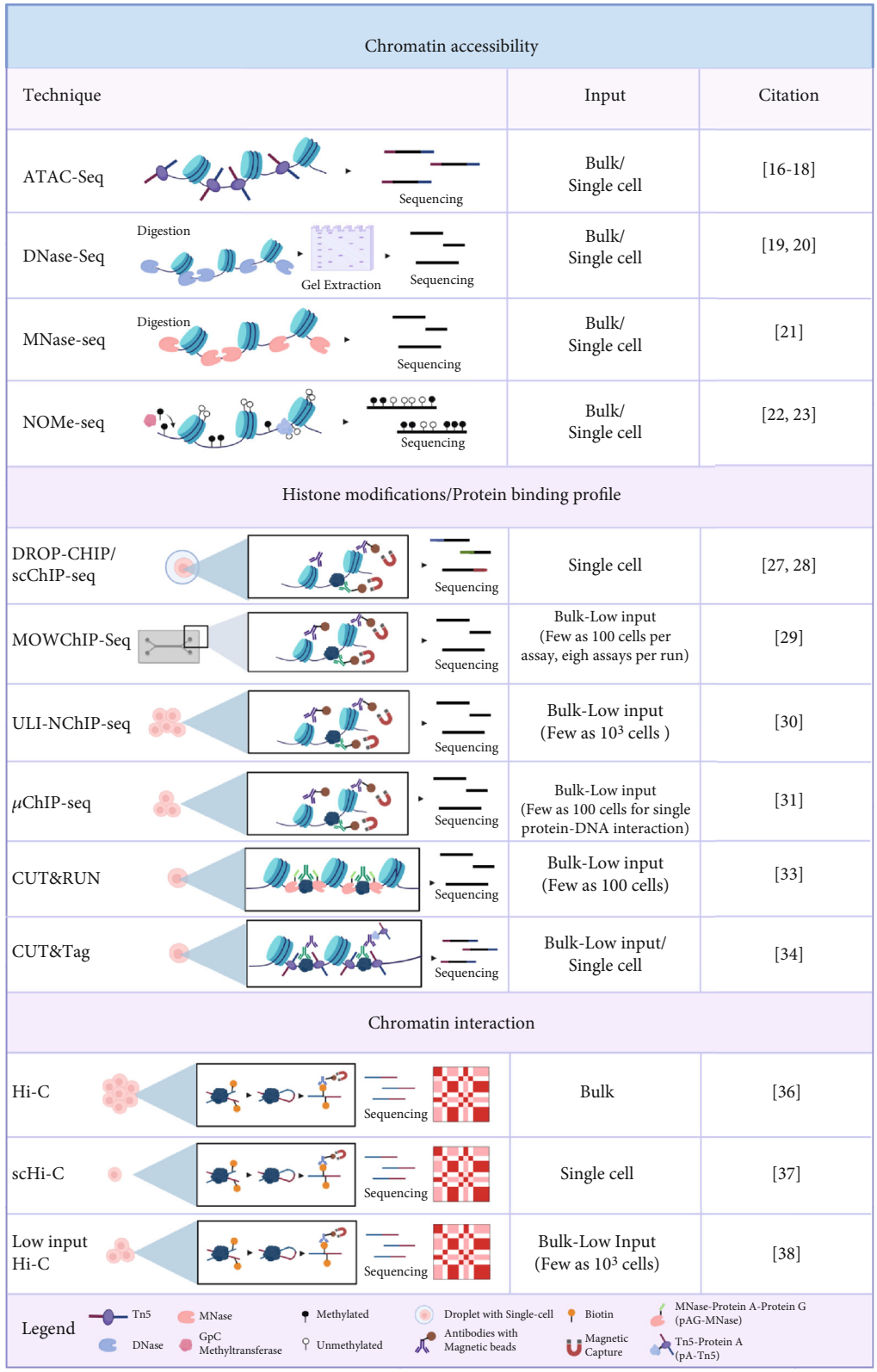


Figure 1: Summary of the comparison of different single-cell and low-input techniques to assess chromatin structure [16–23, 27–31, 33, 34, 36–38]. Created with http://BioRender.com/.

genomic interactions, happens during differentiation. Here, we evaluate the roles of epigenetic factors in chromosome remodelling during differentiation, as well as the differences in the core regulatory network in the transition of human and mouse ESCs to trophoblast stem cells (TSCs) (Figure 2).

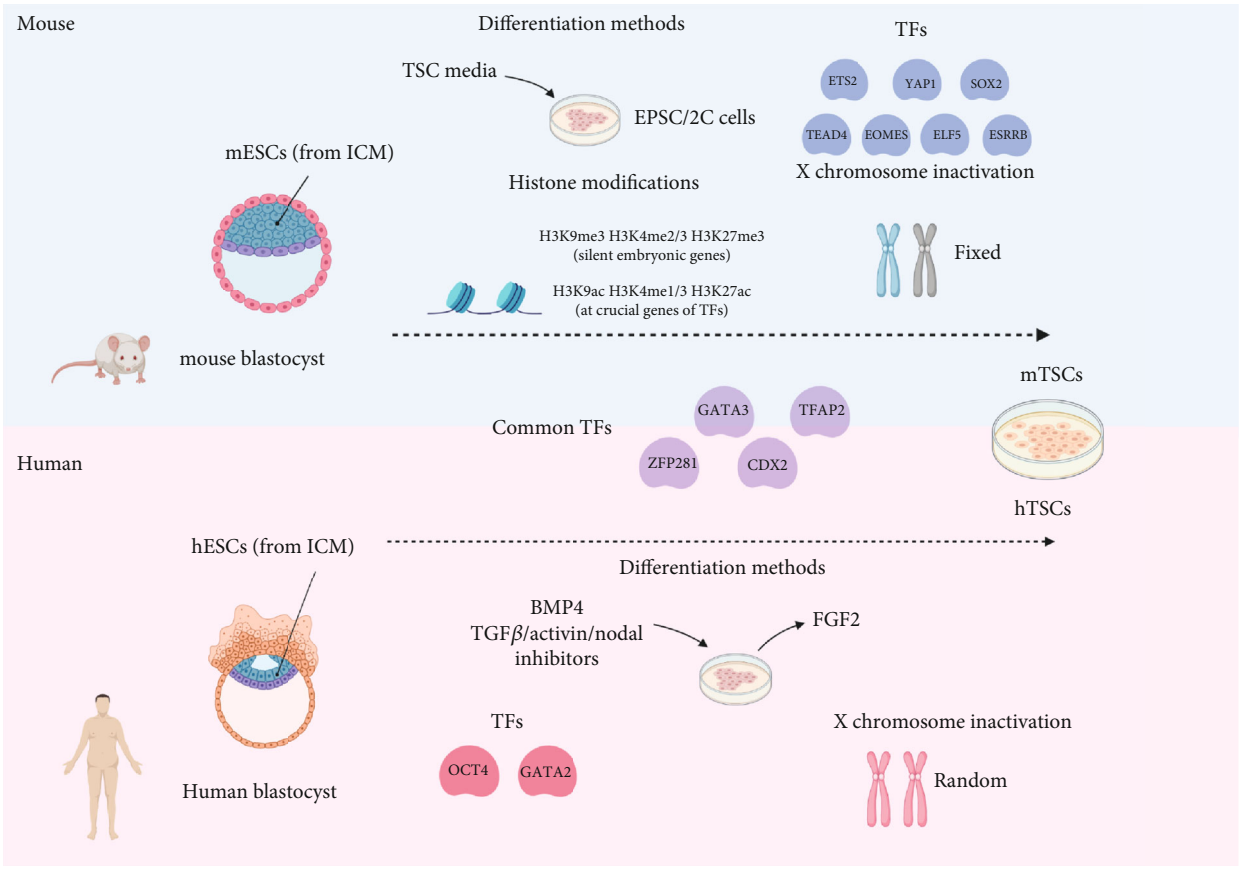


Figure 2: Summary of the comparison in deriving mouse and human ESCs and TSCs from mouse and human ESCs [10, 15, 57, 58, 65, 69–72, 75, 76, 82–85]. Created with http://BioRender.com/.

Chromatin structure is based on the coiling and positioning of the nucleosome, which is made up of two identical subunits consisting of four histone proteins, H2A, H2B, H3, and H4, while the H1 histone binds to linker DNA. After fusion of the two germ cells into a single zygote, the histone composition undergoes rapid changes to be replaced by newly synthesized canonical histones. It has been implicated that expression of zygotic genes is independent of higher order chromatin structure [40, 41]. Cell fate then appears to be marked as early as the 4-cell stage by the core pluripotent markers [42–44]. During the course of embryogenesis, the chromatin progressively loses its open state, gaining a more condensed conformation.

The roles of noncanonical histones have been widely implicated in stem cell differentiation. In hESCs, depletion of histone 3 variant centromere protein A (CENP-A) has no effect on the self-renewal of stem cells but causes cell cycle arrest at the G1 during differentiation. It also impacts the repair mechanism of the stem cell, leading to apoptosis. Whereas in fibroblasts, depletion of CENP-A leads to increased apoptosis and reduced self-renewal capacity [45]. It remains unknown how centromeres are regulated by CENP-B, CENP-C, and CENP-T during the differentiation and self-maintenance of stem cells.

Investigations into histone variant H3.3 have uncovered its crucial role in differentiation, cell fate transition, and the maintenance of heterochromatin integrity at the centromeres, telomeres, and pericentromeric sites [46]. In particular, the H3.3 lysine 4 residue is associated with enhancers and promoters of active genes, facilitating nucleosome deposition, histone replacement, and binding of chromatin remodelers at those sites [47].

On the other hand, the histone 2 variant H2A.Z is essential in marking genes to be downregulated during differentiation by interacting with polycomb repressive complex 2 (PRC2) to deposit repressive H3K27me3 marks [48]. It is enriched at active enhancers and promoters, affecting the accessibility of the transcription start site to the transcription factors [49, 50]; H2A.Z also interacts with lysine acetylation marks on H3 and CHD4 to remodel chromosomes during stem cell maintenance and differentiation [51, 52].

Each species has its own unique H1 variants serving different functions [53]. There are limited studies in this area, and it is currently thought that H1 controls chromatin compaction by regulating H2AK119ub1 during mESC differentiation [54].

4. Histone Modifications

There have been extensive studies on post-translational modifications of Histone 3, which have shown that the pattern of histone modifications is expressed in a lineage-specific

manner in the ESC and TSC state. Bivalent marks, namely the active marker H3K4me3 and repressive marker H3K27me3, are unique characteristics in ESCs [6]. These marks poise genes that are expressed when ESCs are committed to lineage specification, and their roles have been studied for a long time. Recent evidence suggests their crucial role in remodelling chromosome accessibility and chromatin looping [55]. However, their specific functions remain largely unknown [56].

Genome-wide analysis performed by Rugg-Gunn et al. suggests that H3K27me3 and H3K9ac levels are higher in the inner cell mass as compared to the trophoblast lineage, although there is no direct evidence to support the causative relationship between the two [57]. Additionally, either trivalent histone marks such as H3K9me3, H3K4me2/3, and H3K27me3 or bivalent histone marks can be adopted in silencing embryonic genes in cells developing into the trophoblast lineage [58] (Figure 2). CDX2 and EOMES are crucial TFs in the establishment of the TSC cell fate and are enriched with active histone marks such as H3K9ac and H3K4me3 while having lower levels of repressive histone marks [57]. In another study, inducing *CDX2* expression resulted in decreases in the expression of pluripotent genes *OCT4* and *NANOG*, increases in trophoblast lineage genes, and the differentiation of TSCs in the mouse embryo [59].

Moreover, it has been reported that histone methyltransferase *SUV39H1* mediated trimethylation on H3K9 is attributed to the suppression of embryonic genes in TSCs [60]; H3K9me3 also interacts with heterochromatin protein 1 to condense and silence different gene sets during differentiation in hESC and mESC [61, 62], highlighting the indispensable role of histone modifications in the regulation of lineage-specific genes.

Enrichment of H4K20me3 during differentiation leads to formation of pericentric heterochromatin by acting with *SMYD5*, and it has been shown that reduced transcription of endogenous long interspersed nuclear elements (LINEs) and long terminal repeats (LTRs) is key in maintaining pluripotency [63].

5. Comparison in the Development of Human and Mouse Trophoblast-Related Lineage

In the mouse embryo, specification starts at the 4-cell stage [43], whereas current evidence implies that such specification occurs in the early blastocyst stage in human embryos [64, 65]. In the mouse embryo, implantation is initiated by the mural trophectoderm (TE) followed by the polar TE. In the human embryo, implantation is initiated by the polar TE. The TE layer in human and mouse embryo subsequently matures to give rise to the syncytiotrophoblast (ST) and the extravillous cytotrophoblast (EVT) *via* cell fusion and endoreduplication, respectively. The mouse TE subsequently forms three distinct layers of trophoblast derivatives, separating maternal and fetal blood, whereas in human trophoblast analogs, a different structure is formed with only one layer separating maternal and fetal blood [66]. While there are studies aiming at establishing three-dimensional [67] and two-dimensional trophoblast cultures [68] that each are able to differentiate into both the ST and EVT lineages, there is a lack of studies looking into the role of chromatin remodelling and epigenetic regulation in such models.

The similarities and differences in human and mouse TSCs are well manifested through the aforementioned aspect of physiology. While most of the discussion is focused on the signalling pathway that contributes to the successful differentiation from ESCs to TSCs, the underlying conservativeness in the regulation of chromatin and binding of specific transcription factors is still crucial for the transcriptional network that drives the specification of TSCs.

6. Expression of Transcription Factors and Their Binding to Genomic DNA Regions

Transcription factors are known to be bound to specific genes to regulate gene expression directly or indirectly by recruiting other transcription factors (or repressors), or histone modifiers to activate or silence genes. *ZFP281* was identified as a conserved factor critical to the maintenance of human and mouse TSCs. By interacting with MLL and COMPASS subunits and binding to the promoters of target genes, *ZFP281* helps to establish the specific transcriptome necessary for differentiation and specification of mouse TSCs. Moreover, it has been demonstrated that *ZFP281* facilitates the induction of trophoblast stem-like cells from mouse embryonic stem cells upon overexpression. In humans, *ZFP281* helps to stabilise the transcriptome in undifferentiated TSCs [69].

Mouse TSC determination involves genes such as *TEAD4*, *CDX2*, *SOX2*, *ESRRB*, *TFAP2*, *ETS2*, *ELF5*, *GATA3*, and *YAP1* (Figure 2), although it is not known how all these genes interact in this context [10]. On the other hand, a group has recently identified the generation of human induced TSCs through stepwise or direct reprogramming of human dermal fibroblast. TE-associated transcription factors, *TFAP2C* and *GATA2*, are significantly upregulated during reprogramming to naïve state, and supporting their reprogramming to iTSCs [15].

CDX2 is expressed as early as the 8-cell stage and plays a critical role during the differentiation of cells into the TE and subsequent regulation of TE functions. However, it is not essential for the initiation of TE lineage segregation as *CDX2* knock out embryos retained the ability to form blastocoel cavities, implying that other key genes regulated this process. One such gene could be *TEAD4*, where knockout cells failed to differentiate into the TE, and *TEAD4* knockout embryos were unable to develop into blastocysts [70–72]. While expression of *OCT4* and *CDX2* is critical in the human TE, *OCT4* is depleted in the differentiating mouse TE [73]. Contrastingly, in human embryos, *CDX2* is only expressed after blastocyst formation [74].

ELF5 has been described as one of the core genes that regulate the self-renewal and differentiation of TSCs. It interacts with EOMES to recruit TFAP2C to TSC-specific genes, thereby inducing their expression in mouse TSCs [75]. Moreover, *Elf5* was found to be methylated and repressed in mESCs but hypomethylated and activated in mTSCs. It promotes the expression of a network of TFs, including *CDX2* and *EOMES*, that drives the efficient differentiation of ESCs to trophoblast-related lineages [65]. The GATA2/3-TFAP2A/C network was

enriched in regions of inactive placental and pluripotent genes in hESCs after treatment with BMP4, which induces trophoblast-specific genes and suppresses pluripotency during the initial stage of trophoblast differentiation [76].

Super-enhancers (SEs) are also one of the elements that model the transcriptional network. By mapping these SEs in mTSCs, more than 150 TFs, excluding master TFs such as *CDX2*, *GATA2*, and *TEAD4*, were identified as potential contributors to the TE lineage. This approach opens up a new aspect to further elucidate the mechanisms and regulators of mTSC lineage specification [77]. Additionally, it has been proposed that *ESRRB* could directly regulate the core genes of the TSC self-renewal regulatory network such as *CDX2*, *EOMES*, and *SOX2* [78]. Members of the ERV family RLTR13D5 could also act as enhancers; they are bound by H3K4me1 and H3K27ac, therefore providing binding sites for *CDX2*, *EOMES*, and *ELF5* [79].

7. X-Linked Genes

Studies revolving around long noncoding RNA (lncRNA) have shed some light in their roles in regulation of the stem cell pluripotency and lineage segregation. lncRNA recruits chromatin modifiers such as mixed-lineage leukemia 1 (MLL1) and *PRC2* to modulate chromatin structure and gene expression [80, 81]. The study of X-linked genes patterning and X chromosome inactivation by lncRNA X-inactive-specific transcript (XIST) has provided some clues to early developmental events. During lineage segregation in the female mouse embryo, paternal X chromosome is first inactivated, contributing to the TE lineage, followed by reactivation in the inner cell mass (ICM) and finally random X chromosome inactivation. Whereas in the human female embryo, random X chromosome inactivation first occurs in cells contributing to the TE, followed by a second wave of random X chromosome inactivation in ICM [82]. The inactivation is initiated by expression of *XIST* and accompanied by the recruitment of multiple chromatin modifiers to suppress the expression of extra X-linked genes [83]. In contrast to two distinct lineage segregation events in the mouse blastocyst, evidence suggests that the TE, epiblast, and primitive endoderm might arise simultaneously during a single event in human blastocysts [84, 85] (Figure 2).

8. Transposable Elements Function in TSC and ESC

Transposable elements account for at least 40% of the human or mouse genomes [86, 87]. Previously regarded as "junk DNA", it was recently discovered that transposable elements adopt functional roles akin to enhancers, promoters, or insulators, which are essential to gene regulation [88]. Therefore, it is important to explore their regulatory roles in TSCs and ESCs.

Transposable elements have contributed greatly to the gene regulatory network in different lineages or cell types [90]. To explore the overall pattern of different epigenetic modifications that accompany transposable elements, we analysed ATAC-seq data [91], histone modification data including H3K27ac, H3K4me1, H3K4me3, H3K27me3, and H3K9me3 [92], H3K36me3 and H4K20me1 [93], H2BK5ac [94], datasets regarding transcription factors such as *P300* [77], *SOX2* [92], *ELF5*, *EOMES*, and *CDX2* [92], *TET1* [95], *CTCF*, *SP1*, and *TBP* [93], and *LSD1* [96] in mouse TSCs. TE family enrichment analysis were done on these data using the same method suggested by the Cao's team [89]. Result (Figure 3(a)) shows that the endogenous retrovirus-like elements (ERVs) such as the ERVK and ERV1 families are significantly enriched in the open regions of mTSCs and bound by critical TSC-related TFs. Furthermore, transposable elements such as B2, Alu, and MIR (Mammalian-wide interspersed repeats) are bound by active histone marks such as H3K4me1 and H3K27ac, implying possible functions as enhancers. Promoters are conserved across species, whereas enhancers are found to be specific to different organisms or cell types. As enhancers are known to regulate tissue- or cell type-specific gene expression, we overlapped the TE sites with enhancers defined by P300 and H3K27ac. The TE-derived enhancers such as ERVK and ERV1 were significantly bound by transcription factors *SOX2*, *LSD1*, *EOMES*, and *ELF5*. Given the functions of the factors discussed in earlier sections, the analysis suggests that these repeats could act as enhancers to regulate gene expression in TSCs. RLTR13D5 containing ERVK-derived enhancers echoes the significance of ERV in the mTSCs by acting as enhancers and binding sites for TSC-specifying TFs [79, 97].

Understanding the conservation of chromatin accessibility across hTSCs and mTSCs might provide novel insights into their differences. To this end, we analysed ATAC-seq datasets from naïve hESCs, primed hESCs, blastocyst-derived TSCs, and naïve hESCs [98, 99]. The TE families enrichment analysis shows that ERVK and ERV1 were significantly enriched in hESCs and hTSCs (Figure 3(b)), suggesting that ERVK might play conserved and functional roles in TSCs in both species. There are also both unique open ERVKs and shared open ERVKs in ESC and TSC. From the motif analysis, ERVKs with open chromatin state in hTSCs are enriched for TSC-related transcription factor motifs such as *TEAD4* and *GATA3*, suggesting that these ERVKs might have been adopted during evolution to cooperate with TSC-specific transcription factors to regulate transcriptional networks essential for TSC.

Apart from expression of TFs and chromatin accessibilities, recent Hi-C data has revealed the divergence in the repressive and active chromatin interaction between mouse ESCs and TSCs lineages. TSCs genes, which are repressed in ESCs, interact with H3K27me3 associated regions in ESCs through the PRC. Furthermore, enhancer-gene interactions involving key TSC transcription factors are particularly enriched to maintain the expression of TSC-genes [100]. Another recent report correlates the chromatin modifier known as the ChAHP complex (CHD4, ADNP, and HP1) with proper cell differentiation. This complex competes with CTCF binding sites and modulates the formation of TADs in proximal regions, specifically at conserved SINE B2-transposable elements [101]. The role of the ChAHP subunits CHD4 and HP1 in stem cell maintenance and differentiation has been previously reported [51, 61].

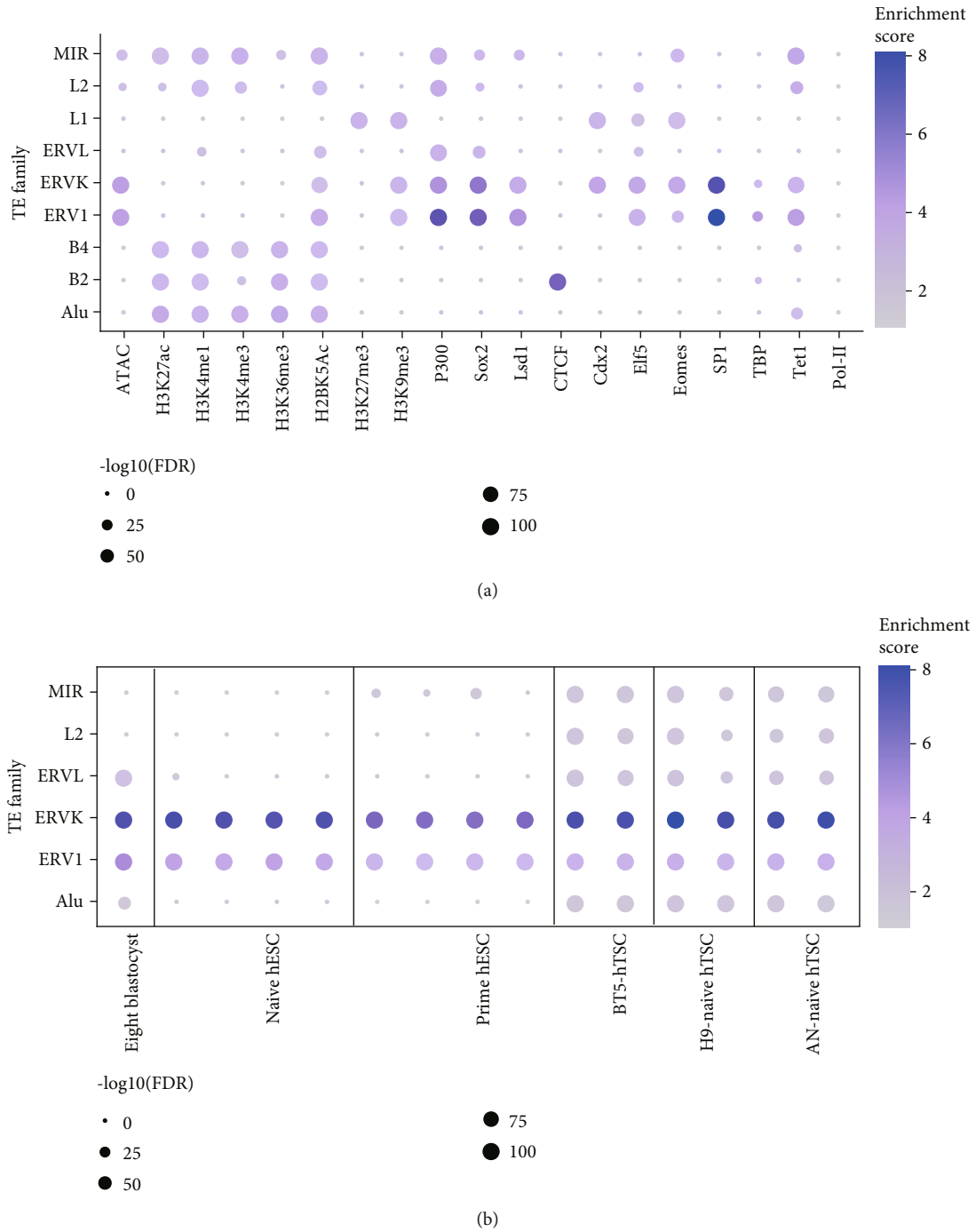


Figure 3: Transposable elements are marked by epigenetic signatures. (a) Dot-plot of the enrichment of transposable elements families in 8 chromatin marks and 11 bound factors in mouse TSCs. The size of the circle represents corrected enrichment *P* values. The colour indicates the enrichment score which was computed with a combination of the binomial test and hypergeometric test [89]. (b) Dot-plot of the enrichment of transposable elements families in open chromatin regions defined by ATAC-seq peaks in human eight-stage blastocysts, naïve ESCs, primed ESCs, blastocyst-derived TSCs, H9-derived TSCs, and AN1 iPSC-derived TSCs. The size of the circle represents corrected enrichment *P* values. The colour indicates the enrichment score which was computed with a combination of the binomial test and hypergeometric test [89].

To identify the target genes regulated by transposable elements, Hi-C or promoter capture techniques could be used to check the putative targets of these TE-derived enhancers. CRISPR interference methods could be used to disrupt the transposable elements followed by validation using RNA-seq or qPCR analysis to check the expression of the putative target genes. As demonstrated by Todd et al. [97], only small subsets of transposable elements are crucial in regulating the

TSC and ESC gene expression. Therefore, it is important to pinpoint those that have been adopted as functional gene regulatory elements during evolution critical in each cell type.

9. Future Perspective

There are numerous pluripotent states reported in human and mouse ESCs, with the most common ones being the naïve and primed state. There are a multitude of studies that attempted to differentiate primed hESC from TSC with varying scales of success [10]. It was reported that during differentiation of human ESCs towards TSCs, FGF2 should be removed completely from the media, and BMP4 and TGFβ/activin/nodal inhibitors should be added as supplements. The size of the initial colonies also affects the outcome of the differentiation process. Meanwhile, 2C-like cells [2] and EPSCs [4] are the only two reported sources of mouse stem cells that are capable of differentiating into TSC *in vitro*, making it worthwhile to dissect the mechanism underlying the derivation of TSCs in the respective state.

It has been established that hESCs exist in the primed pluripotent state with one active X chromosome and one inactive X chromosome. This resembles a closer gene expression and signaling profile to primed mouse epiblast stem cells (mEpiSCs) than mouse ESCs, which is considered an earlier stage of naïve pluripotent state with two active X chromosomes [102, 103]. In humans, naïve pluripotent stem cells express TFs and display open chromatin structure associated with cells from trophoblast-related lineages, which were conversely reported to be able to give rise to self-renewing TSCs, a feat which is unachievable with primed hESCs that are exposed to the same differentiation assays [98, 104]. Similar phenomena were observed in the mouse, when overexpression of *CDX2* in the naïve mESCs drove the cells towards a TSC-like cell fate, but not mEpiSCs [105].

Early studies characterizing hESC-derived trophoblast-like cells focused on human chorionic gonadotropin production and cellular invasion capacity. While some studies claim that mouse or human TSCs derived *in vitro* closely resemble their *in vivo* counterparts, others have provided contradictory results [12, 98]. This might be due to differences in the parameters used by each group during cell type characterization and culturing, as studies have shown that differences in starting colony and chemical providers could drastically alter the results [10]. It will be interesting to apply novel single-cell technologies to improve the characterization and understanding of cellular heterogeneity and help to reconstruct a clearer picture of cellular processes, including chromatin remodelling events during changes in the cell fate.

10. Conclusion

In the last three decades, a considerable amount of effort has been invested to our understanding and capturing cells in different pluripotency states ranging from TSCs, expended potential, 2C-like, naïve, prime, Rosette [106], Founder [107], and many more. Researchers have employed a wide range of methods to delineate their differences and analogues *in vivo* and across different animal species. While single-cell RNA-seq datasets have provided insights into the transcriptome of different cell types and revealed details on rare populations and the trajectory of cells during differentiation, this information is often limited and does not provide sufficient data to derive the factors and mechanisms controlling the specification and determination of each cell type.

Although pluripotency circuitry has been well studied, novel stem cell populations and pluripotent stages are consistently being reported. The ability of cells to form blastocyst-like structures [108] to investigate cell fate changes *ex vivo* has recently gained vast interest. Transposable elements, previously disregarded as an unimportant part of the genome, proved to be essential in controlling totipotency in the mouse, while showing differentially binding to pluripotent and TSC-specifying genes. There is still a broad gap in knowledge regarding the epigenome within each cell in early embryogenesis, priming them to different fates under the same condition.

Conflicts of Interest

The authors declare no conflict of interests. The data used for analysis were sourced from Gene Expression Omnibus. Search was conducted for human and mouse genomic data with relevant key terms such as "ATAC," "ChIP-seq," "ESC," and "TSC." Literature search was conducted using PubMed, Google Scholar, Nature, Cell Press, Science Magazine with a combination of keywords such as "Single-Cell," "Embryogenesis," "Stem cell differentiation," "Trophoblast differentiation," "Histone variant," "Transposable element," "Totipotency," "Transdifferentiation," and "Histone modifier." The search results were considered based on novelty, potential impact, and possible applications.

Authors' Contributions

Zi Hao Zheng and Tsz Wing Sam contributed equally to this work.

Acknowledgments

The authors would like to thank Katherine Anne Lau, Rachel Jun Rou Tan, and Peter Dröge for their editorial suggestions. Figures 1 and 2 are created with http://BioRender.com/. Y.H.L. is supported by the National Research Foundation Investigatorship Award (NRFI2018-02), Agency for Science, Technology and Research JCO Development Program Grant (1534n00153 and 1334k00083), and Singapore National Research Foundation under its Cooperative Basic Research grant administered by the Singapore Ministry of Health's National Medical Research Council (NMRC/CBRG/0092/2015). We thank the Biomedical Research Council, Agency for Science, Technology and Research, Singapore, for research funding.

References

[1] O. L. Kantidze and S. V. Razin, "Weak interactions in higher-order chromatin organization," *Nucleic Acids Research*, vol. 48, no. 9, pp. 4614–4626, 2020.

[2] T. S. Macfarlan, W. D. Gifford, S. Driscoll et al., "Embryonic stem cell potency fluctuates with endogenous retrovirus activity," *Nature*, vol. 487, no. 7405, pp. 57–63, 2012.

[3] Y. Yang, B. Liu, J. Xu et al., "Derivation of pluripotent stem cells with in vivo embryonic and extraembryonic potency," *Cell*, vol. 169, no. 2, pp. 243–257.e25, 2017.

[4] J. Yang, D. J. Ryan, W. Wang et al., "Establishment of mouse expanded potential stem cells," *Nature*, vol. 550, no. 7676, pp. 393–397, 2017.

[5] M. Abad, L. Mosteiro, C. Pantoja et al., "Reprogramming *in vivo* produces teratomas and iPS cells with totipotency features," *Nature*, vol. 502, no. 7471, pp. 340–345, 2013.

[6] B. E. Bernstein, T. S. Mikkelsen, X. Xie et al., "A bivalent chromatin structure marks key developmental genes in embryonic stem cells," *Cell*, vol. 125, no. 2, pp. 315–326, 2006.

[7] M. Minoux, S. Holwerda, A. Vitobello et al., "Gene bivalency at polycomb domains regulates cranial neural crest positional identity," *Science*, vol. 355, no. 6332, article eaal2913, 2017.

[8] P. A. Latos, C. Helliwell, O. Mosaku et al., "NuRD-dependent DNA methylation prevents ES cells from accessing a trophectoderm fate," *Biology Open*, vol. 1, no. 4, pp. 341–352, 2012.

[9] M. Hemberger, T. Nozaki, E. Winterhager et al., "*Parp1*-deficiency induces differentiation of ES cells into trophoblast derivatives," *Developmental Biology*, vol. 257, no. 2, pp. 371–381, 2003.

[10] R. M. Roberts, T. Ezashi, M. A. Sheridan, and Y. Yang, "Specification of trophoblast from embryonic stem cells exposed to BMP4†," *Biology of Reproduction*, vol. 99, no. 1, pp. 212–224, 2018.

[11] X. Fu, M. N. Djekidel, and Y. Zhang, "A transcriptional roadmap for 2C-like–to–pluripotent state transition," *Science Advances*, vol. 6, no. 22, article eaay5181, 2020.

[12] E. Posfai, J. P. Schell, A. Janiszewski et al., "Evaluating totipotency using criteria of increasing stringency," *Nature Cell Biology*, vol. 43, no. 1, 2021.

[13] Z. Hu, D. E. K. Tan, G. Chia et al., "Maternal factor NELFA drives a 2C-like state in mouse embryonic stem cells," *Nature Cell Biology*, vol. 22, no. 2, pp. 175–186, 2020.

[14] S. Collombet, N. Ranisavljevic, T. Nagano et al., "Parental-to-embryo switch of chromosome organization in early embryogenesis," *Nature*, vol. 580, no. 7801, pp. 142–146, 2020.

[15] X. Liu, J. F. Ouyang, F. J. Rossello et al., "Reprogramming roadmap reveals route to human induced trophoblast stem cells," *Nature*, vol. 586, no. 7827, pp. 101–107, 2020.

[16] F. Tang, C. Barbacioru, Y. Wang et al., "mRNA-Seq whole-transcriptome analysis of a single cell," *Nature Methods*, vol. 6, no. 5, pp. 377–382, 2009.

[17] S. J. Hainer, A. Bošković, K. N. McCannell, O. J. Rando, and T. G. Fazzio, "Profiling of pluripotency factors in single cells and early embryos," *Cell*, vol. 177, no. 5, pp. 1319–1329.e11, 2019.

[18] A. Rotem, O. Ram, N. Shoresh et al., "Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state," *Nature Biotechnology*, vol. 33, no. 11, pp. 1165–1172, 2015.

[19] J. D. Buenrostro, B. Wu, U. M. Litzenburger et al., "Single-cell chromatin accessibility reveals principles of regulatory variation," *Nature*, vol. 523, no. 7561, pp. 486–490, 2015.

[20] D. A. Cusanovich, R. Daza, A. Adey et al., "Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing," *Science*, vol. 348, no. 6237, pp. 910–914, 2015.

[21] B. Pijuan-Sala, N. K. Wilson, J. Xia et al., "Single-cell chromatin accessibility maps reveal regulatory programs driving early mouse organogenesis," *Nature Cell Biology*, vol. 22, no. 4, pp. 487–497, 2020.

[22] W. Jin, Q. Tang, M. Wan et al., "Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples," *Nature*, vol. 528, no. 7580, pp. 142–146, 2015.

[23] J. Cooper, Y. Ding, J. Song, and K. Zhao, "Genome-wide mapping of DNase I hypersensitive sites in rare cell populations using single-cell DNase sequencing," *Nature Protocols*, vol. 12, no. 11, pp. 2342–2354, 2017.

[24] W. Gao, B. Lai, B. Ni, and K. Zhao, "Genome-wide profiling of nucleosome position and chromatin accessibility in single cells using scMNase-seq," *Nature Protocols*, vol. 15, no. 1, pp. 68–85, 2020.

[25] T. K. Kelly, Y. Liu, F. D. Lay, G. Liang, B. P. Berman, and P. A. Jones, "Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules," *Genome Research*, vol. 22, no. 12, pp. 2497–2506, 2012.

[26] F. Guo, L. Li, J. Li et al., "Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells," *Cell Research*, vol. 27, no. 8, pp. 967–988, 2017.

[27] S. J. Clark, R. Argelaguet, C. A. Kapourani et al., "scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells," *Nature Communications*, vol. 9, no. 1, p. 781, 2018.

[28] J. Cao, D. A. Cusanovich, V. Ramani et al., "Joint profiling of chromatin accessibility and gene expression in thousands of single cells," *Science*, vol. 361, no. 6409, pp. 1380–1385, 2018.

[29] Q. R. Xing, C. A. E. Farran, Y. Y. Zeng et al., "Parallel bimodal single-cell sequencing of transcriptome and chromatin accessibility," *Genome Research*, vol. 30, no. 7, pp. 1027–1039, 2020.

[30] Q. Wang, H. Xiong, S. Ai et al., "CoBATCH for high-throughput single-cell epigenomic profiling," *Molecular Cell*, vol. 76, no. 1, pp. 206–216.e7, 2019.

[31] B. Zhu, Y. P. Hsieh, T. W. Murphy, Q. Zhang, L. B. Naler, and C. Lu, "MOWChIP-seq for low-input and multiplexed profiling of genome-wide histone modifications," *Nature Protocols*, vol. 14, no. 12, pp. 3366–3394, 2019.

[32] J. Brind'Amour, M. Hudson, C. Chen, M. M. Karimi, M. C. Lorincz, and M. C. Lorincz, "An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations," *Nature Communications*, vol. 6, no. 1, p. 6033, 2015.

[33] J. A. Dahl and P. Collas, "A rapid micro chromatin immunoprecipitation assay (ChIP)," *Nature Protocols*, vol. 3, no. 6, pp. 1032–1045, 2008.

[34] A. Sankar, M. Lerdrup, A. Manaf et al., "KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3 domains from H3K9me3 invasion in oocytes," *Nature Cell Biology*, vol. 22, no. 4, pp. 380–388, 2020.

[35] P. J. Skene and S. Henikoff, "An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites," *eLife*, vol. 6, article e21856, 2017.

[36] H. S. Kaya-Okur, S. J. Wu, C. A. Codomo et al., "CUT&Tag for efficient epigenomic profiling of small samples and single cells," *Nature Communications*, vol. 10, no. 1, p. 1930, 2019.

[37] T. Nagano, Y. Lubling, T. J. Stevens et al., "Single-cell Hi-C reveals cell-to-cell variability in chromosome structure," *Nature*, vol. 502, no. 7469, pp. 59–64, 2013.

[38] M. Chen, Q. Zhu, C. Li et al., "Chromatin architecture reorganization in murine somatic cell nuclear transfer embryos," *Nature Communications*, vol. 11, no. 1, p. 1813, 2020.

[39] X. Chen, Y. Ke, K. Wu et al., "Key role for CTCF in establishing chromatin structure in human embryos," *Nature*, vol. 576, no. 7786, pp. 306–310, 2019.

[40] C. W. Hanna, H. Demond, and G. Kelsey, "Epigenetic regulation in development: is the mouse a good model for the human?," *Human Reproduction Update*, vol. 24, no. 5, pp. 556–576, 2018.

[41] Z. du, H. Zheng, B. Huang et al., "Allelic reprogramming of 3D chromatin architecture during early mammalian development," *Nature*, vol. 547, no. 7662, pp. 232–235, 2017.

[42] M. D. White, J. F. Angiolini, Y. D. Alvarez et al., "Long-lived binding of Sox2 to DNA predicts cell fate in the four-cell mouse embryo," *Cell*, vol. 165, no. 1, pp. 75–87, 2016.

[43] M. Goolam, A. Scialdone, S. J. L. Graham et al., "Heterogeneity in Oct4 and Sox2 targets biases cell fate in 4-cell mouse embryos," *Cell*, vol. 165, no. 1, pp. 61–74, 2016.

[44] S. Guo, X. Cui, X. Jiang et al., "Tracing the origin of the placental trophoblast cells in mouse embryo development†," *Biology of Reproduction*, vol. 102, no. 3, pp. 598–606, 2019.

[45] G. Ambartsumyan, R. K. Gill, S. D. Perez et al., "Centromere protein A dynamics in human pluripotent stem cell self-renewal, differentiation and DNA damage," *Human Molecular Genetics*, vol. 19, no. 20, pp. 3970–3982, 2010.

[46] H.-T. Fang, C. A. el Farran, Q. R. Xing et al., "Global H3.3 dynamic deposition defines its bimodal role in cell fate transition," *Nature Communications*, vol. 9, no. 1, p. 1537, 2018.

[47] M. Gehre, D. Bunina, S. Sidoli et al., "Lysine 4 of histone H3.3 is required for embryonic stem cell differentiation, histone enrichment at regulatory regions and transcription accuracy," *Nature Genetics*, vol. 52, no. 3, pp. 273–282, 2020.

[48] R. Faast, V. Thonglairoam, T. C. Schulz et al., "Histone variant H2A.Z is required for early mammalian development," *Current Biology*, vol. 11, no. 15, pp. 1183–1187, 2001.

[49] G. Hu, K. Cui, D. Northrup et al., "H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation," *Cell Stem Cell*, vol. 12, no. 2, pp. 180–192, 2013.

[50] Z. Li, P. Gadue, K. Chen et al., "Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation," *Cell*, vol. 151, no. 7, pp. 1608–1616, 2012.

[51] H. Zhao, Z. Han, X. Liu et al., "The role of Chd4 in the regulation of ESC self-renewal," *The Journal of Biological Chemistry*, vol. 292, no. 20, pp. 8507–8519, 2017.

[52] C. C. Hsu, D. Zhao, J. Shi et al., "Gas41 links histone acetylation to H2A.Z deposition and maintenance of embryonic stem cell identity," *Cell Discovery*, vol. 4, no. 1, p. 28, 2018.

[53] S. P. Hergeth and R. Schneider, "The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle," *EMBO Reports*, vol. 16, no. 11, pp. 1439–1453, 2015.

[54] J. Zhao, M. Wang, L. Chang et al., "RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division," *Nature Cell Biology*, vol. 22, no. 4, pp. 439–452, 2020.

[55] G. Mas, E. Blanco, C. Ballaré et al., "Promoter bivalency favors an open chromatin architecture in embryonic stem cells," *Nature Genetics*, vol. 50, no. 10, pp. 1452–1462, 2018.

[56] P. Voigt, W.-W. Tee, and D. Reinberg, "A double take on bivalent promoters," *Genes & Development*, vol. 27, no. 12, pp. 1318–1338, 2013.

[57] P. J. Rugg-Gunn, B. J. Cox, A. Ralston, and J. Rossant, "Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 24, pp. 10783–10790, 2010.

[58] C. E. Senner and M. Hemberger, "Regulation of early trophoblast differentiation - lessons from the mouse," *Placenta*, vol. 31, no. 11, pp. 944–950, 2010.

[59] T. S. Carey, I. Choi, C. A. Wilson, M. Floer, and J. G. Knott, "Transcriptional reprogramming and chromatin remodeling accompanies Oct4 and Nanog silencing in mouse trophoblast lineage," *Stem Cells and Development*, vol. 23, no. 3, pp. 219–229, 2014.

[60] O. Alder, F. Lavial, A. Helness et al., "Ring1B and Suv39h1 delineate distinct chromatin states at bivalent genes during early mouse lineage commitment," *Development*, vol. 137, no. 15, pp. 2483–2492, 2010.

[61] A. Mattout, B. S. Sailaja, E. V. Raghu Ram et al., "Heterochromatin protein 1β (HP1β) has distinct functions and distinct nuclear distribution in pluripotent versus differentiated cells," *Genome Biology*, vol. 16, no. 1, pp. 213–213, 2015.

[62] J. L. Golob, S. L. Paige, V. Muskheli, L. Pabon, and C. E. Murry, "Chromatin remodeling during mouse and human embryonic stem cell differentiation," *Developmental Dynamics: an official publication of the American Association of Anatomists*, vol. 237, no. 5, pp. 1389–1398, 2008.

[63] B. L. Kidder, R. He, D. Wangsa et al., "SMYD5 controls heterochromatin and chromosome integrity during embryonic stem cell differentiation," *Cancer Research*, vol. 77, no. 23, pp. 6729–6745, 2017.

[64] D. Meistermann, S. Loubersac, A. Reignier et al., "Spatiotemporal analysis of human preimplantation development reveals dynamics of epiblast and trophectoderm," *bioRxiv*, no. article 604751, 2019.

[65] R. K. Ng, W. Dean, C. Dawson et al., "Epigenetic restriction of embryonic cell lineage fate by methylation of *Elf5*," *Nature Cell Biology*, vol. 10, no. 11, pp. 1280–1290, 2008.

[66] F. Soncin, D. Natale, and M. M. Parast, "Signaling pathways in mouse and human trophoblast differentiation: a comparative review," *Cellular and Molecular Life Sciences*, vol. 72, no. 7, pp. 1291–1302, 2015.

[67] M. Y. Turco, L. Gardner, R. G. Kay et al., "Trophoblast organoids as a model for maternal-fetal interactions during human placentation," *Nature*, vol. 564, no. 7735, pp. 263–267, 2018.

[68] H. Okae, H. Toh, T. Sato et al., "Derivation of human trophoblast stem cells," *Cell Stem Cell*, vol. 22, no. 1, pp. 50–63.e6, 2018.

[69] T. Ishiuchi, H. Ohishi, T. Sato et al., "Zfp281 shapes the transcriptome of trophoblast stem cells and is essential for placental development," *Cell Reports*, vol. 27, no. 6, pp. 1742–1754.e6, 2019.

[70] A. Jedrusik, D. E. Parfitt, G. Guo et al., "Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo," *Genes & Development*, vol. 22, no. 19, pp. 2692–2706, 2008.

[71] G. Wu, L. Gentile, T. Fuchikami et al., "Initiation of trophectoderm lineage specification in mouse embryos is

independent of Cdx2," *Development*, vol. 137, no. 24, pp. 4159–4169, 2010.

[72] R. Yagi, M. J. Kohn, I. Karavanova et al., "Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development," *Development*, vol. 134, no. 21, pp. 3827–3836, 2007.

[73] N. M. E. Fogarty, A. McCarthy, K. E. Snijders et al., "Genome editing reveals a role for OCT4 in human embryogenesis," *Nature*, vol. 550, no. 7674, pp. 67–73, 2017.

[74] K. K. Niakan and K. Eggan, "Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse," *Developmental Biology*, vol. 375, no. 1, pp. 54–64, 2013.

[75] P. A. Latos, A. R. Sienerth, A. Murray et al., "Elf5-centered transcription factor hub controls trophoblast stem cell self-renewal and differentiation through stoichiometry-sensitive shifts in target gene networks," *Genes & Development*, vol. 29, no. 23, pp. 2435–2448, 2015.

[76] C. Krendl, D. Shaposhnikov, V. Rishko et al., "GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency," *Proceedings of the National Academy of Sciences*, vol. 114, no. 45, pp. E9579–E9588, 2017.

[77] B.-K. Lee, Y. . Jang, M. Kim et al., "Super-enhancer-guided mapping of regulatory networks controlling mouse trophoblast stem cells," *Nature Communications*, vol. 10, no. 1, p. 4749, 2019.

[78] H. Gao, R. Gao, L. Zhang et al., "Esrrb plays important roles in maintaining self-renewal of trophoblast stem cells (TSCs) and reprogramming somatic cells to induced TSCs," *Journal of Molecular Cell Biology*, vol. 11, no. 6, pp. 463–473, 2019.

[79] E. B. Chuong, M. A. K. Rumi, M. J. Soares, and J. C. Baker, "Endogenous retroviruses function as species-specific enhancer elements in the placenta," *Nature Genetics*, vol. 45, no. 3, pp. 325–329, 2013.

[80] M. Aich and D. Chakraborty, "Two-Role of lncRNAs in Stem Cell Maintenance and Differentiation," in *Current Topics in Developmental Biology*, T. G. Fazzio, Ed., pp. 73–112, Academic Press, 2020.

[81] J. Chen, Y. Wang, C. Wang, J. F. Hu, and W. Li, "lncRNA functions as a new emerging epigenetic factor in determining the fate of stem cells," *Frontiers in Genetics*, vol. 11, no. 277, 2020.

[82] I. Cantone, "Reversal of X chromosome inactivation: lessons from pluripotent reprogramming of mouse and human somatic cells," *Journal of Translational Genetics and Genomics*, vol. 1, pp. 1–14, 2017.

[83] A. Minajigi, J. E. Froberg, C. Wei et al., "A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation," *Science*, vol. 349, no. 6245, article aab2276, 2015.

[84] G. G. Stirparo, T. Boroviak, G. Guo, J. Nichols, A. Smith, and P. Bertone, "Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast," *Development*, vol. 145, no. 3, 2018.

[85] S. Petropoulos, D. Edsgärd, B. Reinius et al., "Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos," *Cell*, vol. 165, no. 4, pp. 1012–1026, 2016.

[86] C. Biémont, "A brief history of the status of transposable elements: from junk DNA to major players in evolution," *Genetics*, vol. 186, no. 4, pp. 1085–1093, 2010.

[87] A. P. J. de Koning, W. Gu, T. A. Castoe, M. A. Batzer, and D. D. Pollock, "Repetitive elements may comprise over two-thirds of the human genome," *PLoS Genetics*, vol. 7, no. 12, article e1002384, 2011.

[88] P. J. Thompson, T. S. Macfarlan, and M. C. Lorincz, "Long terminal repeats: from parasitic elements to building blocks of the transcriptional regulatory repertoire," *Molecular Cell*, vol. 62, no. 5, pp. 766–776, 2016.

[89] Y. Cao, G. Chen, G. Wu et al., "Widespread roles of enhancer-like transposable elements in cell identity and long-range genomic interactions," *Genome Research*, vol. 29, no. 1, pp. 40–52, 2019.

[90] V. Sundaram, Y. Cheng, Z. Ma et al., "Widespread contribution of transposable elements to the innovation of gene regulatory networks," *Genome Research*, vol. 24, no. 12, pp. 1963–1976, 2014.

[91] A. C. Nelson, A. W. Mould, E. K. Bikoff, and E. J. Robertson, "Mapping the chromatin landscape and Blimp1 transcriptional targets that regulate trophoblast differentiation," *Scientific Reports*, vol. 7, no. 1, p. 6793, 2017.

[92] K. Adachi, I. Nikaido, H. Ohta et al., "Context-dependent wiring of Sox2 regulatory networks for self-renewal of embryonic and trophoblast stem cells," *Molecular Cell*, vol. 52, no. 3, pp. 380–392, 2013.

[93] J. M. Calabrese, W. Sun, L. Song et al., "Site-specific silencing of regulatory elements as a mechanism of X inactivation," *Cell*, vol. 151, no. 5, pp. 951–963, 2012.

[94] R. J. Mobley, D. Raghu, L. D. Duke et al., "MAP3K4 controls the chromatin modifier HDAC6 during trophoblast stem cell epithelial-to-mesenchymal transition," *Cell Reports*, vol. 18, no. 10, pp. 2387–2400, 2017.

[95] S. Chrysanthou, C. E. Senner, L. Woods et al., "A critical role of TET1/2 proteins in cell-cycle progression of trophoblast stem cells," *Stem Cell Reports*, vol. 10, no. 4, pp. 1355–1368, 2018.

[96] J. Castex, D. Willmann, T. Kanouni et al., "Inactivation of *Lsd1* triggers senescence in trophoblast stem cells by induction of *Sirt4*," *Cell Death & Disease*, vol. 8, no. 2, article e2631, 2017.

[97] C. D. Todd, Ö. Deniz, D. Taylor, and M. R. Branco, "Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells," *eLife*, vol. 8, article e44344, 2019.

[98] C. Dong, M. Beltcheva, P. Gontarz et al., "Derivation of trophoblast stem cells from naïve human pluripotent stem cells," *eLife*, vol. 9, 2020.

[99] W. A. Pastor, W. Liu, D. Chen et al., "TFAP2C regulates transcription in human naive pluripotency by opening enhancers," *Nature Cell Biology*, vol. 20, no. 5, pp. 553–564, 2018.

[100] S. Schoenfelder, B. Mifsud, C. E. Senner et al., "Divergent wiring of repressive and active chromatin interactions between mouse embryonic and trophoblast lineages," *Nature Communications*, vol. 9, no. 1, pp. 4189–4189, 2018.

[101] L. J. Kaaij, F. Mohn, R. H. van der Weide, E. de Wit, and M. Bühler, "The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse," *Cell*, vol. 178, no. 6, pp. 1437–1451.e14, 2019.

[102] A. Sahakyan, R. Kim, C. Chronis et al., "Human naive pluripotent stem cells model X chromosome dampening and X inactivation," *Cell Stem Cell*, vol. 20, no. 1, pp. 87–101, 2017.

[103] F. J. Najm, J. G. Chenoweth, P. D. Anderson et al., "Isolation of epiblast stem cells from preimplantation mouse embryos," *Cell Stem Cell*, vol. 8, no. 3, pp. 318–325, 2011.

[104] T. W. Theunissen, M. Friedli, Y. He et al., "Molecular criteria for defining the naive human pluripotent state," *Cell Stem Cell*, vol. 19, no. 4, pp. 502–515, 2016.

[105] S. Blij, A. Parenti, N. Tabatabai-Yazdi, and A. Ralston, "Cdx2 efficiently induces trophoblast stem-like cells in naïve, but not primed, pluripotent stem cells," *Stem Cells and Development*, vol. 24, no. 11, pp. 1352–1365, 2015.

[106] A. Neagu, E. van Genderen, I. Escudero et al., "In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states," *Nature Cell Biology*, vol. 22, no. 5, pp. 534–545, 2020.

[107] M. Nakanishi, R. R. Mitchell, Y. D. Benoit et al., "Human pluripotency is initiated and preserved by a unique subset of founder cells," *Cell*, vol. 177, no. 4, pp. 910–924.e22, 2019.

[108] R. Li, C. Zhong, Y. Yu et al., "Generation of blastocyst-like structures from mouse embryonic and adult cell cultures," *Cell*, vol. 179, no. 3, pp. 687–702.e18, 2019.

Hindawi
Stem Cells International
Volume 2021, Article ID 6663960, 8 pages
https://doi.org/10.1155/2021/6663960

Research Article

Nsd2 Represses Endogenous Retrovirus MERVL in Embryonic Stem Cells

Tingting Gao, Fuquan Chen, Wenying Zhang, Xuan Zhao, Xiao Hu, and Xinyi Lu

State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy, Nankai University, Tianjin 300350, China

Correspondence should be addressed to Xinyi Lu; luxy@nankai.edu.cn

Received 28 October 2020; Revised 31 December 2020; Accepted 5 January 2021; Published 18 January 2021

Academic Editor: Wensheng Zhang

The facilitates chromatin transcription (FACT) complex is a histone H2A/H2B chaperone, which represses endogenous retroviruses (ERVs) and transcription of ERV-chimeric transcripts. It binds to both transcription start site and gene body region. Here, we investigated the downstream targets of FACT complex to identify the potential regulators of MERVL, which is a key 2-cell marker gene. H3K36me2 profile was positively correlated with that of FACT component Ssrp1. Among H3K36me2 deposition enzymes, *Nsd2* was downregulated after the loss of *Ssrp1*. Furthermore, we demonstrated that Nsd2 repressed the expression of ERVs without affecting the expression of pluripotency genes. The expression of MERVL and 2-cell genes was partially rescued by *Nsd2* overexpression. The enrichment of H3K36me2 decreased on MERVL-chimeric gene in ESCs without *Ssrp1*. Our study discovers that Nsd2 is a repressor of MERVL, and FACT partially represses MERVL expression by regulating the expression of *Nsd2* and its downstream H3K36me2.

1. Introduction

Endogenous retroviruses are important components of the mammalian genome [1]. They are usually silenced by host cells to maintain genome stability. However, studies also show that ERVs are functional during development and in mouse embryonic stem cells (ESCs) [2–7]. For example, MERVL marks the 2-cell (2C) embryos and a minority of 2C-like cells within the ESC population [8, 9]. MERVL can be silenced by various epigenetic regulators, such as histone H3 variants, H3K9 methyltransferases, and histone chaperones [10–15]. Recently, we found that H2A/H2B histone chaperone FACT (facilitates chromatin transcription) complex participated in the repression of MERVL and MERVL-derived cryptic transcripts in ESCs [16]. FACT functions partially through Usp7 to remove H2Bub on MERVL and MERVL-fused genes [16]. However, the impact of *Usp7* depletion on MERVL induction is weaker than the loss of FACT complex itself [16]. This implies that there are other ways present for FACT complex to repress the expression of MERVL and its chimeric transcripts. Therefore, in this study, we aim to identify indirect pathways downstream of FACT complex in repressing the expression of MERVL.

We and others previously found that Ssrp1 binding was enriched around transcription start sites and on gene body region [16, 17]. Gene body region can be marked by H3K36me3 and H3K36me2 [18, 19]. An important H3K36 methyltransferase family is Nsd family. Here, we examine the role of Nsd family members at the downstream of FACT complex in repressing the ERV expression.

2. Methods

2.1. Cell Culture. E14 mouse embryonic stem cells (ESCs) were cultured on plates coated with sterile 0.1-0.2% gelatin (G1890, Sigma) in medium containing Dulbecco's modified Eagle's medium (Hyclone), 15% fetal bovine serum (Hyclone), 10 ng/ml leukaemia inhibitory factor (LIF) (Z03077, GenScript), 1% penicillin/streptomycin (P1400, Solarbio), 2 mM L-glutamine (Gibco), 0.1 mM nonessential amino acids (Gibco), and 0.1 mM β-mercaptoethanol (Sigma). ESCs were passaged every two days for maintenance.

Table 1: Primers for qPCR analysis.

Gene	Forward	Reverse
Nsd1	TCCGGTGAATTTAGATGCCTCC	CGGTAACTGCATAGTACACCCAT
Nsd2	GGTGATCCTGGCACAGACAA	GAGCAGAGCCTGTGGACTTT
Nsd3	CCGAGGTTGTGCCAAAGAAG	ACGGAGCTGTCACTGAATCTG
MERVL	AAGAGCCAAGACCTGCTGAG	TCCTCGTTTCTGCAACTGGT
LINE1	GGACCAGAAAAGAAATTCCTCCCG	CTCTTCTGGCTTTCATAGTCTCTGG
SINEB1	GTGGCGCACGCCTTTAATC	GACAGGGTTTCTCTGTGTAG
Oct4	GTGGAAAGCAACTCAGAGG	GGTTCCACCTTCTCCAACT
Sox2	GCGGAGTGGAAACTTTTGTCC	CGGGAAGCGTGTACTTATCCTT
Nanog	TTGCTTACAAGGGTCTGCTACT	ACTGGTAGAAGAATCAGGGCT
Zscan4	GAGATTCATGGAGAGTCTGACTGATGAGTG	GCTGTTGTTTCAAAAGCTTGATGACTTC
Dux	CCCAGCGACTCAAACTCCTTC	GGACTTCGTCCAGCAGTTGAT
ChIP Control	GATTAGCAGCTCCACAGGA	TGGACAATGTGGCCTGTTTA
Zfp809 ChIP	AAGCTGGCTGACTGTAGTGG	GTGAGCCTTCCAATTCCGGA
Foxa2	CCCTACGCCAACATGAACTCG	GTTCTGCCGGTAGAAAGGGA
Sox17	GATGCGGGATACGCCAGTG	CCACCACCTCGCCTTTCAC
Gata4	CCCTACCCAGCCTACATGG	ACATATCGAGATTGGGGTGTCT
Nkx2.5	GACAAAGCCGAGACGGATGG	CTGTCGCTTGCACTTGTAGC
Msx1	TGCTGCTATGACTTCTTTGCC	GCTTCCTGTGATCGGCCAT
Pax6	TACCAGTGTCTACCAGCCAAT	TGCACGAGTATGAGGAGGTCT
Foxd3	GAGTTCATCAGCAACCGTTTTC	CGAAGCTCTGCATCATCAGC
Gata3	CTCGGCCATTCGTACATGGAA	GGATACCTCTGCACCGTAGC

2.2. Analysis of ChIP-seq Data. For ChIP-seq data analysis, all reads were first processed with Cutadapt to trim adaptor sequences and low-quality reads and subsequently mapped to the mouse mm10 genome assembly using Bowtie2. The correlation coefficient between Ssrp1, H3K4me3, H3K36me2, and H3K36me3 was determined by plotCorrelation from Deeptools. The ChIP-seq signal enrichment file was obtained by bamCompare from Deeptools, and the ChIP signal line plot was also generated by Deeptools. Gene structure information was inferred from Gencode.vM21 annotation file.

2.3. Reverse Transcription and qPCR. Total RNA was isolated from cells by RNAiso Reagent (B9109, Takara) in DEPC water (B501005, Sangon Biotech) following by DNase I treatment in RNase-free tubes (401001, NEST Biotechnology). Reverse transcription was performed with 1 μg purified RNA using Transcriptor First Strand cDNA Synthesis Kit (4897030001, Roche) as described previously [20]. qPCR analysis was carried out using SYBR Green qPCR Master Mix (H97410, Yeasen) and a qPCR detection system (CFX384 Real-Time System, Bio-Rad) according to standard protocols. Primers are synthesized by Sangon Biotech and included in Table 1.

2.4. shRNA-Mediated Gene Depletion. The shRNAs targeting *Nsd2* were designed by an online tool (http://sirna.wi.mit.edu/) [21]. The targeting sequences of shRNAs are CCTG GTGCTCATGATACTAAA for shRNA1 and GAGCTG ACTTTCAACTATAA for shRNA2. The shRNAs were synthesized by GENEWIZ corporation and cloned into pSuper-puro. 1 μg plasmid was transfected into mouse ESCs with Polyjet (SignaGen). The cells were further cultured for three days under puromycin selection (1 μg/ml) and harvested for RNA extraction.

2.5. Chromatin Immunoprecipitation (ChIP) Coupled qPCR. ChIP-qPCR was performed as described before [16]. Briefly, ESCs were harvested and crosslinked with 1% formaldehyde, and cell fixation was ceased with the addition of glycine. The cells were primarily lysed, and chromatin extracts were collected and sonicated for obtaining soluble chromatin fragments. The chromatin samples were incubated with specific antibody and immunoprecipitated on protein G magnetic beads (GenScript, the USA). The immunoprecipitated DNA was next eluted, decrosslinked, and analyzed by qPCR. For immunoprecipitation, the antibody used was anti-H3K36me2 (ab9049, Abcam).

2.6. Establishment of ESC Cell Lines Overexpressing Nsd2. Mouse *Nsd2* coding region was cloned into pCAG-3HA vector with hygromycin resistance and purified with a kit (1211-01, Biomiga). *Ssrp1*$^{-/-}$ ESCs were transfected with 1 μg plasmid expressing *Nsd2* via Polyjet reagent (SL100688, SignaGen) following the manufacturer's recommended protocol. The ESCs were continuously selected with 800 μg/ml hygromycin B for 14 days. After selective cell culture, ESCs were collected for downstream experiments.

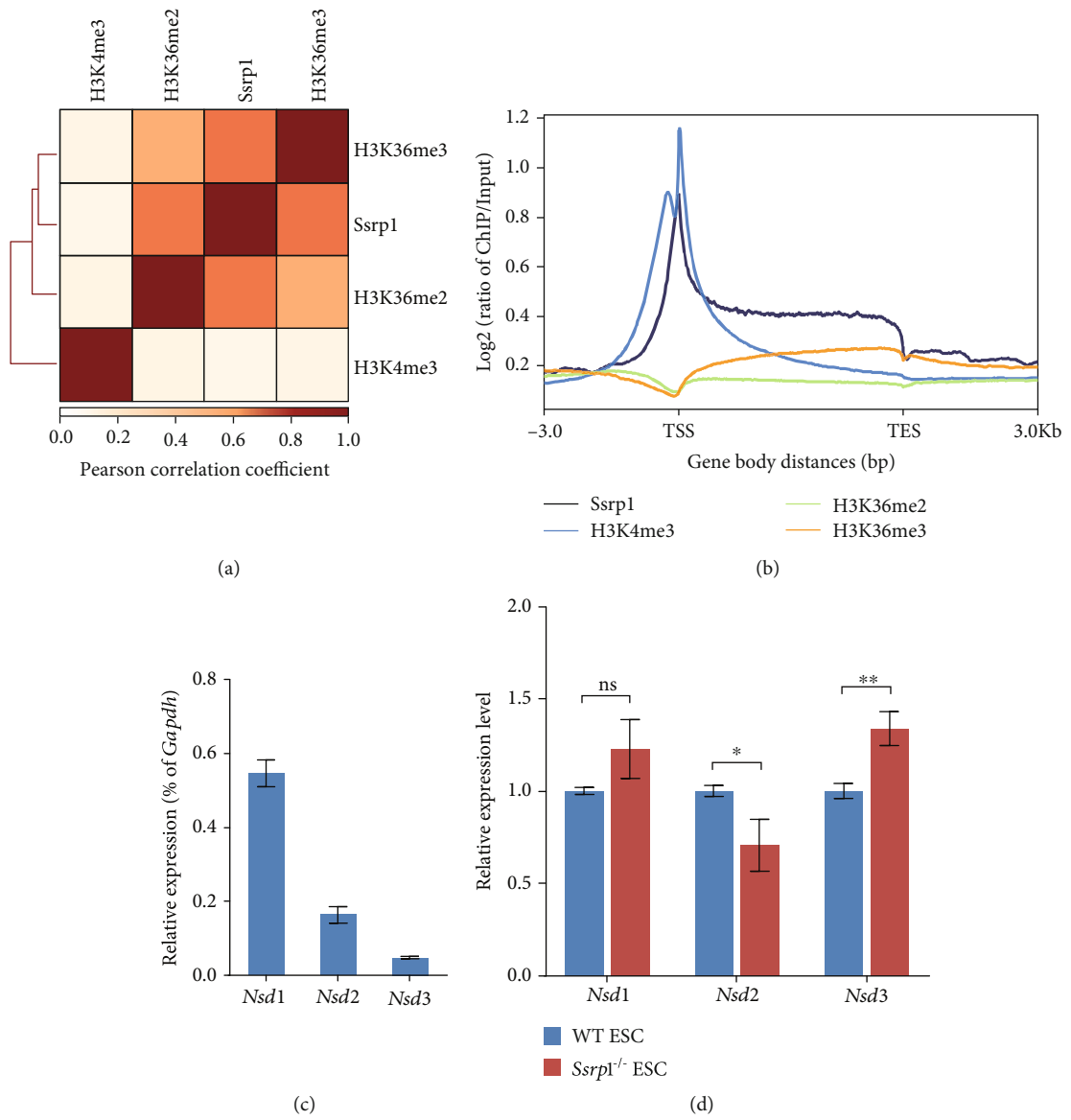


Figure 1: Nsd2-mediated H3K36me2 is correlated with Ssrp1 binding. (a) Genomic distribution profile of H3K36me2, H3K36me3, H3K4me3, and Ssrp1. Color scale represents the strength of Pearson's correlation. (b) ChIP-seq signal density enrichment of Ssrp1, H3K4me3, H3K36me2, and H3K36me3 (Y-axis) on gene body from TSS to TES (X-axis). H3K36me3 enrichment increased from TSS to TES. (c) Relative mRNA expression of *Nsd1*, *Nsd2*, and *Nsd3* in ESCs according to qPCR. (d) qPCR analysis of *Nsd1*, *Nsd2*, and *Nsd3* expression levels in WT ESC and *Ssrp1*$^{-/-}$ ESC; data are shown as mean ± s.e.m.; $n = 3$ biologically independent repeats. ns: nonsignificant; $^{*}p < 0.05$; $^{**}p < 0.01$.

3. Results

3.1. FACT Complex Binding Is Correlated with H3K36 Methylation. Previously, we found that FACT complex interacted with both promoter and gene body regions, which are marked by H3K36me2/3. Interestingly, the genomic distribution profile of H3K36me2 and H3K36me3 was positively correlated with that of Ssrp1 (Figure 1(a)), in contrast with the lower correlation strength of Ssrp1 with H3K4me3 (Figure 1(a)). Moreover, it was noteworthy that H3K36me3 enrichment on the gene body continuously increased from transcription start site (TSS) to transcription end site (TES) whereas the H3K36me2 was preferentially associated with TSS region and gradually decayed from TSS to TES (Figure 1(b)). The distribution profile of H3K36me2 was more similar to that of FACT complex than H3K36me3 (Figures 1(a) and 1(b)). Therefore, we further examined the expression of Nsd family genes (*Nsd1*, *Nsd2*, and *Nsd3*), which are known to mediate H3K36 methylation. *Nsd1* was expressed highest in ESCs while the expression of *Nsd2* and *Nsd3* was lower (Figure 1(c)). The expression of *Nsd1* and *Nsd3* remained unchanged or slightly upregulated in ESCs without FACT complex (Figure 1(d)). However, the *Nsd2* expression was downregulated in *Ssrp1*$^{-/-}$ ESCs (Figure 1(d)), implying *Nsd2* as a potential downstream target gene of FACT complex.

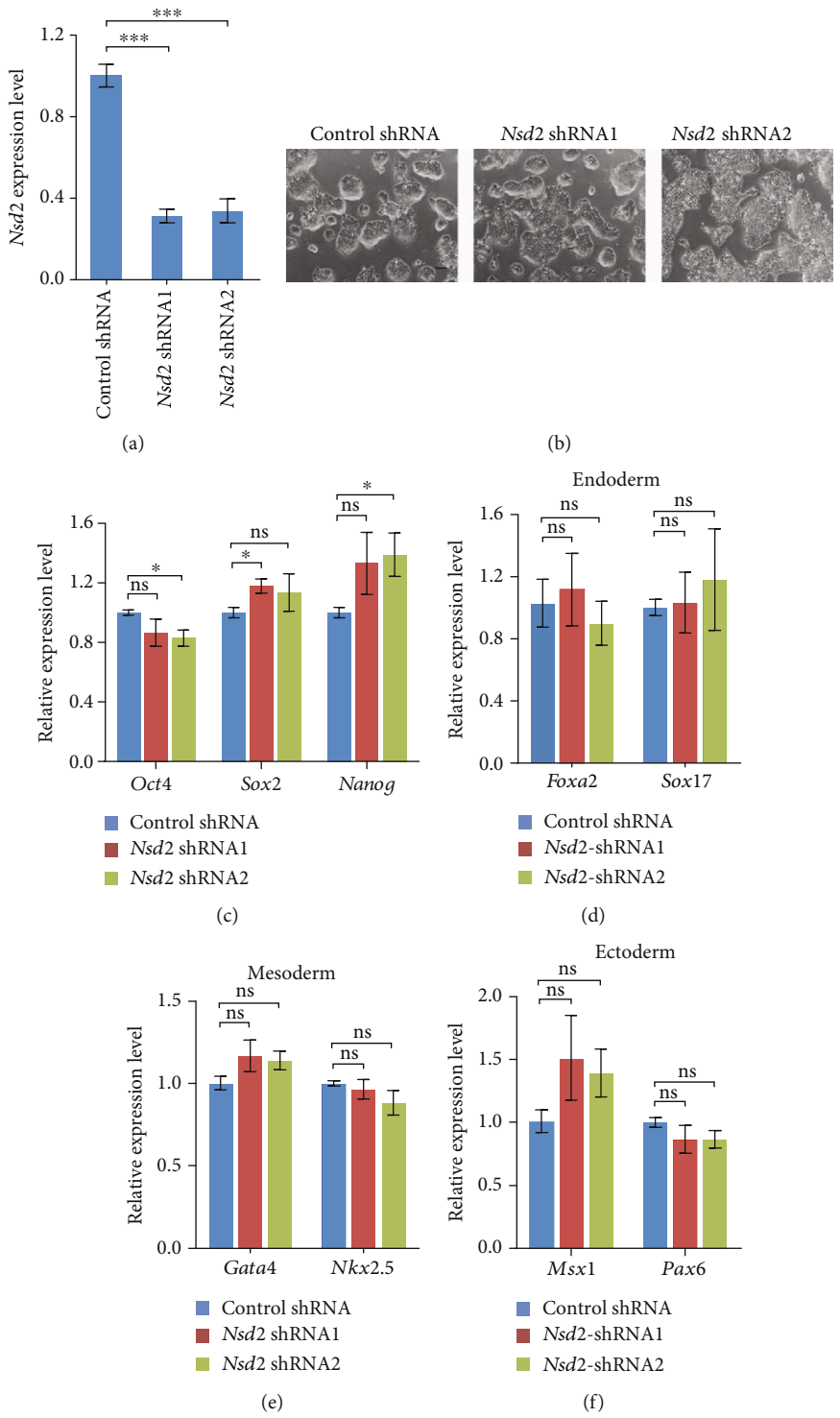


Figure 2: Continued.

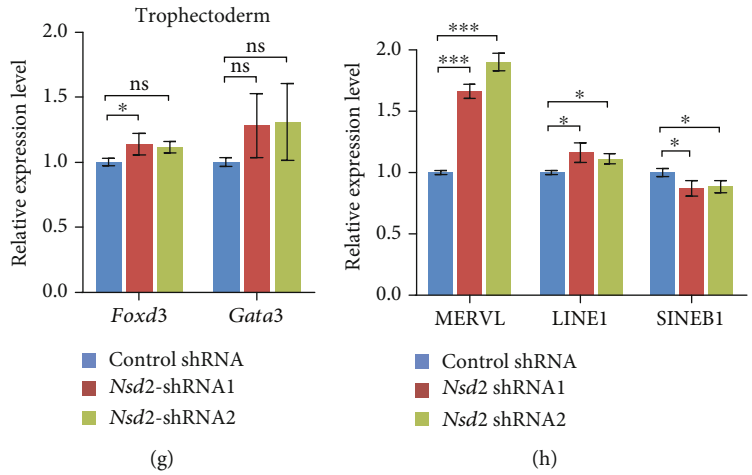


Figure 2: Depletion of *Nsd2* in ESCs. (a) qPCR analysis of the relative expression levels of *Nsd2* after *Nsd2* depletion in ESCs with normalization to *Gapdh*. Data were plotted as mean ± s.e.m., $n = 3$ biological repeats. (b) Cell morphology of *Nsd2* depletion in ESCs. Scale bar, 100 μm. (c) qPCR analysis of pluripotent genes (*Oct4*, *Sox2*, and *Nanog*) in ESCs with *Nsd2* depleted. (d–g) qPCR analysis of endoderm markers *Foxa2* and *Sox17* (d), mesoderm markers *Gata4* and *Nkx2.5* (e), ectoderm markers *Msx1* and *Pax6* (f), and trophectoderm markers *Foxd3* and *Gata3* (g). (h) qPCR analysis of the expression levels of MERVL and other retrotransposons (LINE1, SINE B1) after *Nsd2* depletion in ESCs. The results were normalized to *Gapdh*. Data were shown as mean ± s.e.m.; $n = 3$ biologically independent replicates; ns: nonsignificant; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

3.2. Nsd2 Represses MERVL in ESCs. In agreement with the close resemblance of Ssrp1 binding profile and H3K36me2, the main chromatin-regulatory activity of Nsd2 is mediating the dimethylation of histone H3 at lysine 36 (H3K36me2) [19]. Hence, we depleted *Nsd2* in ESCs with two independent shRNAs to examine whether Nsd2 can regulate the expression of ERVs (Figure 2(a)). The depletion of *Nsd2* did not affect the cell morphology of ESCs (Figure 2(b)). Also, the expression of pluripotency genes (*Oct4*, *Sox2*, *Nanog*) was not disturbed by two *Nsd2* shRNAs at the same time (Figure 2(c)). The suppression of *Nsd2* by two independent shRNAs did not disrupt the expression of differentiation markers for endoderm (*Foxa2* and *Sox17*), mesoderm (*Gata4* and *Nkx2.5*), ectoderm (*Msx1* and *Pax6*), and trophectoderm (*Foxd3* and *Gata3*) at the same time (Figures 2(d)–2(g)), suggesting that ESCs remain undifferentiated without *Nsd2*. Intriguingly, the expression of MERVL was activated to ~2 folds by *Nsd2* depletion (Figure 2(h)), but other retrotransposons (LINE1 or SINE B1) were less activated or downregulated, confirming that Nsd2 acts downstream of FACT complex to repress the ERV expression. These results suggest that Nsd2 represses the expression of MERVL without affecting ESC pluripotency.

3.3. Nsd2 Overexpression Rescues MERVL Expression in Ssrp1$^{-/-}$ ESCs. Since Nsd2 represses the MERVL expression and *Nsd2* is downregulated in *Ssrp1*$^{-/-}$ ESCs, we next asked whether restoration of the *Nsd2* expression in *Ssrp1*$^{-/-}$ ESCs could rescue the expression of MERVL. Hence, we established an *Ssrp1*$^{-/-}$ ESC line with *Nsd2* overexpressed. Our qPCR results showed that *Nsd2* was successfully overexpressed in *Ssrp1*$^{-/-}$ ESCs (Figure 3(a)). Moreover, overexpression of *Nsd2* could partially reduce the expression of MERVL in *Ssrp1*$^{-/-}$ ESCs (Figure 3(b)). Furthermore, the expression of 2-cell marker genes (*Zscan4* and *Dux*) was partially restored (Figure 3(c)). These results suggest that *Nsd2* is an important downstream target gene of Ssrp1 in repressing ERVs and 2-cell genes.

3.4. Nsd2-Mediated H3K36me2 Is Reduced on MERVL-Fused Genes in Ssrp1$^{-/-}$ *ESCs.* We further investigated whether the target of Nsd2, H3K36me2, was affected at MERVL-fusion genes in ESCs without FACT by ChIP-qPCR. Our ChIP-qPCR results revealed that H3K36me2 was enriched on MERVL-fused gene such as *Zfp809* (Figure 4(a)) but not on the control region (Figure 4(b)). However, this enrichment was decreased on MERVL-fusion genes in ESC without the *Ssrp1* expression (Figures 4(a) and 4(b)). Together, these suggest that the decreased enrichment of H3K36me2 on MERVL-fused genes may explain the activation of MERVL-fused genes after *Nsd2* downregulation in *Ssrp1*$^{-/-}$ ESCs.

4. Discussion

In summary, we discovered that Nsd2 was a repressor of MERVL and MERVL-fused 2C genes, and the downregulation of *Nsd2* worked as a secondary regulatory route to activate MERVL after the loss of *Ssrp1*. It is interesting to see that only *Nsd2* (Figure 1(d)), but neither *Nsd1* nor *Nsd3*, is downregulated by the disruption of FACT function, given that all three Nsd genes participate in H3K36 methylation. Nsd2 is an important H3K36me2 methyltransferase [19, 22]. Loss of *Nsd2* mimics H3.3K36M mutation, but not *Nsd1* or *Setd2* mimics the effects of H3.3K36M on adipogenesis [23], implicating a unique role of Nsd2 among Nsd members in gene expression regulation. H3K6me2 was associated with both activation and repression of the gene expression [24]. It was recently reported that Nsd1/Nsd2-mediated intergenic H3K36me2 recruited Dnmt3a for DNA methylation [25, 26]. In yeast cells, H3K36me1/2/3 was also shown

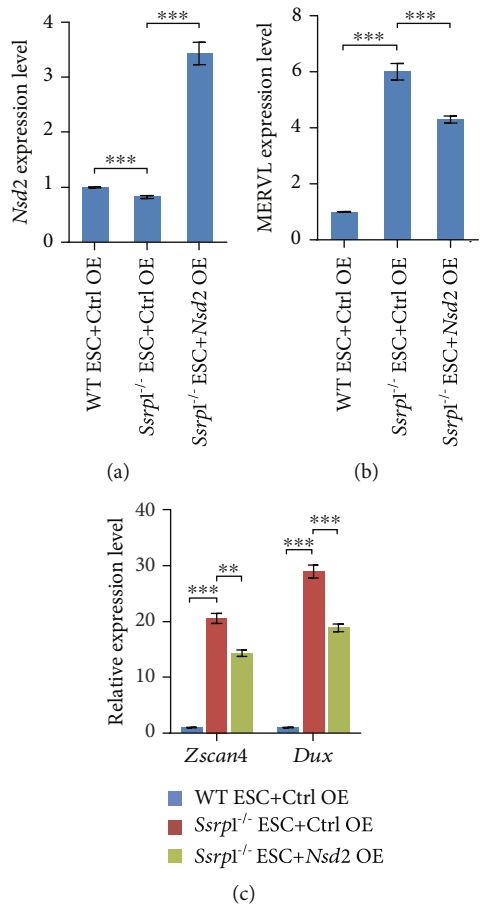


Figure 3: *Nsd2* overexpression rescues the expression of MERVL and 2-cell genes. (a) qPCR analysis of *Nsd2* expression after overexpression of *Nsd2* in *Ssrp1*$^{-/-}$ ESC. The results were shown as mean ± s.e.m., $n = 3$ biologically independent replicates and normalized to *Gapdh*. (b) qPCR analysis of the expression of MERVL after overexpression of *Nsd2* in *Ssrp1*$^{-/-}$ ESC. Mean ± s.e.m., $n = 3$. (c) qPCR analysis of the expression of 2-cell marker genes *Zscan4* and *Dux* after overexpression of *Nsd2* in *Ssrp1*$^{-/-}$ ESC. Ctrl: control; OE: overexpression; mean ± s.e.m., $n = 3$; $^{**}p < 0.01$; $^{***}p < 0.001$.

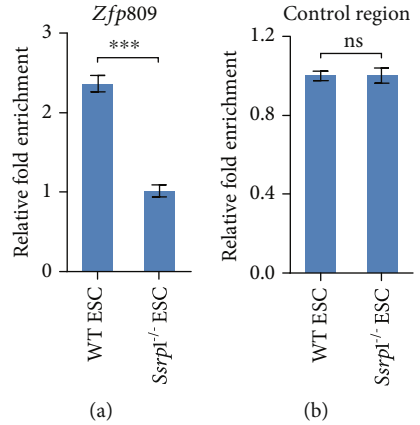


Figure 4: H3K36me2 enrichment on MERVL-fused genes. (a, b) ChIP-qPCR analysis of H3K36me2 enrichment on MERVL-fused gene *Zfp809* (a) and control region (b) in WT ESC and *Ssrp1*$^{-/-}$ ESC with normalization to *Gapdh* and input; data shown as mean ± s.e.m. ($n = 3$ extracts). ns: nonsignificant; $^{***}p < 0.001$.

to repress cryptic transcription [27]. Moreover, H3K36me2 can recruit the Rpd3s histone deacetylase to repress spurious transcription [28]. These are consistent with our finding that H3K36me2 decreased on MERVL-fused genes after the loss of *Ssrp1* (Figures 4(a) and 4(b)), implying a potential repression role of H3K36me2.

Nsd2 is not only involved in gene transcription regulation. It participates in regulating genome stability and methylates non-histone proteins as well. Nsd2-mediated H3K36me2 promotes nonhomologous end-joining at unprotected telomeres and thereby enhances genomic instability caused by telomere dysfunction [29]. Human NSD2-mediated PTEN methylation regulates cell responses to DNA damage [30]. It is recently discovered that DNA damage is induced by the depletion of MERVL activator Zscan4 [3, 31]. Responses of ATR and CHK1 to replication stresses activate *Zscan4* and MERVL [20], implying that DNA damage-induced replication stress and Zscan4 reciprocally regulate each other. It will be interesting to investigate whether Nsd2 is involved DNA damage repair and its relationship with Zscan4 in the future.

5. Conclusion

In conclusion, we found that *Nsd2*, as a downstream gene of FACT, repressed MERVL, without influencing ESC pluripotency. The decreased *Nsd2* in *Ssrp1*$^{-/-}$ ESCs was accompanied by reduced H3K36me2 on MERVL-fused genes while overexpression of *Nsd2* partially rescued the expression of MERVL. These findings establish Nsd2 as an important repressor of MERVL in ESCs and during the loss of FACT function.

Data Availability

Published ChIP-seq data we analyzed are GSE141791 for Ssrp1 [16], GSE117155 for H3K36me2 [32], GSE110321 for H3K36me3 [33], and GSE90893 for H3K4me3 [34].

Conflicts of Interest

We declare that there is no conflict of interest present for this study.

Authors' Contributions

X.L. conceived and designed the study. T.G., F.C., and W.Z. performed most experiments. X.Z. did bioinformatics analysis. T.G., X.H., and X.L. wrote the manuscript.

Acknowledgments

This work is supported by grants from the National Key Research and Development Program of China (2018YFA0107000), the National Natural Science Foundation of China (32070858, 31671352, 31871488), and the Fundamental Research Funds for the Central Universities (63151113, 63161150).

References

[1] R. Enriquez-Gasca, P. A. Gould, and H. M. Rowe, "Host gene regulation by transposable elements: the new, the old and the ugly," *Viruses*, vol. 12, no. 10, p. 1089, 2020.

[2] X. Lu, F. Sachs, L. A. Ramsay et al., "The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity," *Nature Structural & Molecular Biology*, vol. 21, no. 4, pp. 423–425, 2014.

[3] W. Zhang, F. Chen, R. Chen et al., "Zscan4c activates endogenous retrovirus MERVL and cleavage embryo genes," *Nucleic Acids Research*, vol. 47, no. 16, pp. 8485–8501, 2019.

[4] J. Göke, X. Lu, Y. S. Chan et al., "Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells," *Cell Stem Cell*, vol. 16, no. 2, pp. 135–141, 2015.

[5] E. J. Grow, R. A. Flynn, S. L. Chavez et al., "Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells," *Nature*, vol. 522, no. 7555, pp. 221–225, 2015.

[6] S. Mi, X. Lee, X. P. Li et al., "Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis," *Nature*, vol. 403, no. 6771, pp. 785–789, 2000.

[7] A. E. Peaston, A. V. Evsikov, J. H. Graber et al., "Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos," *Developmental Cell*, vol. 7, no. 4, pp. 597–606, 2004.

[8] T. S. Macfarlan, W. D. Gifford, S. Driscoll et al., "Embryonic stem cell potency fluctuates with endogenous retrovirus activity," *Nature*, vol. 487, no. 7405, pp. 57–63, 2012.

[9] D. Kigami, N. Minami, H. Takayama, and H. Imai, "MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos," *Biology of Reproduction*, vol. 68, no. 2, pp. 651–654, 2003.

[10] S. J. Elsässer, K. M. Noh, N. Diaz, C. D. Allis, and L. A. Banaszynski, "Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells," *Nature*, vol. 522, no. 7555, pp. 240–244, 2015.

[11] Y. Hatanaka, K. Inoue, M. Oikawa et al., "Histone chaperone CAF-1 mediates repressive histone modifications to protect preimplantation mouse embryos from endogenous retrotransposons," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 47, pp. 14641–14646, 2015.

[12] T. Ishiuchi, R. Enriquez-Gasca, E. Mizutani et al., "Early embryonic-like cells are induced by downregulating replication- dependent chromatin assembly," *Nature Structural & Molecular Biology*, vol. 22, no. 9, pp. 662–671, 2015.

[13] B. X. Yang, C. A. el Farran, H. C. Guo et al., "Systematic identification of factors for provirus silencing in embryonic stem cells," *Cell*, vol. 163, no. 1, pp. 230–245, 2015.

[14] H. M. Rowe, J. Jakobsson, D. Mesnard et al., "KAP1 controls endogenous retroviruses in embryonic stem cells," *Nature*, vol. 463, no. 7278, pp. 237–240, 2010.

[15] I. A. Maksakova, P. J. Thompson, P. Goyal et al., "Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells," *Epigenetics & Chromatin*, vol. 6, no. 1, p. 15, 2013.

[16] F. Chen, W. Zhang, D. Xie, T. Gao, Z. Dong, and X. Lu, "Histone chaperone FACT represses retrotransposon MERVL and MERVL-derived cryptic promoters," *Nucleic Acids Research*, vol. 48, no. 18, pp. 10211–10225, 2020.

[17] C. Mylonas and P. Tessarz, "Transcriptional repression by FACT is linked to regulation of chromatin accessibility at the promoter of ES cells," *Life Science Alliance*, vol. 1, no. 3, article e201800085, 2018.

[18] T. S. Mikkelsen, M. Ku, D. B. Jaffe et al., "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells," *Nature*, vol. 448, no. 7153, pp. 553–560, 2007.

[19] A. J. Kuo, P. Cheung, K. Chen et al., "NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming," *Molecular Cell*, vol. 44, no. 4, pp. 609–620, 2011.

[20] S. Atashpaz, S. Samadi Shams, J. M. Gonzalez et al., "ATR expands embryonic stem cell fate potential in response to replication stress," *eLife*, vol. 9, 2020.

[21] B. Yuan, R. Latek, M. Hossbach, T. Tuschl, and F. Lewitter, "siRNA Selection Server: an automated siRNA oligonucleotide prediction server," *Nucleic Acids Research*, vol. 32, no. Web Server, pp. W130–W134, 2004.

[22] Y. Li, P. Trojer, C. F. Xu et al., "The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate," *The Journal of Biological Chemistry*, vol. 284, no. 49, pp. 34283–34295, 2009.

[23] L. Zhuang, Y. Jang, Y. K. Park et al., "Depletion of Nsd2-mediated histone H3K36 methylation impairs adipose tissue development and function," *Nature Communications*, vol. 9, no. 1, p. 1796, 2018.

[24] E. J. Wagner and P. B. Carpenter, "Understanding the language of Lys36 methylation at histone H3. Nature reviews," *Nature Reviews Molecular Cell Biology*, vol. 13, no. 2, pp. 115–126, 2012.

[25] D. N. Weinberg, S. Papillon-Cavanagh, H. Chen et al., "The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape," *Nature*, vol. 573, no. 7773, pp. 281–286, 2019.

[26] W. Xu, J. Li, B. Rong et al., "DNMT3A reads and connects histone H3K36me2 to DNA methylation," *Protein & Cell*, vol. 11, no. 2, pp. 150–154, 2020.

[27] J. V. DiFiore, T. S. Ptacek, Y. Wang, B. Li, J. M. Simon, and B. D. Strahl, "Unique and shared roles for histone H3K36 methylation states in transcription regulation functions," *Cell Reports*, vol. 31, no. 10, p. 107751, 2020.

[28] B. Li, J. Jackson, M. D. Simon et al., "Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription," *The Journal of Biological Chemistry*, vol. 284, no. 12, pp. 7970–7976, 2009.

[29] I. de Krijger, J. van der Torre, M. H. Peuscher, M. Eder, and J. J. L. Jacobs, "H3K36 dimethylation by MMSET promotes classical non-homologous end-joining at unprotected telomeres," *Oncogene*, vol. 39, no. 25, pp. 4814–4827, 2020.

[30] J. Zhang, Y. R. Lee, F. Dang et al., "PTEN methylation by NSD2 controls cellular sensitivity to DNA damage," *Cancer Discovery*, vol. 9, no. 9, pp. 1306–1323, 2019.

[31] R. Srinivasan, N. Nady, N. Arora et al., "Zscan4 binds nucleosomal microsatellite DNA and protects mouse two-cell embryos from DNA damage," *Science Advances*, vol. 6, no. 12, p. eaaz9115, 2020.

[32] G. LeRoy, O. Oksuz, N. Descostes et al., "LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells," *Science Advances*, vol. 5, no. 10, p. eaay3068, 2019.

[33] H. Huang, H. Weng, K. Zhou et al., "Histone H3 trimethylation at lysine 36 guides m^6A RNA modification co-transcriptionally," *Nature*, vol. 567, no. 7748, pp. 414–419, 2019.

[34] C. Chronis, P. Fiziev, B. Papp et al., "Cooperative binding of transcription factors orchestrates reprogramming," *Cell*, vol. 168, no. 3, pp. 442–459.e20, 2017, e20.

Hindawi
Stem Cells International
Volume 2021, Article ID 6668866, 7 pages
https://doi.org/10.1155/2021/6668866

Review Article

BAF Complex in Embryonic Stem Cells and Early Embryonic Development

Heyao Zhang,[1] Xuepeng Wang,[2] Jingsheng Li,[1] Ronghua Shi,[3] and Ying Ye[1]

[1]*Cam-Su Genomic Resource Center, Medical College of Soochow University, Suzhou, Jiangsu 215000, China*
[2]*College of Traditional Chinese Medicine, Macau University of Science and Technology, China*
[3]*School of Life Sciences, University of Science and Technology of China, China*

Correspondence should be addressed to Ronghua Shi; rhsh@ustc.edu.cn and Ying Ye; yye@suda.edu.cn

Received 30 October 2020; Revised 30 December 2020; Accepted 5 January 2021; Published 16 January 2021

Academic Editor: Qiang Wu

Embryonic stem cells (ESCs) can self-renew indefinitely and maintain their pluripotency status. The pluripotency gene regulatory network is critical in controlling these properties and particularly chromatin remodeling complexes. In this review, we summarize the research progresses of the functional and mechanistic studies of BAF complex in mouse ESCs and early embryonic development. A discussion of the mechanistic bases underlying the distinct phenotypes upon the deletion of different BAF subunits in ESCs and embryos will be highlighted.

1. Introduction

Embryonic stem cells (ESCs) are derived from the inner cell mass of blastocysts in early embryos [1–3]. With the remarkable abilities to indefinitely self-renew and differentiate to all types of cells in the body, ESCs become an ideal model to study cell fate determination and lineage differentiation, therefore having broad applications in the fields of regenerative medicine and translational medicine.

Since their isolation, the mechanism underlying the self-renewal and pluripotency of ESCs has been the focus of intensive research in the field of stem cell biology [4]. Numerous studies demonstrate that the identity of ESCs is controlled by a core transcriptional regulatory network composed of signaling pathways such as the LIF/STAT3 pathway [4–6], pluripotent transcription factors such as OCT4, SOX2, NANOG, and KLF4 [7–9], protein complexes [10–12], microRNAs [13], and chromatin remodeling complexes [12].

2. Chromatin Remodeling Complexes

Specific transcriptomes expressed in different types of mammalian cells are controlled partly by their unique chromatin states. The regulation of chromatin states selectively causes gene expression or silencing via controlling the access of transcriptional factors to gene regulatory elements. This variation of transcriptional activity according to the chromatin structural changes is called chromatin remodeling [14]. There are two main types of chromatin remodeling: one is covalent histone modification, which includes acetylation, methylation, phosphorylation, and ubiquitination; the other is ATP-dependent physical modification, which is achieved mainly through ATP-dependent protein complexes [14].

The ATP-dependent protein complexes with ATPase activity, termed chromatin remodeling complex, use the energy generated by hydrolysis of ATP to make the four changes in the nucleosomes structure and thereby regulate gene expression (Figure 1) [15].

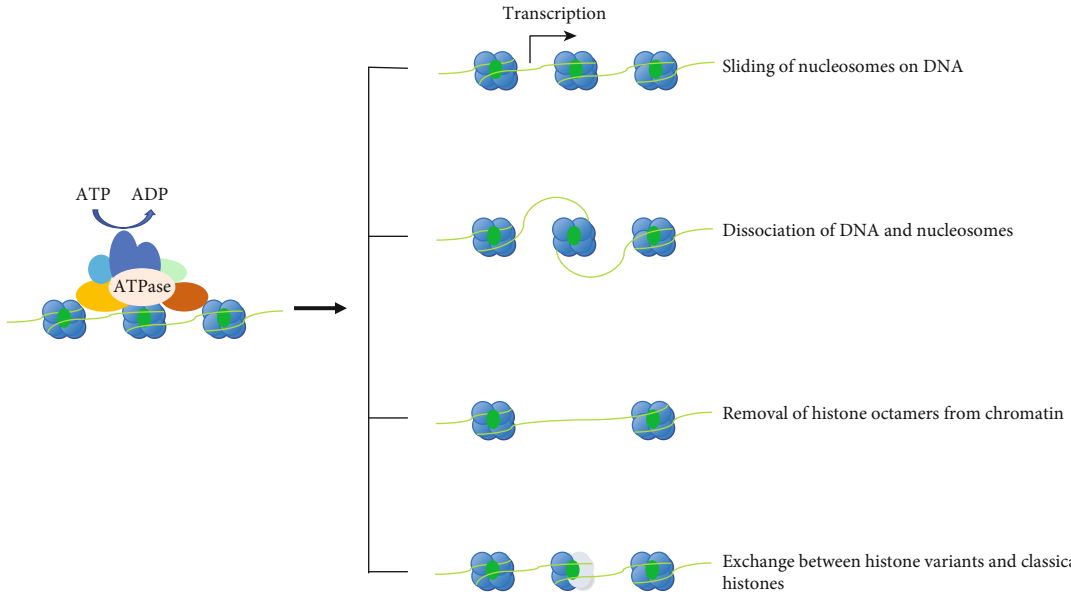


Figure 1: Schematic of the chromatin remodeling complex functional mode. Chromatin remodeling complex, which has ATPase activity, could change the structure of nucleosomes with the energy generated by hydrolyzing ATP, to regulate the accessibility of chromatin and further affect gene expression.

According to the difference in structure and composition of ATPase, chromatin remodeling complexes are divided into four categories: switching (SWI)/sucrose nonfermenting (SNF) [16, 17], INO80 [18], ISWI (imitation SWI) [19], and CHD (chromodomain helicase DNA binding) [20].

3. Structure and Function of SWI/SNF

The SWI/SNF complex is first discovered in yeast [16] and later in Drosophila [21] and mammals [22, 23]. The mammalian SWI/SNF complex, also named BAF (BRG1/BRM-associated factor) complex, is a multi-subunit protein complex of about 2 MDa, which is composed of 12-15 subunits encoded by 29 genes [24]. According to the different composition of subunits, BAF complexes are divided into canonical BAF (cBAF), PBAF, and noncanonical (ncBAF) [25]. The structural characteristics of the three types of SWI/SNF complexes are shown in Figure 2. Recent studies reveal the assembly process of these three types of BAF complexes (Figure 2) [25, 26] In different developmental stages and different tissues, the composition of the BAF complex also changes to regulate distinct gene expression, thereby performing different functions [27].

4. The Role of BAF Complex in mESCs

esBAF, a specific BAF complex in ESCs, consists of 9-11 subunits, which includes the ATPase subunit BRG1 not BRM, BAF250a instead of BAF200, BAF60a/b instead of BAF60c, and BAF155 dimer instead of BAF155 and BAF170 (Figure 3) [28]. Numerous studies reveal the functional importance of the BAF complex in ESCs and embryonic development [28–30]. Here, we summarize the roles of various subunits of the esBAF complex in ESCs (Table 1) and embryonic development (Table 2).

4.1. BRG1. As the core catalytic subunit of the esBAF complex, BRG1 alone can reshape nucleosomes in vitro, but the efficiency is very low. The smallest complex of four subunits, BAF155, BAF170, Baf47, and BRG1, can exert catalytic activity efficiently [44].

BRG1 participates in chromatin remodeling to maintain ESC self-renewal and pluripotency [28, 31]. The absence of Brg1 results in the impairment of ESC self-renewal and pluripotency [28, 31, 32]. Deletion of Brg1 leads to the decreased expression of Oct4 and Sox2 and increased expression of lineage-specific genes, indicating its function in ESC differentiation [28, 31]. The BRG1 null embryos die at the blastocyst stage. ES cells cannot be isolated from Brg1-deficient embryos [40, 45].

BRG1 directly binds the promoter regions of Oct4, Sox2, and Nanog genes, indicating its regulatory roles on the expression of core pluripotency genes. Consistently, BRG1 interacts directly with NANOG, OCT4, and SOX2 and binds with many of their common target genes [30, 46].

In addition, BRG1 also regulates the expression of ESC-related genes by participating in LIF/STAT3 signaling pathways [47]. Leukemia Inhibitory Factor (LIF) is required to maintain the pluripotency of mESCs and naïve human ESCs [5, 6, 48]. In mESCs, the binding of BRG1 and STAT3 colocalize extensively on the genome [30, 47]. The binding of STAT3 to genes associated with pluripotency depends on the presence of the catalytic subunit BRG1 in the esBAF complex, which loosens the chromatin structure at the target gene of STAT3 and thus responds to the LIF signal [47]. BRG1 can enhance the LIF-STAT3 signaling pathway by antagonizing the PcG complex [47]. On the other hand, BRG1 and the

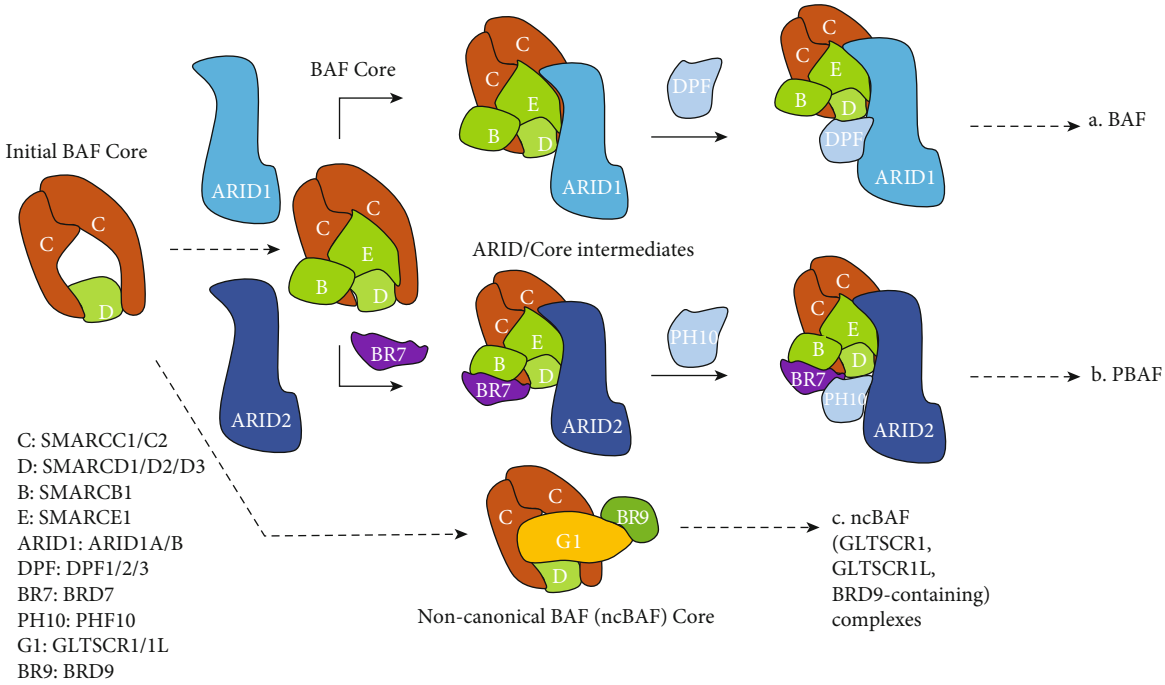


Figure 2: Three types of SWI/SNF assembly processes. (a) BAF: ARID1 and BAF core modules form a subcomplex (ARID1/BAF core), then combine with DPF2, and recruit ATPase modules (including SS18) to complete BAF assembly; (b) PBAF: ARID2 first combines with BAF core to form a subcomplex (ARID2/PBAF core), then combines BRD7 and PHF10, followed by the recruitment of ATPase module (excluding SS18), and finally combines with PBRM1 to complete PBAF assembly; (c) ncBAF: GLTSCAR1/1L BRD9 combines with BAF core module to form the core module of ncBAF and combines BRD9 with ATPase module (containing SS18) to form the ncBAF complex [25, 26].

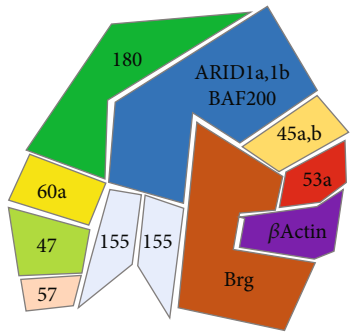


Figure 3: The subunits constituent of the esBAF complex.

PRC complex bind together to four Hox loci, thereby inhibiting the differentiation of ESCs [47].

Recently, YY1 was reported to interact with BRG1 to promote proliferation and pluripotency of mouse ESCs. The knockdown of Yy1 gene downregulates Nanog and upregulates differentiation marker genes Pax3 and Cdx2 [49].

The interaction between BRG1 and TOP2 is required for the initial stage of accessibility induction. Top2 can make the chromatin more accessible for chromatin remodelers as well as transcription factors [50], suggesting that TOP2 may work together with the BAF complex to remodel chromatin and optimize BAF-mediated recruitment of transcriptional factors.

4.2. BAF47. BAF47 (also known as SMARCB1/SNF5/INI1) is involved in the differentiation of stem cells. The knockdown of BAF47 enhances cell pluripotency and prevents differentiation [33]. Overexpressing BAF47 promotes ESC differentiation. BAF47 can fine-tune the level of OCT4 and affect the nucleosome occupation at the regulatory region of OCT4 target genes, thus breaking the balance between pluripotency and differentiation and determining the fate of cells [33]. In contrast, a recent report indicates the upregulated Cdx2 expression in Baf47 KO ESCs [34]. Therefore, further study to clarify the function of Baf47 in ESCs is needed.

The BAF47 null blastocysts do not hatch and cannot implant into the uterus for further development [41, 42], which may cause death of Baf47 null embryos during implantation [41, 42].

4.3. BAF155 and BAF170. BAF155 (also known as SRG3) shares 61.7% amino acid homology with BAF170, but they have different functions [28]. The esBAF complex contains a homodimer of two BAF155 without BAF170 [28]. The deletion of BAF155 resulted in the defects of ESC self-renewal and pluripotency [28]. As expected, overexpression of BAF170 cannot restore the defects of Brg155 KO ESCs [28]. Similarly, knockdown of Baf155 expression also resulted in inhibited ESC proliferation, decreased expression of the pluripotent gene Oct4, and increased apoptosis [28]. Consistently, deletion of BAF155 fails to form inner cell mass [51].

In contrast to mESCs, esBAF in hESCs contains heterodimers composed of BAF155 and BAF170. The contents of BAF155 and BAF170 in the BAF complex seem to determine the fate of the cell [52].

Table 1: The role of BAF subunits in mESCs.

Subunit	Phenotypes	References
BRG1	Knockdown or knockout of Brg1 resulted in ESC differentiation and downregulation of self-renewal and pluripotency genes such as Oct4 and Sox2.	[28, 31, 32]
BAF47	Knockdown of Baf47 blocks differentiation; overexpression of Baf47 enhances differentiation; knockdown of Baf57 upregulates Cdx2 expression.	[33, 34]
BAF155	Depletion of BAF155 resulted in decreased proliferation, decreased Oct4 expression, and increased apoptosis of ESCs.	[28]
BAF250a	The self-renewal ability of mESCs decreases after knocking out BAF250a, and the differentiation of ES cells into the mesoderm and endoderm is inhibited.	[35, 36]
BAF45d	Knockout of BAF45d perturbs ESC self-renewal and impairs its differentiation to three lineages.	[30]
BAF53a	Knockdown of Baf53a reduces the expression of pluripotent genes in ESCs. Baf53a protects mESCs from differentiating into primitive endoderm; knockout of Baf53a represses cell proliferation and induces cell apoptosis.	[37, 38]
BRD9	Preserving the naive pluripotency of ESCs	[39]

Table 2: The role of BAF subunits in early mouse embryonic development.

Subunit	Phenotypes	References
BRG1	Brg1 null embryos die during implantation, and mice heterozygous for Brg1 are prone to cause tumor formation and anencephaly.	[32, 40]
BAF47	Baf47 null mice die during embryo implantation. Baf47 heterozygous mice are prone to cause anencephaly.	[41, 42]
BAF155	Baf155 knockout embryos are lethal during implantation.	[29, 34]
BAF250a	Baf250a knockout embryos die on E6.5. Baf250a regulates heart development.	[35, 43]

The deletion of BAF155 prevented mouse embryos from developing properly and died during implantation [29]. Depletion of BAF155 leads to increased expression of Nanog in the ICM and its ectopic expression in TE. However, the overexpression of BAF155 leads to the development arrested at the E3.5 to E4.5 transition and upregulation of Cdx2 and Sox17 at E4.5 embryos [29].

4.4. BAF53a. BAF53a (also known as ACTL6a or ARP4) is expressed in a variety of stem/progenitor cells, including neural progenitor cells, hematopoietic stem cells, epidermal progenitor cells, and ES cells [37, 38, 53]. The knockdown of BAF53a in ESCs reduces the expression of pluripotent genes such as Oct4 and Nanog and induces ESC differentiation towards the original endoderm [37]. It is interesting that another report indicates that knockout of Baf53a increases the expression of Oct4 and Nanog. Deletion of Baf53a repressed cell proliferation and induced apoptosis [54].

4.5. BAF45. BAF45 has two PHD domains, which can recruit the BAF complex to specific histone modification sites [55]. BAF45 includes four subunits: BAF45a, BAF45b, BAF45c, and BAF45d [30]. Only BAF45a and BAF45d are contained in esBAF [30]. BAF45a plays an important role in the maintenance of hematopoietic stem cells [56], but its role in mESCs is not clear. BAF45d, also known as Dpf2, is widely expressed in a variety of cells [30]. Deletion of Dpf2 in mESCs leads to the differentiation defects, which cannot be restored by BAF45a and BAF45c [30]. Further study demonstrates that Dpf2 regulates ESC differentiation by regulating Tbx3 expression [30].

4.6. BAF250a. BAF250a (ARID1A) is a unique subunit of esBAF, which belongs to the trithorax group (TrxG) family [57, 58]. BAF250a is abundantly expressed in early mouse embryos and ESCs [35, 36]. Deletion of BAF250a inhibits ESC self-renewal and upregulates the expression of the primitive endoderm marker genes in ESCs [35, 36]. The lack of BAF250a prevents ESCs from developing into mesoderm-derived cardiomyocytes, adipocytes, and skeletal muscle cells, but can differentiate into ectoderm-derived nerve cells [35, 43].

BAF250a is necessary for the development of early embryos. The loss of BAF250a caused the development of early embryos (E6.5) of mice to stagnate, and the lack of mesoderm prevented further development of gastrulation embryos [35].

4.7. ncBAF in ESCs. Gatchalian and colleagues found the existence of ncBAF in mESCs that puts BRD9 as the core [39]. Compared with esBAF, ncBAF lacks BAF47, BAF57, and ARID1A subunits. The knockdown of BRD9, the core subunit of the ncBAF complex, inhibited the proliferation of ESCs [39]. Although both esBAF and ncBAF are involved in ESC self-renewal and pluripotency maintenance, ChIP-seq analysis showed that esBAF and ncBAF complexes target distinct sites in the genome and cobound with different pluripotent transcription factors (TFs) [30, 39, 45]. esBAF tends to bind to active enhancers rich in h3k4me1 modification [30], while ncBAF is more likely to bind to promoter regions rich in h3k4me3 [39]. Different from esBAF, ncBAF tends to cobind with KLF4 and CTCF, indicating its distinct mechanisms from cBAF in ESCs [39].

In summary, different components of the BAF complex function differentially in ESC maintenance and differentiation. Deletion of core subunits such as Brg1, Baf155, or Baf250a reduced the expression of Oct4, Sox2, and Nanog, the key pluripotency genes of ES cells [28, 31]. On the contrary, Baf47 negatively regulated Oct4 expression in ESCs [33]. Deletion of Baf250a promoted the expression of endoderm marker genes Gata4 and Gata6 [35], while deletion of Baf45d decreased Tbx3 expression and impaired mesoendoderm differentiation [30]. During embryonic development, knockout of Brg1, Baf155, or Baf250a led to embryo death during peri-implantation [40–42, 51]. Deletion of Baf250a resulted in embryo death in later embryonic development stage [35].

Consistently, BAF complexes also play important roles in the reprogramming of somatic cells to induced pluripotent stem cells (iPSCs). Depletion of Brg1 leads to the failures in reprogramming [59, 60]. Overexpression of Brg1 and Baf155 increases the reprogramming efficiency of MEFs to iPSCs [61], whereas downregulation of Brm and Baf170 improves reprogramming efficiency [62]. Therefore, similar to the distinct roles of different BAF subunits for the maintenance and differentiation of ES cells, different BAF components also play different roles in the reprogramming.

5. Conclusion

BAF complexes are functionally important for the self-renewal and development of ESCs and mouse embryonic development. Deletion of different subunits in ESCs and embryos results in distinct phenotypes in ESC maintenance and differentiation and embryonic development, while the underlying mechanisms are far from clear. Schick et al.'s work reveals that the loss of a single subunit of the BAF complex did not destroy the entire complex, but will change the composition of the BAF complex [24]. Consistently, a recent study shows that deletion of Dpf2 only affects about 8% of BRG1 binding sites on the genome [30]. Therefore, it is attractive to propose that distinct BAF subunit controls the integrity of a part of the BAF complex on the genome, and therefore, its deletion only affects the binding of a part of the BAF complex, which directly changes the expression of distinct pluripotency TFs in both ESCs and differentiating cells with other TFs and chromatin modifiers. It is intriguing to extend the proposed mechanism further to other chromatin remodeling complexes. To confirm the proposal, future works are required to study the deletion of specific subunits on the binding of BRG1 and some other core factors of BAF and other chromatin remodeling complexes.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Heyao Zhang, Xuepeng Wang, and Jinsheng Li are co-first authors.

References

[1] M. J. Evans and M. H. Kaufman, "Establishment in culture of pluripotential cells from mouse embryos," *Nature*, vol. 292, no. 5819, pp. 154–156, 1981.

[2] G. R. Martin, "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, no. 12, pp. 7634–7638, 1981.

[3] J. A. Thomson, J. Itskovitz-Eldor, S. S. Shapiro et al., "Embryonic stem cell lines derived from human blastocysts," *Science*, vol. 282, no. 5391, pp. 1145–1147, 1998.

[4] M. A. Surani, K. Hayashi, and P. Hajkova, "Genetic and epigenetic regulators of pluripotency," *Cell*, vol. 128, no. 4, pp. 747–762, 2007.

[5] T. Matsuda, T. Nakamura, K. Nakao et al., "STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells," *The EMBO Journal*, vol. 18, no. 15, pp. 4261–4269, 1999.

[6] H. Niwa, T. Burdon, I. Chambers, and A. Smith, "Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3," *Genes & Development*, vol. 12, no. 13, pp. 2048–2060, 1998.

[7] K. M. Loh and B. Lim, "A precarious balance: pluripotency factors as lineage specifiers," *Cell Stem Cell*, vol. 8, no. 4, pp. 363–369, 2011.

[8] M. Thomson, S. J. Liu, L. N. Zou, Z. Smith, A. Meissner, and S. Ramanathan, "Pluripotency factors in embryonic stem cells regulate differentiation into germ layers," *Cell*, vol. 145, no. 6, pp. 875–889, 2011.

[9] J. Kim, J. Chu, X. Shen, J. Wang, and S. H. Orkin, "An extended transcriptional network for pluripotency of embryonic stem cells," *Cell*, vol. 132, no. 6, pp. 1049–1061, 2008.

[10] L. A. Boyer, K. Plath, J. Zeitlinger et al., "Polycomb complexes repress developmental regulators in murine embryonic stem cells," *Nature*, vol. 441, no. 7091, pp. 349–353, 2006.

[11] T. I. Lee, R. G. Jenner, L. A. Boyer et al., "Control of developmental regulators by Polycomb in human embryonic stem cells," *Cell*, vol. 125, no. 2, pp. 301–313, 2006.

[12] W. W. Tee and D. Reinberg, "Chromatin features and the epigenetic regulation of pluripotency states in ESCs," *Development*, vol. 141, no. 12, pp. 2376–2390, 2014.

[13] A. Marson, S. S. Levine, M. F. Cole et al., "Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells," *Cell*, vol. 134, no. 3, pp. 521–533, 2008.

[14] A. Harikumar and E. Meshorer, "Chromatin remodeling and bivalent histone modifications in embryonic stem cells," *EMBO Reports*, vol. 16, no. 12, pp. 1609–1619, 2015.

[15] N. Hasan and N. Ahuja, "The emerging roles of ATP-dependent chromatin remodeling complexes in pancreatic cancer," *Cancers*, vol. 11, no. 12, 2019.

[16] L. Neigeborn and M. Carlson, "Genes affecting the regulation of SUC2 gene expression by glucose repression in *Saccharomyces cerevisiae*," *Genetics*, vol. 108, no. 4, pp. 845–858, 1984.

[17] M. Stern, R. Jensen, and I. Herskowitz, "Five _SWI_ genes are required for expression of the _HO_ gene in yeast," *Journal of Molecular Biology*, vol. 178, no. 4, pp. 853–868, 1984.

[18] R. C. Conaway and J. W. Conaway, "The INO80 chromatin remodeling complex in transcription, replication and repair," *Trends in Biochemical Sciences*, vol. 34, no. 2, pp. 71–77, 2009.

[19] K. Klement, M. S. Luijsterburg, J. B. Pinder et al., "Opposing ISWI- and CHD-class chromatin remodeling activities orchestrate heterochromatic DNA repair," *The Journal of Cell Biology*, vol. 207, no. 6, pp. 717–733, 2014.

[20] M. Murawska and A. Brehm, "CHD chromatin remodelers and the transcription cycle," *Transcription*, vol. 2, no. 6, pp. 244–253, 2014.

[21] J. A. Kennison and J. W. Tamkun, "Trans-regulation of homeotic genes in Drosophila," *The New Biologist*, vol. 4, no. 2, pp. 91–96, 1992.

[22] A. P. Wolffe, "Transcriptional activation: switched-on chromatin," *Current Biology*, vol. 4, no. 6, pp. 525–528, 1994.

[23] T. Owen-Hughesand and J. L. Workman, "Experimental analysis of chromatin function in transcription control," *Critical Reviews in Eukaryotic Gene Expression*, vol. 4, no. 4, pp. 403–441, 1994.

[24] S. Schick, A. F. Rendeiro, K. Runggatscher et al., "Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethalities in human cancers," *Nature Genetics*, vol. 51, no. 9, pp. 1399–1410, 2019.

[25] N. Mashtalir, A. R. D'Avino, B. C. Michel et al., "Modular organization and assembly of SWI/SNF family chromatin remodeling complexes," *Cell*, vol. 175, no. 5, pp. 1272–1288.e20, 2018.

[26] S. He, Z. Wu, Y. Tian et al., "Structure of nucleosome-bound human BAF complex," *Science*, vol. 367, no. 6480, pp. 875–881, 2020.

[27] E. Y. Sonand and G. R. Crabtree, "The role of BAF (mSWI/SNF) complexes in mammalian neural development," *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, vol. 166C, no. 3, pp. 333–349, 2014.

[28] L. Ho, J. L. Ronan, J. Wu et al., "An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 13, pp. 5181–5186, 2009.

[29] M. Panamarova, A. Cox, K. B. Wicher et al., "The BAF chromatin remodelling complex is an epigenetic regulator of lineage specification in the early mouse embryo," *Development*, vol. 143, no. 8, pp. 1271–1283, 2016.

[30] W. Zhang, C. Chronis, X. Chen et al., "The BAF and PRC2 complex subunits Dpf2 and Eed antagonistically converge on Tbx3 to control ESC differentiation," *Cell Stem Cell*, vol. 24, no. 1, pp. 138–152.e8, 2019.

[31] B. L. Kidder, S. Palmer, and J. G. Knott, "SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells," *Stem Cells*, vol. 27, no. 2, pp. 317–328, 2009.

[32] N. Singhal, D. Esch, M. Stehling, and H. R. Schöler, "BRG1 is required to maintain pluripotency of murine embryonic stem cells," *Biores Open Access*, vol. 3, no. 1, pp. 1–8, 2014.

[33] J. S. You, D. D. De Carvalho, C. Dai et al., "SNF5 is an essential executor of epigenetic regulation during differentiation," *PLoS Genetics*, vol. 9, no. 4, article e1003459, 2013.

[34] M. Sakakura, S. Ohta, M. Yagi et al., "Smarcb1 maintains the cellular identity and the chromatin landscapes of mouse embryonic stem cells," *Biochemical and Biophysical Research Communications*, vol. 519, no. 4, pp. 705–713, 2019.

[35] X. Gao, P. Tate, P. Hu, R. Tjian, W. C. Skarnes, and Z. Wang, "ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 18, pp. 6656–6661, 2008.

[36] I. Lei, S. Tian, V. Chen, Y. Zhao, and Z. Wang, "SWI/SNF component BAF250a coordinates OCT4 and WNT signaling pathway to control cardiac lineage differentiation," *Frontiers in Cell and Developmental Biology*, vol. 7, p. 358, 2020.

[37] W. Lu, L. Fang, B. Ouyang et al., "Actl6a protects embryonic stem cells from differentiating into primitive endoderm," *Stem Cells*, vol. 33, no. 6, pp. 1782–1793, 2015.

[38] V. Krasteva, M. Buscarlet, A. Diaz-Tellez, M. A. Bernard, G. R. Crabtree, and J. A. Lessard, "The BAF53a subunit of SWI/SNF-like BAF complexes is essential for hemopoietic stem cell function," *Blood*, vol. 120, no. 24, pp. 4720–4732, 2012.

[39] J. Gatchalian, S. Malik, J. Ho et al., "A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells," *Nature Communications*, vol. 9, no. 1, p. 5139, 2018.

[40] S. Bultman, T. Gebuhr, D. Yee et al., "A _Brg1_ null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes," *Molecular Cell*, vol. 6, no. 6, pp. 1287–1295, 2000.

[41] A. Klochendler-Yeivin, L. Fiette, J. Barra, C. Muchardt, C. Babinet, and M. Yaniv, "The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression," *EMBO Reports*, vol. 1, no. 6, pp. 500–506, 2000.

[42] C. J. Guidi, A. T. Sands, B. P. Zambrowicz et al., "Disruption of Ini1 leads to peri-implantation lethality and tumorigenesis in mice," *Molecular and Cellular Biology*, vol. 21, no. 10, pp. 3598–3603, 2001.

[43] I. Lei, J. West, Z. Yan et al., "BAF250a protein regulates nucleosome occupancy and histone modifications in priming embryonic stem cell differentiation," *Journal of Biological Chemistry*, vol. 290, no. 31, pp. 19343–19352, 2015.

[44] M. L. Phelan, S. Sif, G. J. Narlikar, and R. E. Kingston, "Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits," *Molecular Cell*, vol. 3, no. 2, pp. 247–253, 1999.

[45] S. J. Bultman, J. I. Herschkowitz, V. Godfrey et al., "Characterization of mammary tumors from _Brg1_ heterozygous mice," *Oncogene*, vol. 27, no. 4, pp. 460–468, 2008.

[46] L. Ho, R. Jothi, J. L. Ronan, K. Cui, K. Zhao, and G. R. Crabtree, "An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 13, pp. 5187–5191, 2009.

[47] L. Ho, E. L. Miller, J. L. Ronan, W. Q. Ho, R. Jothi, and G. R. Crabtree, "esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function," *Nature Cell Biology*, vol. 13, no. 8, pp. 903–913, 2011.

[48] C. Buecker, R. Srinivasan, Z. Wu et al., "Reorganization of enhancer patterns in transition from naive to primed pluripotency," *Cell Stem Cell*, vol. 14, no. 6, pp. 838–853, 2014.

[49] J. Wang, X. Wu, C. Wei et al., "YY1 positively regulates transcription by targeting promoters and super- enhancers through the BAF complex in embryonic stem cells," *Stem Cell Reports*, vol. 10, no. 4, pp. 1324–1339, 2018.

[50] E. L. Miller, D. C. Hargreaves, C. Kadoch et al., "TOP2 synergizes with BAF chromatin remodeling for both resolution and

formation of facultative heterochromatin," *Nature Structural & Molecular Biology*, vol. 24, no. 4, pp. 344–352, 2017.

[51] J. K. Kim, S. O. Huh, H. Choi et al., "Srg3, a mouse homolog of yeast SWI3, is essential for early embryogenesis and involved in brain development," *Molecular and Cellular Biology*, vol. 21, no. 22, pp. 7787–7795, 2001.

[52] X. Zhang, B. Li, W. Li et al., "Transcriptional repression by the BRG1-SWI/SNF complex affects the pluripotency of human embryonic stem cells," *Stem Cell Reports*, vol. 3, no. 3, pp. 460–474, 2014.

[53] X. Bao, J. Tang, V. Lopez-Pajares et al., "ACTL6a enforces the epidermal progenitor state by suppressing SWI/SNF- dependent induction of KLF4," *Cell Stem Cell*, vol. 12, no. 2, pp. 193–203, 2013.

[54] B. Zhu, A. Ueda, X. Song, S. I. Horike, T. Yokota, and T. Akagi, "Baf53a is involved in survival of mouse ES cells, which can be compensated by Baf53b," *Scientific Reports*, vol. 7, no. 1, p. 14059, 2017.

[55] C. Kadoch and G. R. Crabtree, "Mammalian SWI/SNF chromatin remodeling complexes and cancer: mechanistic insights gained from human genomics," *Science Advances*, vol. 1, no. 5, article e1500447, 2015.

[56] V. Krasteva, G. R. Crabtree, and J. A. Lessard, "The BAF45a/PHF10 subunit of SWI/SNF-like chromatin remodeling complexes is essential for hematopoietic stem cell maintenance," *Experimental Hematology*, vol. 48, pp. 58–71.e15, 2017.

[57] Z. Nie, Y. Xue, D. Yang et al., "A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex," *Molecular and Cellular Biology*, vol. 20, no. 23, pp. 8879–8888, 2000.

[58] M. Vazquez, L. Moore, and J. A. Kennison, "The trithorax group gene Osa encodes an ARID-domain protein that genetically interacts with the brahma chromatin-remodeling factor to regulate transcription," *Development*, vol. 126, no. 4, pp. 733–742, 1999.

[59] C. Hansis, G. Barreto, N. Maltry, and C. Niehrs, "Nuclear reprogramming of human somatic cells by *Xenopus* egg extract requires BRG1," *Current Biology*, vol. 14, no. 16, pp. 1475–1480, 2004.

[60] D. Egli and K. Eggan, "Recipient cell nuclear factors are required for reprogramming by nuclear transfer," *Development*, vol. 137, no. 12, pp. 1953–1963, 2010.

[61] N. Singhal, J. Graumann, G. Wu et al., "Chromatin-remodeling components of the BAF complex facilitate reprogramming," *Cell*, vol. 141, no. 6, pp. 943–955, 2010.

[62] Z. Jiang, Y. Tang, X. Zhao, M. Zhang, D. M. Donovan, and X. C. Tian, "Knockdown of Brm and Baf170, components of chromatin remodeling complex, facilitates reprogramming of somatic cells," *Stem Cells and Development*, vol. 24, no. 19, pp. 2328–2336, 2015.