



Review

# Endogenous Retroviruses Function as Gene Expression Regulatory Elements During Mammalian Pre-implantation Embryo Development

Bo Fu <sup>1,2</sup> , Hong Ma <sup>1,2</sup> and Di Liu <sup>1,2,\*</sup>

<sup>1</sup> Institute of Animal Husbandry Research, HeiLongJiang Academy of Agricultural Sciences, Harbin 150086, China; fubohao810@163.com (B.F.); mahong197400@163.com (H.M.)

<sup>2</sup> Key Laboratory of Combine of Planting and Feeding, Ministry of Agriculture of the People's Republic of China, Harbin 150086, China

\* Correspondence: liudi1963616@163.com; Tel.: +86-138-4512-0192

Received: 24 December 2018; Accepted: 5 February 2019; Published: 12 February 2019



**Abstract:** Pre-implantation embryo development encompasses several key developmental events, especially the activation of zygotic genome activation (ZGA)-related genes. Endogenous retroviruses (ERVs), which are regarded as “deleterious genomic parasites”, were previously considered to be “junk DNA”. However, it is now known that ERVs, with limited conservatism across species, mediate conserved developmental processes (e.g., ZGA). Transcriptional activation of ERVs occurs during the transition from maternal control to zygotic genome control, signifying ZGA. ERVs are versatile participants in rewiring gene expression networks during epigenetic reprogramming. Particularly, a subtle balance exists between ERV activation and ERV repression in host–virus interplay, which leads to stage-specific ERV expression during pre-implantation embryo development. A large portion of somatic cell nuclear transfer (SCNT) embryos display developmental arrest and ZGA failure during pre-implantation embryo development. Furthermore, because of the close relationship between ERV activation and ZGA, exploring the regulatory mechanism underlying ERV activation may also shed more light on the enigma of SCNT embryo development in model animals.

**Keywords:** pre-implantation embryo; endogenous retroviruses; zygotic genome activation; epigenetic reprogramming; somatic cell nuclear transfer

## 1. Introduction

Transposable elements (TEs), which are the descendants of ancestral viruses, have colonized genomes and now make up about half of mammalian genomes [1–3]. TEs can be divided into DNA transposons and retrotransposons [4]. DNA transposons have not yet been characterized during early development, so they are not discussed in this review. In most mammals, retrotransposons are the predominant TEs. They occupy nearly 40% of mammalian genomes [1] and include long terminal repeat (LTR) retrotransposons (also known as endogenous retroviruses (ERVs)) and non-LTR retrotransposons represented by long interspersed elements (LINEs) and short interspersed elements (SINEs) [5–7]. The potential transposition activities of unconstrained ERV mobilization can cause insertional mutagenesis or chromosomal abnormality. Therefore, ERVs should be repressed through DNA methylation in terminally differentiated somatic cells, otherwise aberrant activation threatens genomic integrity and results in cancer and autoimmune disorders [8,9]. However, the transcriptional activation of retrotransposons, especially ERVs, is a species-specific and conserved biological process in early mammalian embryos. Transcriptional activation of ERVs induces zygotic genome activation (ZGA) when maternal control gives way to zygotic control, indicating

totipotency [10–15]. An important question is how ERVs, with limited conservatism across species, mediate conserved developmental processes such as ZGA. Early embryos provide an environment suitable for the transcription of ERVs during epigenetic reprogramming and ERVs may take advantage of developmental epigenomic reprogramming windows to evade silencing by chromatin-modifying enzymes. ERVs, as regulators of gene networks, play multiple critical roles during ZGA. In host–virus interplay, ERV activation is controlled by a multilayered regulatory network that maintains a balance between ERV activation and ERV repression, which results in stage-specific ERV expression during pre-implantation embryo development. The SCNT technique, an asexual reproductive tool, is still inefficient in domestic animal breeding and the establishment of animal disease models. Exploring the molecular properties of ERV activation may also provide clues on the mechanism underlying SCNT embryo development.

## 2. General Survey of ERV Elements in Genomes

Exogenous retroviruses that infect germ cells are integrated into the host genome and inherited by offspring. ERVs derived from exogenous retroviruses are the most abundant transposable elements, accounting for 10% and 8% of the mouse and human genomes, respectively [1,2]. ERVs evolve more rapidly than other TEs, and this has been confirmed by the finding that orthologous ERVs in humans and chimpanzees exhibit signatures of directional selection after the human–chimpanzee divergence that occurred five million years ago [16]. Intact ERV genes encode *env*, *gag*, and *pol* proteins that play key roles in replication and mobilization. These genes are flanked by LTRs that regulate ERV transcription. *Env* encodes *env*, an envelope protein that enables retroviruses to undergo an extracellular infectious phase. However, most ERVs (e.g., murine endogenous retrovirus-like MuERV-L/MERVL) do not have an *Env* gene and are incapable of horizontal transfer [17–19]. *Gag* encodes a specific retroviral antigen and *Pol* encodes integrase, ribonuclease, and reverse transcriptase [20]. ERVs can be divided into class I, class II, and class III elements based on the sequence of their reverse transcriptase gene [1]. The transcription of ERVs is initiated by a 5′LTR promoter, generating a terminally redundant mRNA that is translated into Gag and Gag–Pro–Pol fusion proteins. Then, the ERV mRNA is reverse transcribed into double-stranded cDNA containing the LTR. It is this cDNA copy that is finally integrated into the host genome via ERV integrase [21]. Unlike DNA transposons, ERVs are transcribed into an RNA intermediate that may then be reverse transcribed into cDNA and reintegrated into the host genome at another location. This “copy-and-paste” mechanism tends to increase the copy number of ERVs and thus tends to increase the size of the host genome. However, because of the loss of regulatory elements or protein-coding sequences, most ERVs in mammalian genomes cannot continue with genomic expansion and horizontal transfer [6]. For example, nearly all human ERVs (HERVs) have lost their transposition ability in the human genome [22,23].

Because of ERVs’ viral origin and the history of exposure of hosts to different exogenous retroviruses, genomic ERV content, as genomic parasites, varies significantly between species [6]. Unlike housekeeping activity, the expression of ERVs is species-specific because the LTR elements contain species-specific transcription factor binding sites [24]. ERVs that still have a replication ability rely on the host cell machinery to express their genes. In LTR elements of ERVs that flank the coding sequence of ERVs, *cis*-regulatory sequences and RNA polymerase II promoters are present [25]. During evolution, some ERV elements drift from their original location, whereas other ERVs remain intact, leaving a host genome that is patched with full-length and truncated, active, and inactivated ERVs elements [26]. Following insertion, ERVs often undergo non-allelic homologous recombination between flanking LTRs in *cis* sequences, which results in loss of the coding regions of ERVs, leaving solitary LTRs. Particularly, nearly 90% of HERVs exist in the human genome as solitary LTRs and often contain transcription factor binding sites that are species-specific [2,27]. Thus, transposition also provided an opportunity for genomes to gain novel transcription factor binding sites during mammalian evolution.

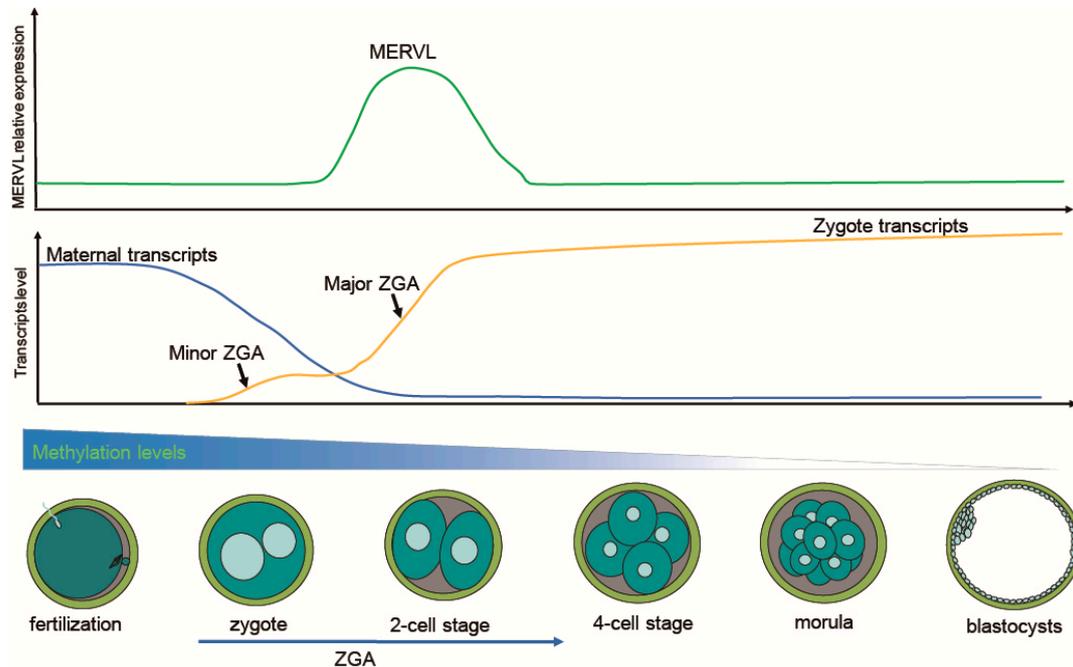
### 3. Transcriptional Activation of ERVs Signifies ZGA

Despite the different transcription profiles of ERVs between species, transcriptional activation of ERVs is a conserved event in early embryos [28]. Although ERVs have limited conservation across species, they have the potential to regulate ZGA. Intriguingly, it is not known how such a divergent element can mediate conserved developmental processes such as ZGA. ERVs were identified previously as deleterious genetic elements, and early embryos also employ numerous mechanisms to restrict the retrotransposition of ERVs during development [28]. Paradoxically, ERVs are broadly transcribed into tissue-specific genes or ERV-derived sequences in early embryos and stem cells [10–12,29–35]. The expression of ERVs contributes to the activation of the embryonic genome and to cellular plasticity [32], which is associated with the establishment of totipotency and pluripotency. A large number of ERV-derived sequences are activated, especially in embryonic and cancer cells, and this cell type-specific activation is associated with cell type-specific expression of neighboring genes [36]. MERVL is not expressed in oocytes, but its expression begins to increase after fertilization and peaks at the 2-cell stage, before gradually decreasing until the blastocyst stage [10]. HERV families, such as HERV-H and HERV-K, which are associated with early embryonic development, then signify an undifferentiated state, which indicates that expression profiles of HERV families may herald cell identity [30]. Recently, the high level expression of bovine endogenous retroviruses BERV-K1 and BERV-K2 was also detected in the embryonic blastomeres (2-cell to 16-cell stages) [37]. These suggest that the role of ERV activation is conservative between species. Cleavage stage embryos provide an environment particularly suitable for the transcription of ERVs, where most of the epigenetic markers are wiped off and then reestablished [38]. Remodeling of heterochromatic marks and a relaxed chromatin structure during pre-implantation embryo development provide a time window for the expression of ERVs [39]. The parental DNA methylation and histones are reset across the genome of the zygote, preparing for ZGA and ERV activation. The hydroxylation of 5mC into 5hmC is catalyzed by dioxygenases, the ten-eleven translocation (TET) proteins. The paternal genome undergoes genome-wide loss of DNA methylation via an active mechanism, because TET3-mediated hydroxylation of 5mC accounts for some of the active DNA demethylation of the paternal genome. DNA methyltransferase (DNMT) 1, which methylates hemi-methylated cytosines in CpG sequences, contributes to maintaining genome-wide methylation patterns during replication. Owing to the exclusion of the DNMT1 from the nucleus, the maternal genome is passively de-methylated in subsequent divisions [40]. Hyperaccessibility of chromatin, which is largely determined by histone modifications of its N-terminal tails, is a prerequisite for ZGA. Following fertilization, H3K4me3 and H4 acetylation, which make the male pronucleus permissive for active transcription, are responsible for a minor ZGA [41,42].

After fertilization, maternal inherited transcripts are progressively degraded to give way to the ZGA, setting transition from maternal to zygotic control [43,44]. In mouse embryos, the onset of ZGA has been shown to occur during the 1-cell stage (minor ZGA) [45]. Very weak transcription was observed during the minor ZGA stage [46–48]. The 2-cell stage mouse embryos (major ZGA) went through the switch from maternal control to zygotic genome control and were associated with the transient upregulation of many “2C genes” that contain promoters derived from the LTR elements of ERVs [49]. Dramatic transcriptional activation and robust translational activity occur during the major ZGA stage [44,50]. When ZGA failed, embryos failed to develop further [50]; therefore, ZGA may be one of the first critical events during pre-implantation embryo development.

Transcripts of ERVs that can drive cell transcriptional landscape and developmental stage-specific gene expression, occupy a large portion of the transcriptome during ZGA. For example, ERVs have contributed hundreds of thousands of novel regulatory elements to the human transcriptional landscape [36]. In particular, many critical ZGA-specific genes are regulated by LTR elements of ERVs [11]. Activation of the MERVL gene can be observed as early as 8 h after fertilization (minor ZGA) [10]. Downregulation of MERVL through RNA interference can result in developmental arrest at the 2-cell stage [51]. The transcriptional profile of 2-cell stage mouse embryos is characterized by

robust activation of MERVL and 2-cell-specific genes, such as *Zscan4*, *Zfp352*, and *Tdpoz* [52]. In general, the expression of ERVs signifies the onset of ZGA [10]. Several key events during ZGA have been shown in Figure 1.



**Figure 1.** Schematic illustration showing the dynamics of transcription and DNA demethylation during murine pre-implantation embryo development. Shortly after fertilization, embryos undergo extensive global DNA demethylation from the zygote to the blastocyst. Degradation of maternal transcripts is required for ZGA. Minor ZGA occurs at the one-cell stage, while major ZGA takes place at the two-cell stage. Particularly, the expression of MERVL peaks at the two-cell stage, and then gradually decreases until the blastocyst stage.

Although this paper mainly discusses the role of ERV activation in the ZGA process, we shall also bear in mind that ERV-derived transcripts play key roles in trophoblasts and the placenta. Sub-families of ERV proviruses, such as HERV-W and HERV-FRD, are expressed in trophoblasts. This expression is required for cell–cell fusion to enable the formation of syncytiotrophoblasts that are essential for invasive placental development, and the prevention of immune rejection of the fetus at the fetomaternal interface [53]. BERV-K3, belonging to bovine ERVs, is located between interleukin enhancer-binding factor 3 (*Ilf3*) and *Qtrt1* genes on chromosome 7. BERV-K3 has integrated within LOC100848658, from which noncoding RNA could be transcribed. After conceptus attachment to the endometrial epithelium, a high expression of BERV-K3 was detected in the placenta, which may associate with the bovine conceptus epithelial–mesenchymal transition (EMT) and/or its attachment to the uterine epithelium [54].

#### 4. ERV Elements Act as Gene Expression Regulators and are Versatile

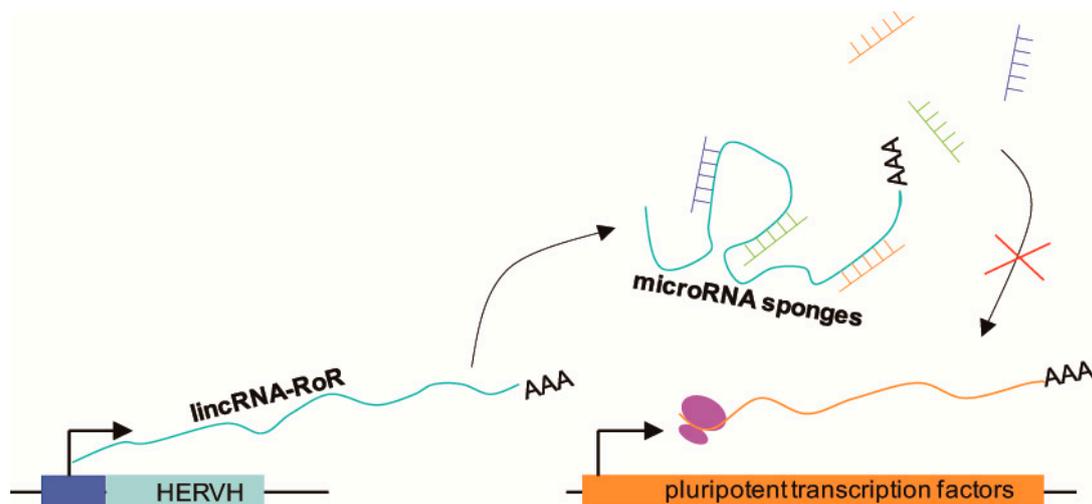
ERV-derived sequences create alternative promoters during ZGA. LTR elements provide a promoter at the 5' end of ERVs and transcriptional termination and polyadenylation signals at their 3' end. It should be noted that LTRs also contain bidirectional promoters that are able to initiate transcription in both the sense and antisense orientations. For example, in Human T-cell leukemia virus type 1 (HTLV-1), antisense transcription from the 3' LTR of HTLV-1 regulates the expression of *Hbz*, while sense transcription from the 5' LTR of HTLV-1 controls the expression of *Tax* [55]. Many ZGA-associated genes were controlled by LTR elements of ERVs. In mice, the LTR promoters of MERVL elements regulate gene networks that are specific to the 2-cell stage of embryonic development,

indicating totipotency [13]. A large percentage of ZGA-associated genes are located in proximity to LTR elements [13]. LTR elements contain a binding motif for DUX/DUX4 that is a pioneer transcription factor [24,56,57]. The solitary LTRs that occupy the majority of all HERVs in the human genome still maintain transcriptional and regulatory functions, which affect the expression of neighboring genes [58]. The conserved splice donor site located within LTRs may make LTRs into alternative promoters. During ZGA, these kinds of elements serve as powerful alternative promoters that enable the transcription of neighboring genes, resulting in 'chimeric' LTR–host transcripts [11]. Two-cell mouse embryos contain large numbers of chimeric gene transcripts that are identical to mRNA sequences of known host genes, except for the sequence at the 5' end, which is derived from MERV1 or other rodent ERV-L elements (e.g., MT2B and MT2C) [11]. In addition, several host genes such as *Tcstv1*, *Tcstv3*, *Ubtfl1*, *Chit1*, *Eif1a*, and *Zfp352* have transcripts initiating from LTR elements in mouse ESCs, and LTRs have been co-opted by cellular genes as promoters [59].

With pluripotent transcription factor binding sites located within LTRs, the expression of ERVs may also be induced by pluripotent transcription factors. For example, LTR7 that acts as a flanking LTR element in HERV-H contains Oct4, Nanog, Klf4, and Lbp9 transcription factor binding sites and can induce the transcriptional activation of HERV-H, which is a hallmark of naive-like human embryonic stem cells (hESCs) [33].

After fertilization, the assembly of high-order chromatin structure occurs, drastically changing from a condensed maternal and a naive paternal genome to a totipotent state. ERVs provide binding sites for a chromatin organizer, and then participate in the formation of a high-order chromatin structure. For example, CCCTC binding factor (CTCF) regulates the high-order chromatin structure in multiple ways, including participation in chromatin loop formation, connecting the long-range enhancer-promoter, and insulating epigenetic spreading [60]. Binding sites of CTCF are enriched in ERV elements [61]. When bound by CTCF, ERV elements can organize a high-order chromatin structure, thus promoting the global remodeling of chromatin architecture during pre-implantation embryo development [62,63].

ERV-derived long noncoding RNAs (lncRNAs) participate in the control of pluripotency. Biomarkers of pluripotency, such as Oct4, Sox2, and Nanog, promote the expression of ESC-specific genes and suppress differentiation [64,65]. In posttranscriptional networks, microRNAs (miRNAs) act as posttranscriptional modifiers and contribute to restraining pluripotency. For example, miR-145 represses Oct4, Sox2, and Klf4 by binding to their 3'UTRs, and miR-134, miR-296, and miR-470 target the coding DNA sequence of Nanog, Oct4, and Sox2 [66,67]. LincRNA-RoR is a long intergenic noncoding RNA derived from human endogenous retrovirus subfamily H (HERVH) [68]. LincRNA-RoR acts as a competing endogenous miRNA sponge and protects pluripotent transcription factors from miRNA-mediated degradation; therefore, it is necessary and sufficient to maintain pluripotency and control ESC differentiation [69] (Figure 2). Jens [70] found that the lncRNA HPAT5 can also function as an miRNA sponge for the let-7 miRNA family to positively regulate pluripotency in ESCs. In addition, Fort et al. [31] found numerous pluripotent lncRNAs that contained unannotated antisense, intergenic, and intronic transcripts derived from ERV elements. Particularly, many stem cell-specific transcription start sites are not associated with protein-coding genes, but with these kinds of ERV elements in mice and humans.



**Figure 2.** Schematic illustration showing how pluripotent transcription factors are protected from miRNA-mediated degradation. LincRNA-RoR, a long intergenic noncoding RNA, is transcribed from human endogenous retrovirus subfamily H (HERVH), and then acts as an miRNA sponge to prevent miRNA-mediated degradation of pluripotent transcription factor mRNAs. The intact pluripotent transcription factor mRNAs can then be translated.

The activation of ERVs may contribute to genome defense. LTR5HS, a subclass of HERV-K, is transcribed from an LTR at the 8-cell stage (ZGA in human embryos) and contains the Oct4-binding motif. Grow et al. [35] found that by binding to LTR5HS, Oct4 drove the expression of HERV-K proviruses, producing viral-like particles and Gag proteins in pre-implantation embryos. Then, the overexpression of Rec, HERV-K accessory protein, is sufficient to increase the level of virus restriction factors such as IFITM1 during this process and contributes to fighting against exogenous viral infections.

## 5. Regulation of ERV Activation

How exactly species-specific ERVs are activated to function as gene regulators in host cells and, conversely, how host cells defend themselves against ERV activation during the window of epigenetic reprogramming (pre-implantation embryo development) to prevent widespread retrotransposition, has been investigated. ERV transcripts are under acute surveillance by multilayered and interleaved systems that ensure a subtle balance between ERV activation and ERV repression, resulting in stage-specific ERV expression during pre-implantation embryo development.

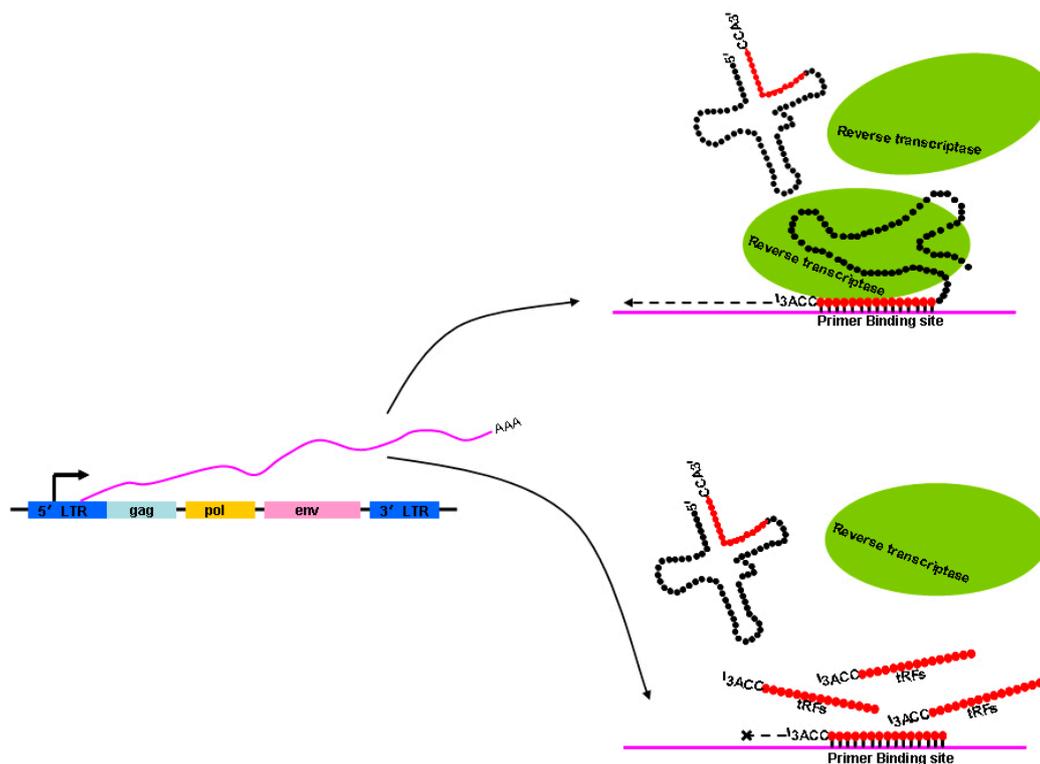
Unlike differentiated somatic cells with high DNA methylation levels, pre-implantation embryos undergo DNA methylation reprogramming, which is associated with establishment of the pluripotent state [71]. TET proteins catalyze oxidation reactions that convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), or further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [72–75]. The activity of TET proteins in a pre-implantation embryo allows ERVs to evade DNA methylation-mediated transcriptional repression. In response, numerous histone modifications predominate in silencing ERV expression in pre-implantation embryos and ESCs where DNA methylation mediated silencing is compromised [28,75,76]. Acquiring H3K9me3 while losing DNA methylation in LTR loci is a mark of epigenetic transformation [77]. H3K9me3 is enriched at the LTRs of many class I and class II ERVs in mouse ESCs (mESCs) [78]. Histone methyltransferase SETDB1 catalyzes the addition of methyl groups to H3K9 [79]. When SETDB1 was knocked out from mESCs, significant upregulation of several class I and class II ERVs was detected [80]. Correspondingly, removing KRAB-interacting scaffold protein KAP1, an SETDB1-interacting protein, also resulted in upregulation of the same ERVs in mESCs [80,81], indicating that KAP1, acting as a SETDB1–KAP1 complex recruiter, is also required for silencing ERVs [82]. KAP1 or its partner SETDB1 can repress

several classes of ERVs in ESCs, but not in embryonic fibroblasts, because DNA methylation takes over control of the ERVs in differential somatic cells [81,83,84].

Indeed, KAP1-mediated repression is sequence-specific, because KAP1 is recruited to repetitive sequences through site-specific Kruppel-associated box zinc-finger proteins (KRAB-ZFPs), the largest transcription factor family in mouse and human genomes. KRAB-ZFPs are characterized by the N-terminal KRAB domain and a tandem array of C2H2 zinc fingers [28,76]. KAP1 represses multiple classes of ERVs, including interstitial A-type particles and MERVK in mice or HERVK in humans, which indicates that ERVs may recruit KAP1 through different site-specific KRAB-ZFPs [85,86]. This hypothesis is supported by the observation that KRAB-ZFPs encoded from host genomes are species-specific and rapidly evolving, corresponding to the rapid evolution of viral sequences [87–91]. ERVs and the host genome are often referred to as an evolutionary arms race [92,93]. The continuous cycle of KRAB-ZFP evolution against ERVs may provide a driving force for new adaptations in mammals [83]. For example, the expansion of the *Znf91* subfamily across primate lineages reflects KRAB-ZFP adaptive evolution [94].

Unlike class I and class II ERVs, the repression of class III ERVs, including MERVL elements, is KDM1A (H3K4 demethylase)-mediated. MERVL sequences or LTR elements are targets of KDM1A repression. MERVL elements display increased methylation of histone H3K4, increased acetylation of H3K27, and decreased methylation of H3K9 in *Kdm1a* mutant ESCs. Therefore, significant upregulation of MERVL elements and their flanking LTRs occurs in *Kdm1a* mutant ESCs [59].

In addition to numerous pathways based on histone modifications, other mechanisms may function to control ERV activation. LTR retrotransposons use mature tRNAs that are selectively packaged into the intracellular virus-like particles (VLPs), and then bound by primer binding sites of the viral RNA as primers for reverse transcription, thus initiating the synthesis of minus-strand cDNAs [95]. Using the 3' terminus of mature tRNAs, the special primer for reverse transcription, to bind the primer binding sequence in ERV transcripts, ERVs start-up the reverse transcription process. However, tRNA-derived fragments (tRFs) that are 18-nt long can strongly inhibit the reverse transcription processes of ERV transcripts when 18-nt 3'-tRF binds to the primer binding site. Namely, 18-nt 3'-tRFs and mature tRNAs compete for primer binding sites at the same time [96] (Figure 3). In addition, the transposition activity of ERVs is inhibited by 22-nt 3'-tRFs that degrade the protein-coding mRNA of autonomous ERVs through post-transcriptional silencing with miRNA-like mechanisms that tolerate mismatch [88]. The 18-nt 3'-tRF and 22-nt 3'-tRF both contain a 3'-CCA at the end of their tRNA precursors, indicating that they both originated from mature tRNAs. A rapid increase in tRFs at the 8-cell stage was found in mouse pre-implantation embryos [97]. Particularly, tRFs are abundant in mature sperm [98,99] and can be delivered to a pre-implantation embryo through fertilization [100]. Therefore, tRFs derived from sperm may also inhibit the retrotransposition activity of ERVs in pre-implantation embryos where the genome lacks epigenetic markers such as DNA methylation and histone modification [101].



**Figure 3.** Schematic illustration showing how the reverse transcription processes of endogenous retrovirus (ERV) transcripts are repressed. The ERV genes *env*, *gag*, and *pol* are flanked by LTRs that regulate ERV transcription. The 3' terminus of intact mature tRNA is used as the special primer to complete the reverse transcription process. However, the process is interrupted when 18-nt 3'-tRF (tRNA-derived fragment) binds to the primer binding site.

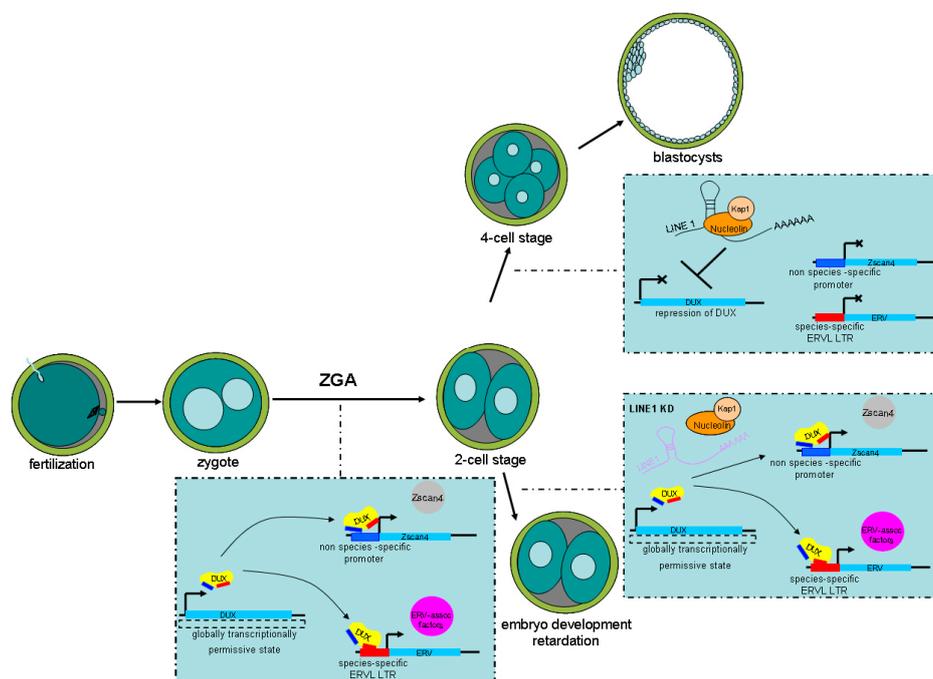
Homeodomain is a DNA-binding motif present in transcription factors. Double homeodomain-containing DUX-C proteins, located at the top of a transcriptional hierarchy, are expressed before the ZGA phase in placental mammals [24,102,103]. The human DUX-C homologue is DUX4, and the mouse DUX-C homologue is DUX. After fertilization, a globally transcriptionally permissive state caused by the loosening of chromatin and increased histone mobility may induce DUX-C activation [104,105]. In mammals, DUX-C family homologues (*DUX-C*, *DUX4*, and *DUX*) show a microsatellite tandem-array organization that, thus, has high copy numbers in the genomes [106]. The weak transcriptional activation of each *Dux-c* gene copy will result in a surge of the DUX-C protein. The two separate DNA-binding domains in DUX-C orthologues could provide an enhanced and synergistic binding ability. DUX4 can also bind at DNase I inaccessible sites, and then acetylate histone H3 at lysine 27, to open up chromatin via recruiting p300/CBP [107]. All these permit DUX-C orthologues to function as pioneer regulators of ZGA. DUX4, which is known to be associated with facioscapulohumeral dystrophy (FSHD), activates a large portion of the genes normally expressed at the onset of ZGA, particularly including cleavage-specific genes that have an upstream enriched binding motif for DUX4 [108]. When overexpressed in human-induced pluripotent stem cells, DUX4 also activated ZGA-related genes and retrotransposons, particularly HERV-L. Similarly, the expression of DUX peaked in 2-cell stage embryos (ZGA in mouse), and overexpression of DUX in mouse ESCs also activated ZGA-related genes. Deleting DUX in mouse zygotes caused ZGA failure and impaired developmental progression [56]. DUX upregulation also appeared in 2C-like cells that were identified by the reactivation of ERV elements such as MERVL, which then drives the transcription of many ZGA-related genes in embryos [13]. Together, the above phenomena indicate that DUX-C-family homologues serve as pioneer transcription factors that seem to drive the expression of ZGA-related genes such as *Zscan4* and regulate ERV activation, including MERVL. Furthermore, because of the

versatility of ERVs, their control by DUX proteins strengthens the transcriptional regulation potential of DUX in return.

Although mouse DUX and human DUX4 can both activate orthologous genes such as *Zscan4* in mouse and human myoblasts respectively, the overexpression of human DUX4 did not activate ZGA-related murine MERVL LTRs [57]. The species-specific transcription factor binding sites located within LTRs elements and the more divergent homeodomain (the first homeodomain) in DUX-C family proteins means that the activation of ERVs by DUX is species-specific, although DUX proteins still exhibit functional conservation because DUX-C family proteins have a more conserved homeodomain (the second homeodomain) [57]. The ancestral DUX proteins can regulate embryonic transcription conservatively, whereas the divergence of the DNA-binding motif of DUX-C family proteins indicates the adaption to species-specific activation of ERVs during evolution [24].

Expression of MERVL and activation of 2-cell stage genes induced by DUX are rapidly repressed after the 2-cell stage. Therefore, embryos exit the 2-cell state, otherwise ZGA will fail and embryonic development will be blocked [59,104]. LINE 1 is expressed throughout murine pre-implantation development [109]. In addition, human transposable elements such as

LINE 1 were also detected in human pre-implantation embryos [110]. Further research demonstrated that LINE1 knockdown caused persistence of the 2-cell state and failure of ZGA in murine pre-implantation embryos. In addition, LINE 1 RNA served as a nuclear RNA scaffold that recruited Nucleolin/Kap1 to repress DUX; thus, indirectly repressing the transcription of ERVs. LINE 1 depletion leads to failure to repress DUX/2-cell programming, which then leads to defective ZGA, resulting in embryo development retardation. Therefore, persistent LINE 1 RNA may be required for the repression of DUX and its associated ERV elements during pre-implantation embryo development [111] (Figure 4).



**Figure 4.** Schematic illustration showing how endogenous retrovirus (ERVs) activation is controlled by double homeobox DUX during murine pre-implantation development. After fertilization, a globally transcriptionally permissive state caused by the loosening of chromatin activates DUX. Then, DUX drives the expression of zygotic genome activation (ZGA)-related genes such as *Zscan4* and regulates ERV activation. LINE 1 RNA represses DUX by recruiting Nucleolin/Kap1, thereby indirectly repressing ZGA-related genes and ERV elements. This allows 2-cell embryos to develop into the 4-cell stage. In contrast, LINE 1 knockdown (KD) causes persistence of the 2-cell state and failure of ZGA.

## 6. Insights into Improving the Developmental Potential of SCNT Embryos

During somatic cell nuclear transfer (SCNT) reprogramming, differentiated somatic nuclei were transferred into enucleated oocytes and the reconstructed embryos acquired totipotency to produce viable cloned animals [112,113]. Although many mammalian species have been successfully cloned from various somatic cells, the SCNT technique is limited by its low efficiency because differentiated somatic cells often retain somatic chromatin states, even in reconstructed embryos [114]. In mouse SCNT embryos, nearly half of them display developmental arrest during pre-implantation embryo development [115]. The gene expression profiles of the somatic cells are reprogrammed to those of the 2-cell mouse embryos during the ZGA process. However, abnormal gene expression profiles were detected in early SCNT embryos, particularly in ZGA-stage SCNT embryos [116,117]. Besides mice, developmental defects have also been detected at the time of ZGA in other cloned species [118–120]. Unlike mouse embryos, ZGA occurs at the 8-cell stage in bovine embryos [121], which may give bovine embryos a longer time for ZGA-associated genes to be reprogrammed and activated sufficiently. Indeed, the developmental rates of cloned bovine embryos, which are similar to those of IVF bovine embryos [122], are relatively higher than the developmental rates of other cloned species [114]. It has been postulated that ZGA is necessary for somatic cell reprogramming [123]. Insufficient reprogramming and failure of ZGA are common defects of SCNT embryos, which impairs the final developmental potential of SCNT embryos [118–120]. ERVs such as MERVL can be activated specifically during ZGA [51,124–127]. With LTR elements serving as alternative promoters, ERVs drive the expression of hundreds of chimeric transcripts [11]. Therefore, the expression of ERVs is a hallmark of the ZGA process. However, compared with fertilized embryos, MERVL was downregulated at the 2-cell stage in the cloned mouse embryos. Furthermore, ZGA-associated genes such as *interferon- $\gamma$* , *Dub-1*, and *Spz1* were also inhibited in SCNT embryos [128]. The red fluorescent protein tandem dimeric tomato (tdTomato) reporter under the control of MERVL-LTR (MERVL:tdTomato) has been used to monitor ZGA events in real time [13]. Using transgenic mouse lines as the source of donor cells and sperm that contain the MERVL::tdTomato reporter, the expression pattern of ERVs such as MERVL can be observed in early SCNT embryos and intracytoplasmic sperm injection (ICSI) embryos in real time. Only 12% of SCNT embryos exhibited the reactivation of somatic MERVL::tdTomato during ZGA, whereas 92% of ICSI embryos exhibited the reactivation of MERVL::tdTomato during that stage [129]. Moreover, single-cell RNA-seq data also indicated that ZGA-associated genes could not be properly activated in tdTomato-SCNT embryos compared with normal fertilization embryos [129]. Given the key roles played by ERVs during ZGA, downregulated ERV expression may cause the failure of ZGA, leading to defects of SCNT embryo development.

The failure of ERV expression during ZGA in SCNT embryos may be induced by preexisting epigenetic barriers in the genomes of differentiated donor cells. Matoba et al. [130] found that ERVs are suppressed in SCNT embryos compared to IVF embryos and identified hundreds of reprogramming resistant regions (RRRs) in SCNT embryos. These RRRs, where ERV elements such as LTRs are located, are gene-poor regions and are specifically enriched with H3K9me3. H3K9me3-initiated heterochromatin assembly may prevent the activation of ERV elements in RRRs. In general, the conserved epigenetic markers that originated from differentiated donor cells preexisted in RRRs. Unfortunately, conserved epigenetic barriers impede the expression of ERVs and ZGA during SCNT reprogramming. Elimination of epigenetic barriers located in ERVs of RRRs may improve ZGA and SCNT efficiency. For instance, the ectopically expressed H3K9me3 demethylase Kdm4d can remove the inhibitory chromatin state in RRRs, thereby improving ZGA and SCNT efficiency [130]. In another case, Yang et al. [129] found that enrichment of H3K27me3 was also associated with the failure of ERV activation. KDM6A and KDM6B are H3K27me3-specific demethylases that are functionally redundant and compensate for each other, so a decrease in KDM6A or KDM6B expression is accompanied by an increase in KDM6B or KDM6A expression, respectively. H3K27me3 levels decrease when injected with KDM6B small interference RNA (siRNA), which increased KDM6A expression and correspondingly

increased MERVL levels. Finally, KDM6B siRNA injection improved the developmental potential of SCNT embryos by increasing ZGA-related ERV expression.

## 7. Conclusions

ERVs behave like retroviruses and play a variety of functions. The subtle balance between the activation and silencing of ERVs during pre-implantation embryo development implies a multilayered regulatory network is involved in host–virus interplay. An increase in ERV expression at specific developmental stages marks key events such as ZGA, and ERV activation can be inhibited by various mechanisms, such as DNA methylation, histone modification, post-transcriptional silencing, and transcription factor inactivation. From an evolutionary perspective, ERVs and the host genome are often regarded as being an arms race that is embodied by KRAB-ZFP adaptive evolution. ERVs and the host genome also have a cooperative symbiotic relationship that supports the amplification of ERVs in the host genome and provides the regulatory elements in gene networks. In particular, specific ERV activation (also known as the marker of ZGA) provides clues to elucidate the mechanism underlying SCNT embryo development. Although, due to ethical and technical constraints, regulating ERV activation should not be applied in clinical practice, ERV activation may be a key developmental event across species. Therefore, precisely regulating ERV activation may provide a new perspective for investigating the molecular mechanism underlying the ZGA process and improve the developmental potential of SCNT embryos in model animals.

**Acknowledgments:** This work was supported by grants from Heilongjiang Academy of Agricultural Sciences Incubation Project for National Natural Science Foundation of China (Grant Numbers:2018JJPY001), Heilongjiang Academy of Agricultural Sciences Doctoral Scientific Research Project (Grant Numbers:201507-32), The Postdoctoral Scientific Research Project of Heilongjiang Province (Grant Numbers:LBH-Q15130), National Natural Science Foundation of China (Grant Numbers:31872980,31671289,31201804), China Agricultural Research System (Grant Numbers:CARS-37), Project for Improving Innovative Capability of Scientific Institutions in Heilongjiang Province (Grant Numbers:YC2016D001), and Natural Science Foundation of Heilongjiang Province of China (Grant Numbers:JJ2018ZZ0082).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mouse Genome Sequencing Consortium; Chinwalla, A.T.; Cook, L.L.; Delehaunty, K.D.; Fewell, G.A.; Fulton, L.A.; Fulton, R.S.; Graves, T.A.; Hillier, L.W.; Mardis, E.R.; et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* **2002**, *420*, 520–562.
2. International Human Genome Sequencing Consortium; Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [[CrossRef](#)]
3. Koning, A.P.J.; De Wanjun, G.; Castoe, T.A.; Batzer, M.A.; Pollock, D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* **2011**, *7*, e1002384. [[CrossRef](#)] [[PubMed](#)]
4. Finnegan, D.J. Eukaryotic transposable elements and genome evolution. *Trends Genet.* **1989**, *5*, 103–107. [[CrossRef](#)]
5. Ostertag, E.M.; Kazazian, H.H., Jr. Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* **2001**, *35*, 501–538. [[CrossRef](#)]
6. Mager, D.L.; Stoye, J.P. Mammalian Endogenous Retroviruses. *Microbiol. Spectr.* **2015**, *3*. [[CrossRef](#)]
7. Richard, C.; Batzer, M.A. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **2009**, *10*, 691–703.
8. Lamprecht, B.; Walter, K.S.; Kumar, R.; Hummel, M.; Lenze, D.; Kochert, K.; Bouhleb, M.A.; Richter, J.; Soler, E.; Stadhouders, R. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat. Med.* **2010**, *16*, 571–579. [[CrossRef](#)]
9. Nexø, B.A.; Christensen, T.; Frederiksen, J.; Møller-Larsen, A.; Oturai, A.B.; Villesen, P.; Hansen, B.; Nissen, K.K.; Laska, M.J.; Petersen, T.S. The etiology of multiple sclerosis: Genetic evidence for the involvement of the human endogenous retrovirus HERV-Fc1. *PLoS ONE* **2011**, *6*, e16652. [[CrossRef](#)]

10. Daisuke, K.; Naojiro, M.; Hanae, T.; Hiroshi, I. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Boil. Reprod.* **2003**, *68*, 651–654.
11. Peaston, A.E.; Evsikov, A.V.; Graber, J.H.; Vries, W.N.D.; Holbrook, A.E.; Solter, D.; Knowles, B.B. Retrotransposons Regulate Host Genes in Mouse Oocytes and Preimplantation Embryos. *Dev. Cell* **2004**, *7*, 597–606. [[CrossRef](#)]
12. Bui, L.C.; Evsikov, A.V.; Khan, D.R.; Archilla, C.; Peynot, N.; Hénaut, A.; Le Bourhis, D.; Vignon, X.; Renard, J.P.; Duranthon, V. Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. *Reproduction* **2009**, *138*, 289–299. [[CrossRef](#)] [[PubMed](#)]
13. Macfarlan, T.S.; Gifford, W.D.; Shawn, D.; Karen, L.; Rowe, H.M.; Dario, B.; Amy, F.; Oded, S.; Didier, T.; Pfaff, S.L. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **2012**, *487*, 57–63. [[CrossRef](#)] [[PubMed](#)]
14. Latham, K.E.; Schultz, R.M. Embryonic genome activation. *Front. Biosci.* **2001**, *6*, D748–D759. [[CrossRef](#)] [[PubMed](#)]
15. Graf, A.; Krebs, S.; Heininen-Brown, M.; Zakhartchenko, V.; Blum, H.; Wolf, E. Genome activation in bovine embryos: Review of the literature and new insights from RNA sequencing experiments. *Anim. Reprod. Sci.* **2014**, *149*, 46–58. [[CrossRef](#)] [[PubMed](#)]
16. Gemmell, P.; Hein, J.; Katzourakis, A. Orthologous endogenous retroviruses exhibit directional selection since the chimp-human split. *Retrovirology* **2015**, *12*, 52. [[CrossRef](#)] [[PubMed](#)]
17. Bénéit, L.; De Parseval, N.; Casella, J.F.; Callebaut, I.; Cordonnier, A.; Heidmann, T. Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the Fv1 restriction gene. *J. Virol.* **1997**, *71*, 5652–5657. [[PubMed](#)]
18. Bénéit, L.; Lallemand, J.B.; Casella, J.F.; Philippe, H.; Heidmann, T. ERV-L elements: A family of endogenous retrovirus-like elements active throughout the evolution of mammals. *J. Virol.* **1999**, *73*, 3301–3308. [[PubMed](#)]
19. Johnson, W.E. Endogenous Retroviruses in the Genomics Era. *Annu. Rev. Virol.* **2015**, *2*, 135–159. [[CrossRef](#)]
20. Eickbush, T.; Jamburuthugoda, V. The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res.* **2008**, *134*, 221–234. [[CrossRef](#)]
21. Schwartzberg, P.; Colicelli, J.; Goff, S.P. Construction and analysis of deletion mutations in the pol gene of Moloney murine leukemia virus: A new viral function required for productive infection. *Cell* **1984**, *37*, 1043–1052. [[CrossRef](#)]
22. Stoye, J.P. Endogenous retroviruses: Still active after all these years? *Curr. Biol.* **2001**, *11*, R914–R916. [[CrossRef](#)]
23. Boeke, J.D.; Stoye, J.P. *Retrotransposons, Endogenous Retroviruses, and the Evolution of Retroelements*; Cold Spring Harbor Lab Press: Cold Spring Harbor, NY, USA, 1997; pp. 343–436.
24. Hendrickson, P.G.; Doráis, J.A.; Grow, E.J.; Whiddon, J.L.; Lim, J.W.; Wike, C.L.; Weaver, B.D.; Pflueger, C.; Emery, B.R.; Wilcox, A.L. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **2017**, *49*, 925–934. [[CrossRef](#)] [[PubMed](#)]
25. Chuong, E.B.; Elde, N.C.; Feschotte, C. Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* **2017**, *18*, 71–86. [[CrossRef](#)] [[PubMed](#)]
26. Gifford, W.D.; Pfaff, S.L.; Macfarlan, T.S. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Boil.* **2013**, *23*, 218–226. [[CrossRef](#)] [[PubMed](#)]
27. Friedli, M.; Trono, D. The Developmental Control of Transposable Elements and the Evolution of Higher Species. *Annu. Rev. Cell Dev. Boil.* **2015**, *31*, 429–451. [[CrossRef](#)] [[PubMed](#)]
28. Rowe, H.M.; Didier, T. Dynamic control of endogenous retroviruses during development. *Virology* **2011**, *411*, 273–287. [[CrossRef](#)] [[PubMed](#)]
29. Mari, O.; Koji, T.; Kenta, S.; Ito, T.; Yuka, S.; Megumi, N.; Michiko, N.; Yumie, T.; Masahiro, N.; Akira, W. Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12426–12431.
30. Göke, J.; Lu, X.; Chan, Y.S.; Ng, H.H.; Ly, L.H.; Sachs, F.; Szczerbinska, I. Dynamic Transcription of Distinct Classes of Endogenous Retroviral Elements Marks Specific Populations of Early Human Embryonic Cells. *Cell Stem Cell* **2015**, *16*, 135–141. [[CrossRef](#)] [[PubMed](#)]
31. Fort, A.; Hashimoto, K.; Yamada, D.; Salimullah, M.; Keya, C.A.; Saxena, A.; Bonetti, A.; Voineagu, I.; Bertin, N.; Kratz, A. Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat. Genet.* **2014**, *46*, 558–566. [[CrossRef](#)] [[PubMed](#)]

32. Lu, X.; Sachs, F.; Ramsay, L.; Jacques, P.É.; Göke, J.; Bourque, G.; Ng, H.H. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.* **2014**, *21*, 423–425. [[CrossRef](#)] [[PubMed](#)]
33. Wang, J.; Xie, G.; Singh, M.; Ghanbarian, A.T.; Raskó, T.; Szvetnik, A.; Cai, H.; Besser, D.; Prigione, A.; Fuchs, N.V.; et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **2014**, *516*, 405–409. [[CrossRef](#)] [[PubMed](#)]
34. Fuchs, N.V. Human endogenous retrovirus K (HML-2) RNA and protein expression is a marker for human embryonic and induced pluripotent stem cells. *Retrovirology* **2013**, *10*, 115. [[CrossRef](#)] [[PubMed](#)]
35. Grow, E.J.; Flynn, R.A.; Chavez, S.L.; Bayless, N.L.; Mark, W.; Wesche, D.J.; Lance, M.; Ware, C.B.; Blish, C.A.; Chang, H.Y. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **2015**, *522*, 221–225. [[CrossRef](#)] [[PubMed](#)]
36. Pierre-Étienne, J.; Justin, J.; Guillaume, B. The majority of primate-specific regulatory sequences are derived from transposable elements. *PLoS Genet.* **2013**, *9*, e1003504.
37. Khazaei, E.; Farzaneh, N.; Mirshokraei, P.; Tabatabaeizadeh, S.-E.; Dehghani, H. Expression of endogenous retroviruses in pre-implantation stages of bovine embryo. *Reprod. Domest. Anim.* **2018**, *53*, 1405–1414. [[CrossRef](#)]
38. Burton, A.; Torrespadilla, M.E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 723–734. [[CrossRef](#)]
39. Bošković, A.; Eid, A.; Pontabry, J.; Ishiuchi, T.; Spiegelhalter, C.; Ram, E.V.; Meshorer, E.; Torres-Padilla, M.E. Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. *Genes Dev.* **2014**, *28*, 1042–1047.
40. Anifandis, G.; Messini, C.I.; Dafopoulos, K.; Messinis, I.E. Genes and Conditions Controlling Mammalian Pre- and Post-implantation Embryo Development. *Curr. Genom.* **2015**, *16*, 32–46. [[CrossRef](#)]
41. Adenot, P.G.; Mercier, Y.; Renard, J.P.; Thompson, E.M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **1997**, *124*, 4615–4625.
42. Zhang, B.; Zheng, H.; Huang, B.; Li, W.; Xiang, Y.; Peng, X.; Ming, J.; Wu, X.; Zhang, Y.; Xu, Q.; et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **2016**, *537*, 553. [[CrossRef](#)] [[PubMed](#)]
43. Schultz, R.M. Regulation of zygotic gene activation in the mouse. *Bioessays* **1993**, *15*, 531–538. [[CrossRef](#)] [[PubMed](#)]
44. Flach, G.; Johnson, M.H.; Braude, P.R.; Taylor, R.A.; Bolton, V.N. The transition from maternal to embryonic control in the 2-cell mouse embryo. *Embo J.* **1982**, *1*, 681–686. [[CrossRef](#)] [[PubMed](#)]
45. Latham, K.E.; Solter, D.; Schultz, R.M. Acquisition of a transcriptionally permissive state during the 1-cell stage of mouse embryogenesis. *Dev. Biol.* **1992**, *149*, 457–462. [[CrossRef](#)]
46. Ram, P.T.; Schultz, R.M. Reporter gene expression in G2 of the 1-cell mouse embryo. *Dev. Biol.* **1993**, *156*, 552–556. [[CrossRef](#)]
47. Matsumoto, K.; Anzai, M.; Nakagata, N.; Takahashi, A.; Takahashi, Y.; Miyata, K. Onset of paternal gene activation in early mouse embryos fertilized with transgenic mouse sperm. *Mol. Reprod. Dev.* **1994**, *39*, 136–140. [[CrossRef](#)]
48. Bouniol, C.; Nguyen, E.; Debey, P. Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Exp. Cell Res.* **1995**, *218*, 57–62. [[CrossRef](#)]
49. Schoorlemmer, J.; Pérez-Palacios, R.; Climent, M.; Guallar, D.; Muniesa, P. Regulation of Mouse Retroelement MuERV-L/MERVL Expression by REX1 and Epigenetic Control of Stem Cell Potency. *Front. Oncol.* **2014**, *4*, 14. [[CrossRef](#)]
50. Latham, K.E.; Garrels, J.I.; Chang, C.; Solter, D. Quantitative analysis of protein synthesis in mouse embryos. I. Extensive reprogramming at the one- and two-cell stages. *Development* **1991**, *112*, 921–932.
51. Huang, Y.; Kim, J.K.; Do, D.V.; Lee, C.; Penfold, C.A.; Zyllicz, J.J.; Marioni, J.C.; Hackett, J.A.; Surani, M.A. STELLA modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition. *eLife* **2017**, *6*, e22345. [[CrossRef](#)]
52. Evsikov, A.V.; de Vries, W.N.; Peaston, A.E.; Radford, E.E.; Fancher, K.S.; Chen, F.H.; Blake, J.A.; Bult, C.J.; Latham, K.E.; Solter, D. Systems biology of the 2-cell mouse embryo. *Cytogenet. Genome Res.* **2004**, *105*, 240–250. [[CrossRef](#)] [[PubMed](#)]

53. Moore, H.; Soygur, B. Expression of Syncytin 1 (HERV-W), in the preimplantation human blastocyst, embryonic stem cells and trophoblast cells derived in vitro. *Hum. Reprod.* **2016**, *31*, 1455–1461. [[CrossRef](#)]
54. Sakurai, T.; Nakagawa, S.; Bai, H.; Bai, R.; Kusama, K.; Ideta, A.; Aoyagi, Y.; Kaneko, K.; Iga, K.; Yasuda, J.; et al. Novel endogenous retrovirus-derived transcript expressed in the bovine placenta is regulated by WNT signaling. *Biochem. J.* **2017**, *474*, 3499–3512. [[CrossRef](#)] [[PubMed](#)]
55. Laverdure, S.; Polakowski, N.; Hoang, K.; Lemasson, I. Permissive Sense and Antisense Transcription from the 5' and 3' Long Terminal Repeats of Human T-Cell Leukemia Virus Type 1. *J. Virol.* **2016**, *90*, 3600–3610. [[CrossRef](#)]
56. Iaco, A.D.; Planet, E.; Coluccio, A.; Verp, S.; Duc, J.; Trono, D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **2017**, *49*, 941–945. [[CrossRef](#)]
57. Whiddon, J.L.; Langford, A.T.; Wong, C.J.; Zhong, J.W.; Tapscott, S.J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **2017**, *49*, 935–940. [[CrossRef](#)]
58. Rebollo, R.; Romanish, M.T.; Mager, D.L. Transposable elements: An abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* **2011**, *46*, 21–42. [[CrossRef](#)]
59. Macfarlan, T.S.; Gifford, W.D.; Saurabh, A.; Shawn, D.; Karen, L.; Wang, J.; Andrews, S.E.; Laura, F.; Rosenfeld, M.G.; Bing, R. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* **2011**, *25*, 594–607. [[CrossRef](#)]
60. Satou, Y.; Paola, M.; Ishihara, K.; Fukuda, A.; Nosaka, K.; Watanabe, T.; Rowan, A.; Nakao, M.; Bangham, C.R. HTLV-1 inserts an ectopic CTCF-binding site into the human genome. *Proc. Natl. Acad. Sci. USA* **2015**, *113*, 3054–3059. [[CrossRef](#)]
61. Ito, J.; Sugimoto, R.; Nakaoka, H.; Yamada, S.; Kimura, T.; Hayano, T.; Inoue, I. Systematic identification and characterization of regulatory elements derived from human endogenous retroviruses. *PLoS Genet.* **2017**, *13*, e1006883. [[CrossRef](#)]
62. Ke, Y.; Xu, Y.; Chen, X.; Feng, S.; Liu, Z.; Sun, Y.; Yao, X.; Li, F.; Zhu, W.; Gao, L. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell* **2017**, *170*, 367–381. [[CrossRef](#)] [[PubMed](#)]
63. Schmidt, D.; Schwalie, P.; Wilson, M.; Ballester, B.; Gonçalves, Â.; Kutter, C.; Brown, G.; Marshall, A.; Flicek, P.; Odom, D. Waves of Retrotransposon Expansion Remodel Genome Organization and CTCF Binding in Multiple Mammalian Lineages. *Cell* **2012**, *148*, 335–348. [[CrossRef](#)]
64. Boyer, L.A.; Ihn, L.T.; Cole, M.F.; Johnstone, S.E.; Levine, S.S.; Zucker, J.P.; Guenther, M.G.; Kumar, R.M.; Murray, H.L.; Jenner, R.G. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **2005**, *122*, 947–956. [[CrossRef](#)] [[PubMed](#)]
65. Chen, X.; Xu, H.; Yuan, P.; Fang, F.; Huss, M.; Vega, V.B.; Wong, E.; Orlov, Y.L.; Zhang, W.; Jiang, J.; et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **2008**, *133*, 1106–1117. [[CrossRef](#)] [[PubMed](#)]
66. Xu, N.; Papagiannakopoulos, T.; Pan, G.; Thomson, J.A.; Kosik, K.S. MicroRNA-145 Regulates OCT4, SOX2, and KLF4 and Represses Pluripotency in Human Embryonic Stem Cells. *Cell* **2009**, *137*, 647–658. [[CrossRef](#)]
67. Yvonne, T.; Jinqui, Z.; Thomson, A.M.; Bing, L.; Isidore, R. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **2008**, *455*, 1124–1128.
68. Loewer, S.; Cabili, M.N.; Guttman, M.; Loh, Y.H.; Thomas, K.; Park, I.H.; Garber, M.; Curran, M.; Onder, T.; Agarwal, S.; et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* **2010**, *42*, 1113–1117. [[CrossRef](#)]
69. Wang, Y.; Xu, Z.; Jiang, J.; Chen, X.; Kang, J.; Xiao, L.; Wu, M.; Xiong, J.; Guo, X.; Liu, H. Endogenous miRNA Sponge lincRNA-RoR Regulates Oct4, Nanog, and Sox2 in Human Embryonic Stem Cell Self-Renewal. *Dev. Cell* **2013**, *25*, 69–80. [[CrossRef](#)]
70. Durruthydurruthy, J.; Sebastiano, V.; Wossidlo, M.; Cepeda, D.; Cui, J.; Grow, E.J.; Davila, J.; Mall, M.; Wong, W.H.; Wysocka, J. The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. *Nat. Genet.* **2016**, *48*, 44–52. [[CrossRef](#)]
71. Myriam, H.; Wendy, D.; Wolf, R. Epigenetic dynamics of stem cells and cell lineage commitment: Digging Waddington's canal. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 526–537.
72. Tahiliani, M.; Koh, K.P.; Shen, Y.; Pastor, W.A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L.M.; Liu, D.R.; Aravind, L.; et al. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **2009**, *324*, 930–935. [[CrossRef](#)] [[PubMed](#)]

73. Khursheed, I.; Seung-Gi, J.; Pfeifer, G.P.; Szabó, P.E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3642–3647.
74. He, Y.F.; Li, B.Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **2011**, *333*, 1303–1307. [[CrossRef](#)]
75. Ito, S.; Shen, L.; Dai, Q.; Wu, S.C.; Collins, L.B.; Swenberg, J.A.; He, C.; Zhang, Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **2011**, *333*, 1300–1303. [[CrossRef](#)] [[PubMed](#)]
76. Leung, D.C.; Lorincz, M.C. Silencing of endogenous retroviruses: When and why do histone marks predominate? *Trends Biochem. Sci.* **2012**, *37*, 127–133. [[CrossRef](#)] [[PubMed](#)]
77. Wang, C.; Liu, X.; Gao, Y.; Yang, L.; Li, C.; Liu, W.; Chen, C.; Kou, X.; Zhao, Y.; Chen, J. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat. Cell Biol.* **2018**, *20*, 620–631. [[CrossRef](#)] [[PubMed](#)]
78. Mikkelsen, T.S.; Ku, M.; Jaffe, D.B.; Issac, B.; Lieberman, E.; Giannoukos, G.; Alvarez, P.; Brockman, W.; Kim, T.K.; Koche, R.P.; et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **2007**, *448*, 553–560. [[CrossRef](#)]
79. Schultz, D.C.; Kasirajan, A.; Dmitri, N.; Maul, G.G.; Rauscher, F.J. SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **2002**, *16*, 919–932. [[CrossRef](#)]
80. Matsui, T.; Leung, D.; Miyashita, H.; Maksakova, I.A.; Miyachi, H.; Kimura, H.; Tachibana, M.; Lorincz, M.C.; Shinkai, Y. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **2010**, *464*, 927–931. [[CrossRef](#)]
81. Rowe, H.M.; Johan, J.; Daniel, M.; Jacques, R.; Séverine, R.; Tugce, A.; Maillard, P.V.; Hillary, L.L.; Sonia, V.; Julien, M. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **2010**, *463*, 237–240. [[CrossRef](#)]
82. Wolf, D.; Cammas, F.; Losson, R.; Goff, S.P. Primer Binding Site-Dependent Restriction of Murine Leukemia Virus Requires HP1 Binding by TRIM28. *J. Virol.* **2008**, *82*, 4675–4679. [[CrossRef](#)] [[PubMed](#)]
83. Yang, P.; Wang, Y.; Macfarlan, T.S. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. *Trends Genet.* **2017**, *33*, 871–881. [[CrossRef](#)] [[PubMed](#)]
84. Turelli, P.; Castro-Diaz, N.; Marzetta, F.; Kapopoulou, A.; Raclot, C.; Duc, J.; Tieng, V.; Quenneville, S.; Trono, D. Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res.* **2014**, *24*, 1260–1270. [[CrossRef](#)] [[PubMed](#)]
85. Rowe, H.M.; Kapopoulou, A.; Corsinotti, A.; Fasching, L.; Macfarlan, T.S.; Tarabay, Y.; Viville, S.; Jakobsson, J.; Pfaff, S.L.; Trono, D. TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. *Genome Res.* **2013**, *23*, 452–461. [[CrossRef](#)] [[PubMed](#)]
86. Karimi, M.; Goyal, P.; Maksakova, I.; Bilenky, M.; Leung, D.; Tang, J.X.; Shinkai, Y.; Mager, D.; Jones, S.; Hirst, M. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **2011**, *8*, 676–687. [[CrossRef](#)] [[PubMed](#)]
87. Corsinotti, A.; Kapopoulou, A.; Gubelmann, C.; Imbeault, M.; de Sio, F.R.; Rowe, H.M.; Mouscaz, Y.; Deplancke, B.; Trono, D. Global and stage specific patterns of Krüppel-associated-box zinc finger protein gene expression in murine early embryonic cells. *PLoS ONE* **2013**, *8*, e56721. [[CrossRef](#)] [[PubMed](#)]
88. Emerson, R.O.; Thomas, J.H. Adaptive Evolution in Zinc Finger Transcription Factors. *PLoS Genet.* **2009**, *5*, e1000325. [[CrossRef](#)]
89. Liu, H.; Chang, L.H.; Sun, Y.; Lu, X.; Stubbs, L. Deep Vertebrate Roots for Mammalian Zinc Finger Transcription Factor Subfamilies. *Genome Biol. Evol.* **2014**, *6*, 510–525. [[CrossRef](#)]
90. Lukic, S.; Nicolas, J.-C.; Levine, A.J. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. *Cell Death Differ.* **2014**, *21*, 381–387. [[CrossRef](#)]
91. Thomas, J.H.; Schneider, S. Coevolution of retroelements and tandem zinc finger genes. *Genome Res.* **2011**, *21*, 1800–1812. [[CrossRef](#)]
92. Dawkins, R.; Krebs, J.R. Arms races between and within species. *Proc. R. Soc. Lond. B Biol. Sci.* **1979**, *205*, 489–511. [[PubMed](#)]
93. Valen, L.V. A New Evolutionary Law. *Evol. Theory* **1973**, *1*, 1–30.

94. Hamilton, A.T.; Huntley, S.; Tran-Gyamfi, M.; Baggott, D.M.; Gordon, L.; Stubbs, L. Evolutionary expansion and divergence in the ZNF91 subfamily of primate-specific zinc finger genes. *Genome Res.* **2006**, *16*, 584–594. [[CrossRef](#)] [[PubMed](#)]
95. Mak, J.; Kleiman, L. Primer tRNAs for reverse transcription. *J. Virol.* **1997**, *71*, 8087–8095. [[PubMed](#)]
96. Schorn, A.J.; Gutbrod, M.J.; Leblanc, C.; Martienssen, R. LTR-Retrotransposon Control by tRNA-Derived Small RNAs. *Cell* **2017**, *170*, 61–71. [[CrossRef](#)] [[PubMed](#)]
97. Yang, Q.; Lin, J.; Liu, M.; Li, R.; Tian, B.; Zhang, X.; Xu, B.; Liu, M.; Zhang, X.; Li, Y. Highly sensitive sequencing reveals dynamic modifications and activities of small RNAs in mouse oocytes and early embryos. *Sci. Adv.* **2016**, *2*, e1501482. [[CrossRef](#)] [[PubMed](#)]
98. Qi, C.; Wei, Y.; Duan, E. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat. Rev. Genet.* **2016**, *17*, 733–743.
99. Peng, H.; Shi, J.; Ying, Z.; He, Z.; Liao, S.; Wei, L.; Li, L.; Han, C.; Ning, L.; Cao, Y. A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Res.* **2012**, *22*, 1609–1612. [[CrossRef](#)]
100. Chen, Q.; Yan, M.; Cao, Z.; Li, X.; Zhang, Y.; Shi, J.; Feng, G.H.; Peng, H.; Zhang, X.; Zhang, Y. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* **2016**, *351*, 397–400. [[CrossRef](#)]
101. Sharma, U.; Conine, C.C.; Shea, J.M.; Boskovic, A.; Derr, A.G.; Bing, X.Y.; Belleannee, C.; Kucukural, A.; Serra, R.W.; Sun, F. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* **2016**, *351*, 391–396. [[CrossRef](#)]
102. Gehring, W.J.; Affolter, M.; Bürglin, T. HOMEODOMAIN PROTEINS. *Annu. Rev. Biochem.* **1994**, *63*, 487–526. [[CrossRef](#)] [[PubMed](#)]
103. Clapp, J.; Mitchell, L.M.; Bolland, D.J.; Fantes, J.; Corcoran, A.E.; Scotting, P.J.; Armour, J.A.; Hewitt, J.E. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* **2007**, *81*, 264–279. [[CrossRef](#)]
104. Wu, J.; Huang, B.; Chen, H.; Yin, Q.; Liu, Y.; Xiang, Y.; Zhang, B.; Liu, B.; Wang, Q.; Xia, W. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **2016**, *534*, 652–657. [[CrossRef](#)] [[PubMed](#)]
105. Ooga, M.; Fulka, H.; Hashimoto, S.; Suzuki, M.G.; Aoki, F. Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics* **2016**, *11*, 85–94. [[CrossRef](#)] [[PubMed](#)]
106. Leidenroth, A.; Clapp, J.; Mitchell, L.M.; Coneyworth, D.; Dearden, F.L.; Iannuzzi, L.; Hewitt, J.E. Evolution of gene macrosatellites in placental mammals. *Chromosoma* **2012**, *121*, 489–497. [[CrossRef](#)] [[PubMed](#)]
107. Choi, S.H.; Gearhart, M.D.; Cui, Z.; Bosnakovski, D.; Kim, M.; Schennum, N.; Kyba, M. DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res.* **2016**, *44*, 5161–5173. [[CrossRef](#)] [[PubMed](#)]
108. Young, J.M.; Whiddon, J.L.; Yao, Z.; Kasinathan, B.; Snider, L.; Geng, L.N.; Balog, J.; Tawil, R.; Maarel, S.M.V.D.; Tapscott, S.J. DUX4 Binding to Retroelements Creates Promoters That Are Active in FSHD Muscle and Testis. *PLoS Genet.* **2013**, *9*, e1003947. [[CrossRef](#)] [[PubMed](#)]
109. Fadloun, A.; Le, G.S.; Jost, B.; Ziegler-Birling, C.; Takahashi, H.; Gorab, E.; Carninci, P.; Torres-Padilla, M.E. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat. Struct. Mol. Boil.* **2013**, *20*, 332–338. [[CrossRef](#)] [[PubMed](#)]
110. Adjaye, J.; Daniels, R.; Monk, M. The construction of cDNA libraries from human single preimplantation embryos and their use in the study of gene expression during development. *J. Assist. Reprod. Genet.* **1998**, *15*, 344–348. [[CrossRef](#)] [[PubMed](#)]
111. Percharde, M.; Lin, C.J.; Yin, Y.; Guan, J.; Peixoto, G.A.; Bulut-Karslioglu, A.; Biechele, S.; Huang, B.; Shen, X.; Ramalho-Santos, M. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell* **2018**, *174*, 391–405. [[CrossRef](#)]
112. Wilmut, I.; Schnieke, A.E.; McWhir, J.; Kind, A.J.; Campbell, K.H.S. Viable offspring derived from fetal and adult mammalian cells. *Nature* **1997**, *385*, 810–813. [[CrossRef](#)]
113. Wakayama, T.; Perry, A.C.F.; Zuccotti, M.; Johnson, K.R.; Yanagimachi, R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **1998**, *394*, 369–374. [[CrossRef](#)] [[PubMed](#)]

114. Rodriguez-Osorio, N.; Urrego, R.; Cibelli, J.B.; Eilertsen, K.; Memili, E. Reprogramming mammalian somatic cells. *Theriogenology* **2012**, *78*, 1869–1886. [[CrossRef](#)] [[PubMed](#)]
115. Atsuo, O.; Kimiko, I.; Teruhiko, W. Recent advancements in cloning by somatic cell nuclear transfer. *Philos. Trans. R. Soc. Lond.* **2013**, *368*, 20110329.
116. Liu, Y.; Wu, F.; Zhang, L.; Wu, X.; Li, D.; Xin, J.; Xie, J.; Kong, F.; Wang, W.; Wu, Q.; et al. Transcriptional defects and reprogramming barriers in somatic cell nuclear reprogramming as revealed by single-embryo RNA sequencing. *BMC Genom.* **2018**, *19*, 734. [[CrossRef](#)] [[PubMed](#)]
117. Liu, W.; Liu, X.; Wang, C.; Gao, Y.; Gao, R.; Kou, X.; Zhao, Y.; Li, J.; Wu, Y.; Xiu, W. Identification of key factors conquering developmental arrest of somatic cell cloned embryos by combining embryo biopsy and single-cell sequencing. *Cell Discov.* **2016**, *2*, 16010. [[CrossRef](#)] [[PubMed](#)]
118. Li, S.; Chen, X.; Fang, Z.; Shi, J.; Sheng, H.Z. Rabbits generated from fibroblasts through nuclear transfer. *Reproduction* **2006**, *131*, 1085–1090. [[CrossRef](#)]
119. Zhao, J.; Ross, J.W.; Hao, Y.; Spate, L.D.; Walters, E.M.; Samuel, M.S.; Rieke, A.; Murphy, C.N.; Prather, R.S. Significant improvement in cloning efficiency of an inbred miniature pig by histone deacetylase inhibitor treatment after somatic cell nuclear transfer. *Boil. Reprod.* **2009**, *81*, 525–530. [[CrossRef](#)]
120. Noggle, S.; Fung, H.L.; Gore, A.; Martinez, H.; Satriani, K.C.; Prosser, R.; Oum, K.; Paull, D.; Druckenmiller, S.; Freeby, M. Human oocytes reprogram somatic cells to a pluripotent state. *World Neurosurg.* **2012**, *77*, 9–11. [[CrossRef](#)]
121. Telford, N.A.; Watson, A.J.; Schultz, G.A. Transition from maternal to embryonic control in early mammalian development: A comparison of several species. *Mol. Reprod. Dev.* **1990**, *26*, 90–100. [[CrossRef](#)]
122. Wei, S.; Zakhartchenko, V.; Wolf, E. Epigenetic reprogramming in mammalian nuclear transfer. *Differentiation* **2003**, *71*, 91–113.
123. Inoue, K.; Oikawa, M.; Kamimura, S.; Ogonuki, N.; Nakamura, T.; Nakano, T.; Abe, K.; Ogura, A. Trichostatin A specifically improves the aberrant expression of transcription factor genes in embryos produced by somatic cell nuclear transfer. *Sci. Rep.* **2015**, *5*, 10127. [[CrossRef](#)] [[PubMed](#)]
124. Eckersley-Maslin, M.; Svensson, V.; Krueger, C.; Stubbs, T.M.; Giehr, P.; Krueger, F.; Miragaia, R.; Kyriakopoulos, C.; Berrens, R.; Milagre, I. MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs. *Cell Rep.* **2016**, *17*, 179–192. [[CrossRef](#)] [[PubMed](#)]
125. Iturbide, A.; Torres-Padilla, M.E. Starting embryonic transcription for the first time. *Nat. Genet.* **2017**, *49*, 820–821. [[CrossRef](#)] [[PubMed](#)]
126. David, R.; Sophie, L.V.; Francis, H.; Nathalie, D.P.; Marie, D.; Odile, H.; Gérard, P.; Bernard, M.; Thierry, H. Murine endogenous retrovirus MuERV-L is the progenitor of the “orphan” epsilon viruslike particles of the early mouse embryo. *J. Virol.* **2008**, *82*, 1622–1625.
127. Svoboda, P.; Stein, P.; Anger, M.; Bernstein, E.; Hannon, G.J.; Schultz, R.M. RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. *Dev. Biol.* **2004**, *269*, 276–285. [[CrossRef](#)]
128. Suzuki, T.; Minami, N.; Kono, T.; Imai, D.H. Zygotically Activated Genes Are Suppressed in Mouse Nuclear Transferred Embryos. *Cloning Stem Cells* **2006**, *8*, 295–304. [[CrossRef](#)]
129. Yang, L.; Song, L.; Liu, X.; Bai, L.; Li, G. KDM6A and KDM6B play contrasting roles in nuclear transfer embryos revealed by MERVL reporter system. *Embo Rep.* **2018**, *19*, e46240. [[CrossRef](#)]
130. Matoba, S.; Liu, Y.; Lu, F.; Iwabuchi, K.; Shen, L.; Inoue, A.; Zhang, Y. Embryonic Development following Somatic Cell Nuclear Transfer Impeded by Persisting Histone Methylation. *Cell* **2014**, *159*, 884–895. [[CrossRef](#)]

