Minireview

Impact of Retrotransposons in Pluripotent Stem Cells

Yoshiaki Tanaka, Leeyup Chung[1], and In-Hyun Park*

Retrotransposons, which constitute approximately 40% of the human genome, have the capacity to 'jump' across the genome. Their mobility contributes to oncogenesis, evolution, and genomic plasticity of the host genome. Induced pluripotent stem cells as well as embryonic stem cells are more susceptible than differentiated cells to genomic aberrations including insertion, deletion and duplication. Recent studies have revealed specific behaviors of retrotransposons in pluripotent cells. Here, we review recent progress in understanding retrotransposons and provide a perspective on the relationship between retrotransposons and genomic variation in pluripotent stem cells.

INTRODUCTION

Embryonic stem cells (ESCs) are derived from the inner cell mass of the mammalian blastocyst and have the capacity to self-renew and differentiate into all types of cells and tissues (Thomson et al., 1998). The pluripotency of ESCs makes them a promising system for *in vitro* studies of cell biology, screening drugs for efficacy and toxicity, and regenerative medicine (Jung et al., 2012). Recent studies succeeded in reprogramming human and mouse somatic cells into ESC-like cells by ectopic expression of defined transcription factors (Oct4, Sox2, Klf4, and c-Myc) (Park et al., 2008; Takahashi and Yamanaka, 2006). Such induced pluripotent stem cells (iPSCs) avoid the ethical issues involved in deriving ESCs and provide autologous sources for cell therapy.

Despite their tremendous potential in cell-based therapies, several challenges remain to be overcome before pluripotent cells can be used in the clinic. Long-term culture of ESCs and iPSCs induces karyotypic changes (Draper et al., 2004; Lefort et al., 2008; Spits et al., 2008) and copy number variations (CNVs) (Hussein et al., 2011; Liang et al., 2008; Närvä et al., 2010). Overexpression of Klf4 and c-Myc represses the p53/p21 signaling pathway, resulting in increased DNA damage (Deng and Xu, 2009; Hong et al., 2009; Kawamura et al., 2009). In addition, ESCs and iPSCs contain fewer genomic domains enriched with repressive histone modifications, such as H3K27me3 and H3K9me3, compared to somatic cells such as fibroblasts

(Hawkins et al., 2010). DNA methylation status in regions outside promoters also changes during reprogramming into iPSCs, to a greater extent than it changes within promoters (Lister et al., 2011; Meissner et al., 2008). This genomic and epigenetic instability increases the risk of diseases of the genome, such as tumors, and makes these cells less than ideally suitable for fundamental research and stem cell-based transplantation therapy.

Transposable elements (TEs) are mobile DNA sequences that can move from one chromosomal site to another. TEs are a major source of CNVs (Fig. 1) (Huang et al., 2010). Retrotransposons, the most abundant TEs, can be classified according to the presence or absence of a long terminal repeat (LTR). Non-LTR retrotransposons are further separated into Short and Long INterspersed Elements (SINEs and LINEs) (Finnegan, 1997). SINEs and LINEs are transcribed by RNA polymerases (Pol) III and II, respectively, and the resulting transcripts are converted into DNA by reverse transcriptase (RT). These DNA copies are subsequently integrated into genomic DNA. Since RT is encoded by LINEs but not SINEs, reverse transcription of SINEs must be accomplished by RT encoded by other elements. LTR retrotransposons are also transcribed by RNA polymerase II, using a promoter within the LTR itself (Havecker et al., 2004). Following transcription, retrotransposon RNA is translated into proteins that form the virus-like particle (VLR), which encapsulates the RNA within the cytoplasm. Reverse transcription proceeds within the VLR, and the resulting retrotransposon cDNA is integrated into the host genome.

Most retrotransposons in the human genome have been mutated and transcriptionally silenced. In addition, many retrotransposons are attenuated by epigenetic modifications such as DNA methylation (Kuramochi-Miyagawa et al., 2008), enrichment of repressive histone modifications (Kondo and Issa, 2003; Martens et al., 2005; Mikkelsen et al., 2007), depletion of active histone modifications (Huda et al., 2010), and strong nucleosome positioning (Englander and Howard, 1995, 1996; Tanaka et al., 2010). Thus, retrotransposons have been historically disregarded as "junk DNA." However, despite the presence of repressive machinery, recent studies have demonstrated that younger retrotransposons can be activated in specific cell types, including neuronal cells (Baillie et al., 2011; Coufal et

Springer

al., 2009), tumors (Balaj et al., 2011), germ cells (Watanabe et al., 2006), and undifferentiated cells (Macia et al., 2011). In particular, for a retrotransposition event to be passed on to the next generation, the retrotransposition must occur in a germ cell or a stem cell that subsequently differentiates into the germ cell lineage.

Several recent studies have described the specific activities and mechanisms of retrotransposons in undifferentiated cells. Here, we address the virulence of retrotransposons in the human genome, review recent advances in understanding the behaviors of retrotransposons in undifferentiated cells, and discuss their potential roles in pluripotent cell types (Fig. 2).

IMPACT OF ALU ELEMENTS

The Alu retrotransposons, a family of primate-specific SINEs, are the most abundant repetitive elements (>1 million copies) in the human genome (Grover et al., 2004). A typical Alu element is approximately 300 bp long, and comprises two distinct GC-rich regions linked with an A-rich region. According to their evolutionary ages, Alu elements are classified into subfamilies: old (AluJ), middle (AluS), and young (AluY) (Batzer and Deininger, 2002). Whereas all AluJ and many AluS elements are currently inactive, a portion of AluS and AluY elements (approximately 6,000 copies) are still active (Bennett et al., 2008). Alu elements do not encode any proteins, but contain two internal promoter elements (A and B box) in the left monomer. These promoters are not detectable in older Alu elements (Alemán et al., 2000). The activity of the A box may be repressed in Alu elements that harbor nucleosomes in specific positions (Englander and Howard, 1995, 1996; Tanaka et al., 2010). In addition, younger Alu elements contain more CpG dinucleotides, but these are hypermethylated in somatic cells (Xie et al., 2009).

Alu insertion in the vicinity of a gene is likely to influence post-transcriptional regulation of the neighboring gene, for example by influencing alternative splicing (Häsler et al., 2007). Alu-mediated changes in regulatory mechanisms confer both advantages and disadvantages in humans (Lev-Maor et al., 2008). As an example of a positive effect, the insertion of an Alu element within the fifth intron of the plakophilin 2 (*PKP2*) gene introduced a new splice site, yielding a novel protein isoform in the hominid lineage (Krull et al., 2005). On the other hand, approximately 0.1% of genetic disorders are related to Alu insertion (Deininger and Batzer, 1999). Retinitis pigmentosa (RP) is an inherited eye disease characterized by photoreceptor loss, decline of visual field, and night blindness. Approximately 1.2% of RP patients have an Alu insertion in the ninth exon of male germ cell-associated kinase (MAK) (Tucker et al., 2011). Tucker *et al.* recently developed iPSCs from RP patients and monitored MAK gene and protein expression during retinal cell differentiation. They showed that the Alu insertion causes a splicing defect, resulting in failure to produce mature MAK protein in retinal precursor cells. A recent study demonstrated that Alu RNA levels are increased in human ESCs (hESCs) (Macia et al., 2011). Therefore, it is possible that pathogenic Alu insertions such as the one responsible for RP occurred in undifferentiated cells or during differentiation of the germ cell lineage.

Adenosine-to-inosine (A-to-I) RNA editing is a form of post-transcriptonal regulation that increases transcriptome and proteome diversity. The majority of A-to-I editing occurs in Alu elements, within non-coding RNAs, and in the introns and untranslated regions (UTRs) of coding RNAs (Athanasiadis et al., 2004; Peng et al., 2012). Editing in coding regions is rare, but is significantly increased during brain development, suggesting

that it may contribute to functional changes in neuronal proteins and the development of brain complexity (Wahlstedt et al., 2009). The Adenosine Deaminase Acting on RNA (ADAR) proteins are enzymes that catalyze the A-to-I reaction in double-stranded regions of RNA regions. Three genes encoding ADAR proteins (*ADAR1*, *ADAR2,* and *ADAR3*) have been identified in the human genome (Hogg et al., 2011). In mouse, *ADAR1* is required in embryogenesis, and its loss induces widespread apoptosis (Wang et al., 2000; 2004). The level of A-to-I RNA editing in Alu elements within non-coding regions is higher in hESCs, but decreases during neuronal differentiation whereas the level of editing in coding regions is low in both undifferentiated and differentiated cells (Osenberg et al., 2010). Furthermore, the *ADAR1* mRNA level in hESCs is dramatically higher than those of *ADAR2* and *ADAR3*. Knockdown of *ADAR1* decreases the level of A-to-I editing in Alu RNAs and increases expression of genes related to cell development and differentiation.

The A-to-I editing of inverted Alu pairs in 3' UTRs promotes mRNA silencing by nuclear retention (Chen et al., 2008). LIN28, one of the reprogramming factors that have been used to derive iPSCs, contains inverted Alu elements in its 3' UTR (Yu et al., 2007). Despite the high level of Alu A-to-I editing in hESCs *via* ADAR1, LIN28 protein is highly expressed in these cells. Chen et al. recently demonstrated that expression of NEAT non-coding RNA is essential in assembly of paraspeckles, which enhances nuclear retention of edited mRNAs (Chen and Carmichael, 2009). They found that hESCs do not express NEAT and do not form paraspeckles. Thus, LIN28 is not retained in the nuclei even with the A-to-I editing. Taken together, these findings suggest that Alu elements play an important role in human embryonic development.

IMPACT OF L1 ELEMENTS

L1 is the major LINE in mammals, and occupies approximately 17% (approximately 500,000 copies) of the human genome (Lander et al., 2001). Although there are three subfamilies of LINEs in the human genome, only L1 is capable of autonomous amplification. Full-length L1 is approximately 6 kb long and encodes two proteins, ORF1p and ORF2p. ORF1p forms a ribonucleoprotein particle by binding single-stranded L1 DNA and RNA, thereby serving as a nucleic acid chaperone (Kolosha and Martin, 1997; Martin and Bushman, 2001). ORF2p is a multifunctional protein with endonuclease and RT activity (Feng et al., 1996; Mathias et al., 1991). Both proteins are required for L1 retrotransposition (Martin et al., 2005). ORF1p is not crucial for retrotransposition of Alu elements, but can enhance Alu mobilization (Wallace et al., 2008). In addition to ORFs, L1 has sense and antisense RNA Pol II promoters in the 5′ UTR and polyadenlyation signals in the 3′ UTR. Therefore, L1 insertion can confer promoter and polyadenylation activity on neighboring genes (Mätlik et al., 2006; Nigumann et al., 2002; Roy-Engel et al., 2005; Speek, 2001).

There are approximately 4,000 full-length L1 elements in the human genome. As with Alu retrotransposition, L1 retrotransposition contributes to both genomic evolution and pathogenesis (Chen et al., 2006; Cordaux and Batzer, 2009). In some haemophilia A patients, L1 insertions are present in exon 14 of the gene encoding Factor VIII, which is required for blood coagulation (Kazazian et al., 1988). In a patient with Duchenne muscular dystrophy, a *de novo* insertion of L1 into exon 44 of the gene encoding dystrophin results in exon skipping (Musova et al., 2006). More than 90% of disease-causing insertions are
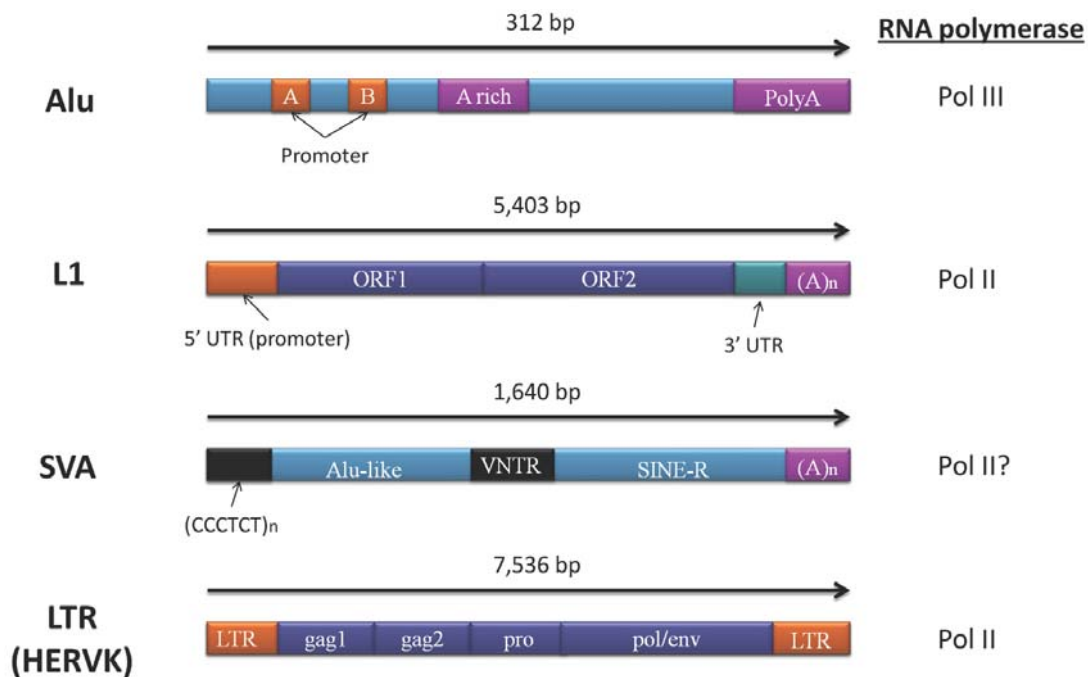
Fig. 1. Schematics of the major classes of retrotransposons. Typical sequence length of each retrotransposon is according to the Repbase database (http://www.girinst.org/repbase/). Alu contains two RNA polymerase III promoters, but does not encode proteins. L1 has a polymerase II promoter in the 5′ UTR region and two ORFs. SVA consists of hexamer repeat regions, Alu-like regions, variable number of tandem repeats (VNTR), SINE-R, and polyA regions, but does not encode proteins. Because there is no internal promoter, SVA may be transcribed using 5′-flanking promoters. LTR retrotransposon contains a promoter in the LTR, and the transcript produces group specific antigen (gag), protease (pro), reverse transcriptase (pol), and envelope (env) proteins.

due to Alu and L1 elements (Hancks and Kazazian, 2012). Because L1 mediates the retrotransposition of other elements, L1 expression is related to almost all disease-causing insertions in the human genome.

L1 RNAs are highly transcribed in germ cells (Kano et al., 2009), hESCs (Garcia-Perez et al., 2007; Macia et al., 2011), and neuronal cells (Coufal et al., 2009). In hESCs, highly expressed L1 elements are located within intragenic regions (Macia et al., 2011). Despite high L1 expression in these cell types, the genomic integration of L1 is not detectable in germ cells (Kano et al., 2009). Rather, L1 RNAs are carried over from the germ cells into the embryo, and integration occurs during embryogenesis. These results suggest that most L1 retrotranspositions occur in cells of the early embryo, including ESCs, and lead to mosaicism in somatic and germ-line tissues (van den Hurk et al., 2007).

L1 elements also play a key role in dosage compensation. In order to compensate for the double dose of X-chromosome genes in females relative to males, one copy of the X chromosome in female somatic cells must be inactivated. During differentiation of ESCs, X-chromosome inactivation (XCI) is initiated by coating the X chromosome with the non-coding XIST RNA; L1 elements are thought to provide binding sites for XIST (Bailey et al., 2000; Lyon, 2000). L1 elements are two-fold more abundant on the X chromosome than on autosomes; conversely, regions that escape XCI are relatively depleted of L1 elements (Lyon, 1998). The majority of L1s are involved in forming the heterochromatic nuclear compartment of the inactive X chromosome. A subset of young full-length L1s are active

outside the XIST-mediated compartment during early stages of differentiation, but they are expressed within the compartment during later stages. Although it is clear that L1 elements are associated with XCI during ESC differentiation, the mechanism underlying expression of young L1 elements during ESC differentiation remains unknown. According to one current model, L1 transcription assists in the spreading of the silent compartment along the inactive X (Chow et al., 2010).

Previously, we and other groups demonstrated that the inactive X chromosome may be reactivated during reprogramming of fibroblast cells derived from Rett syndrome (RTT) patients into iPSCs (RTT-iPSCs) (Cheung et al., 2011; Kim et al., 2011; Marchetto et al., 2010; Pomp et al., 2011). RTT is a neurodevelopmental disorder, arising mostly in girls. RTT patients carry mutations in the gene encoding Methyl-CpG binding protein 2 (MeCP2) (Amir et al., 1999). RTT-iPSCs exhibit increased susceptibility to retrotransposition of L1 elements (Muotri et al., 2010). By contrast, Alu transcription is not regulated by MeCP2 (Yu et al., 2001).

TET1 protein was recently identified as an enzyme that converts methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), a process that is upregulated during the generation of iPSCs (Ruzov et al., 2011; Tahiliani et al., 2009). TET1 is highly expressed in ESCs and is essential for maintenance of pluripotency (Ito et al., 2010). The level of 5hmC is elevated in L1 promoters in ESCs, but decreases after differentiation into embryonic bodies (EBs) (Ficz et al., 2011). By contrast, the level of 5mC in L1 promoter increases during differentiation of ESCs into EBs. 5hmC is enriched in binding sites of pluripotency fac-
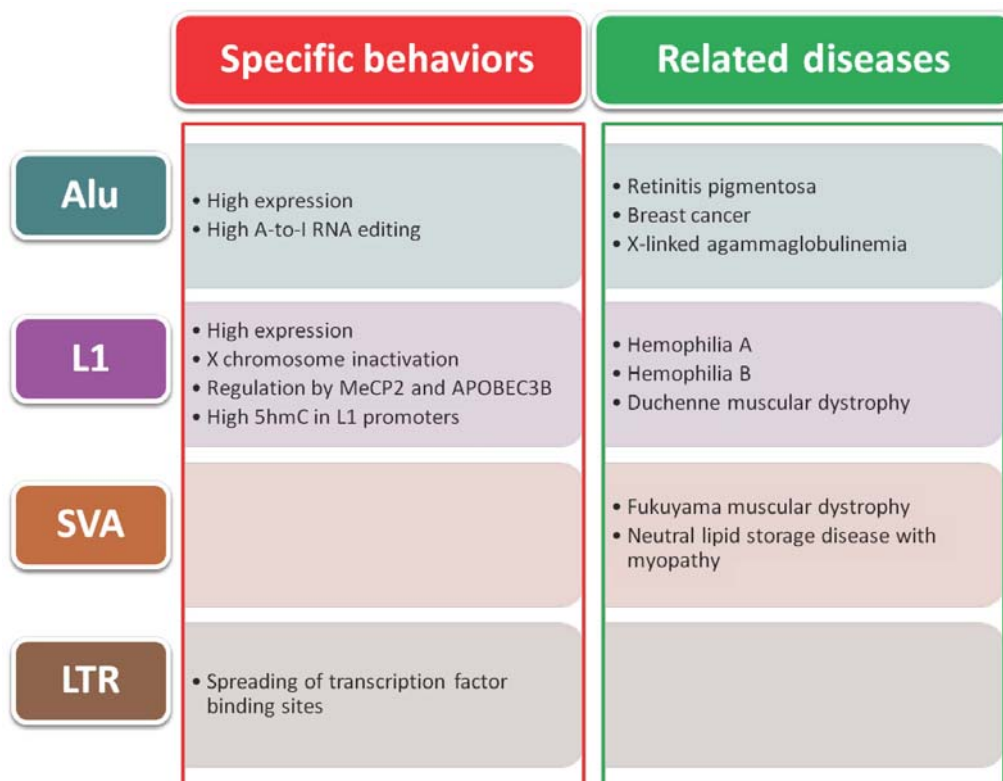
Fig. 2. Summary of ESCs/iPSCs-specific behaviors of retrotransposons. The first column represents the contributions of retrotransposons to undifferentiated cells. The second column gives examples of diseases related to each retrotransposon.

tors (Oct4, Nanog, and Sox2), enhancers marked by H3K4me1 and H3K27ac, promoters, and gene coding regions in human (Stroud et al., 2011; Szulwach et al., 2011) and mouse ESCs (Pastor et al., 2011; Wu et al., 2011a; Xu et al., 2011). Because the reaction catalyzed by TET1 involves demethylation of methylated cytosine that marks inactive genes, 5hmC is presumed to be distributed around active genes. However, recent studies in ESCs have revealed that 5hmC is also enriched in developmental genes with bivalent histone modifications (H3K4me3 and H3K27me3) (Wu et al., 2011a; 2011b). Furthermore, chromatin immunoprecipitation with massively parallel sequencing (ChIP-seq) analysis has revealed that TET1 is co-localized with the SIN3A co-repressor (Williams et al., 2011). Consequently, TET1 (as well as 5hmC) may function as both activator and repressor. It is still unclear whether 5hmC in L1 promoters activates L1 transcription. However, because SINE and LINE elements are significantly depleted around PcG-bound promoters in hESCs (Estécio et al., 2010), one could infer that 5hmC may activate L1 transcription in undifferentiated cells.

In addition to mechanisms for activating L1 elements, mechanisms also exist to repress L1 in ESCs/iPSCs. APOBEC (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide) is a cytidine deaminase that removes 5hmC by converting it into 5hmU (Branco et al., 2012). Members of the human APOBEC3 family inhibit L1 retrotransposons in HeLa cells (Kinomoto et al., 2007). Despite high expression of almost all members of the APOBEC3 family in hESCs, only APOBEC3B restricts L1 retrotransposition in those cells (Wissing et al., 2011). APOBEC3B mRNA is not expressed in somatic tissues

(Bogerd et al., 2006). Thus, APOBEC3B-mediated cytidine deamination represents a mechanism for regulating L1 retrotransposition that is unique to undifferentiated cells.

IMPACT OF SVA ELEMENTS

SVA is a newly defined family comprising three other types of repetitive elements (SINE-R, VNTR, and Alu) (Ostertag et al., 2003). SVA elements evolved recently in the hominid lineage, and many of them are currently active in the human genome. Their mobilization is mediated in *trans* by L1 elements (Hancks et al., 2011). Several SVA insertions disrupted gene structures, resulting in genetic diseases (Hancks and Kazazian, 2010). In Fukuyama muscular dystrophy, SVA insertion into the 3′ UTR of the gene encoding fukutin provokes an abnormal splicing event that changes the position of the stop codon (Taniguchi-Ikeda et al., 2011). Although the activity of SVA in undifferentiated cells has not been elucidated, the evidence regarding L1 activity in ESCs/iPSCs suggests that increased retrotransposition of SVA elements may occur during pluripotent stages.

IMPACT OF LTR RETROTRANSPOSONS

In contrast to other retrotransposons, LTR retrotransposons are thought to be static in the human genome at present. However, recent studies have suggested that LTR retrotransposons have contributed to the spread of binding sites for pluripotency factors during evolution (Bourque et al., 2008). Kunarso et al. (2010) performed ChIP-seq for CTCF, OCT4, and NANOG in

hESCs and found that more than 2,000 and 4,000 OCT4 and NANOG binding sites, respectively, are within endogenous retrovirus 1. Most of these repeat-associated binding sites were not present in homologous regions of the mouse genome. Despite shared expression of numerous genes in human and mouse ESCs/iPSCs (Chin et al., 2009; Hanna et al., 2010), the transcriptional and epigenetic profiles of the human cells exhibit several distinct patterns (Bernstein et al., 2005; Ginis et al., 2004; Richards et al., 2004). Genes associated with leukemia inhibitory factor signaling pathways, cell cycle and death pathways, and cytokines are highly expressed only in mouse (Ginis et al., 2004; Richards et al., 2004). In addition, despite similar patterns of histone H3 lysine 4 methylation (H3K4me) in orthologous regions, many of the H3K4me-enriched regions are not conserved (Bernstein et al., 2005). It may be speculated that the human-specific pluripotency regulatory networks have arisen as a result of retrotransposition events within transcription factor binding sites over the course of human evolution.

CONCLUSION

Several lines of new evidence have revealed that retrotransposon expression is increased, and that retrotransposon-mediated gene regulation occurs, in undifferentiated cells. Because retrotransposition also promotes genomic evolution (Cordaux and Batzer, 2009), retrotransposition is a double-edged sword with respect to the host genome. Whereas integration events in germ cells can be passed down to future generations, only a fraction of the events in ESCs are heritable. Piwi-interacting RNAs, which are capable of retrotransposon silencing, are highly expressed in germ cells but diminished in ESCs (Ohnishi et al., 2010). It is possible that the early embryonic stage is more tolerant to retrotransposition than germ cell lineage to increase genomic diversity.

Techniques for generating iPSCs are improving rapidly, and the current methods exhibit a significantly lower frequency of CNVs (Quinlan et al., 2011). Understanding the mechanisms of retrotransposon activation and repression in undifferentiated cells should enable us to more effectively develop non-pathogenic iPSCs for use in research and cell-based therapies.

REFERENCES

Alemán, C., Roy-Engel, A.M., Shaikh, T.H., and Deininger, P.L. (2000). Cis-acting influences on Alu RNA levels. Nucleic Acids Res. *28*, 4755-4761.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat. Genet. *23*, 185-188.

Athanasiadis, A., Rich, A., and Maas, S. (2004). Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol. *2*, e391.

Bailey, J.A., Carrel, L., Chakravarti, A., and Eichler, E.E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc. Natl. Acad. Sci. USA *97*, 6634-6639.

Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P.M., Rizzu, P., Smith, S., Fell, M., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. Nature *479*, 534-537.

Balaj, L., Lessard, R., Dai, L., Cho, Y.J., Pomeroy, S.L., Breakefield, X.O., and Skog, J. (2011). Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. Nat. Commun. *2*, 180.

Batzer, M.A., and Deininger, P.L. (2002). Alu repeats and human genomic diversity. Nat. Rev. Genet. *3*, 370-379.

Bennett, E.A., Keller, H., Mills, R.E., Schmidt, S., Moran, J.V., Weichenrieder, O., and Devine, S.E. (2008). Active Alu retrotransposons in the human genome. Genome Res. *18*, 1875-1883.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., Gingeras, T.R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V., and Cullen, B.R. (2006). Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proc. Natl. Acad. Sci. USA *103*, 8780-8785.

Bourque, G., Leong, B., Vega, V.B., Chen, X., Lee, Y.L., Srinivasan, K.G., Chew, J.L., Ruan, Y., Wei, C.L., Ng, H.H., et al. (2008). Evolution of the mammalian transcription factor binding repertoire *via* transposable elements. Genome Res. *18*, 1752-1762.

Branco, M.R., Ficz, G., and Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat. Rev. Genet. *13*, 7-13.

Chen, L.L., and Carmichael, G.G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol. Cell *35*, 467-478.

Chen, J.M., Férec, C., and Cooper, D.N. (2006). LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease: mutation detection bias and multiple mechanisms of target gene disruption. J. Biomed. Biotechnol. *2006*, 56182.

Chen, L.L., DeCerbo, J.N., and Carmichael, G.G. (2008). Alu element-mediated gene silencing. EMBO J. *27*, 1694-1705.

Cheung, A.Y., Horvath, L.M., Grafodatskaya, D., Pasceri, P., Weksberg, R., Hotta, A., Carrel, L., and Ellis, J. (2011). Isolation of MECP2-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. Hum. Mol. Genet. *20*, 2103-2115.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell *5*, 111-123.

Chow, J.C., Ciaudo, C., Fazzari, M.J., Mise, N., Servant, N., Glass, J.L., Attreed, M., Avner, P., Wutz, A., Barillot, E., et al. (2010). LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell *141*, 956-969.

Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. *10*, 691-703.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127-1131.

Deininger, P.L., and Batzer, M.A. (1999). Alu repeats and human disease. Mol. Genet. Metab. *67*, 183-193.

Deng, W., and Xu, Y. (2009). Genome integrity: linking pluripotency and tumorgenicity. Trends Genet. *25*, 425-427.

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat. Biotechnol. *22*, 53-54.

Englander, E.W., and Howard, B.H. (1995). Nucleosome positioning by human Alu elements in chromatin. J. Biol. Chem. *270*, 10091-10096.

Englander, E.W., and Howard, B.H. (1996). A naturally occurring T14A11 tract blocks nucleosome formation over the human neurofibromatosis type 1 (NF1)-Alu element. J. Biol. Chem. *271*, 5819-5823.

Estécio, M.R., Gallegos, J., Vallot, C., Castoro, R.J., Chung, W., Maegawa, S., Oki, Y., Kondo, Y., Jelinek, J., Shen, L., et al. (2010). Genome architecture marked by retrotransposons modulates predisposition to DNA methylation in cancer. Genome Res. *20*, 1369-1382.

Feng, Q., Moran, J.V., Kazazian, H.H., and Boeke, J.D. (1996). Human L1 retrotransposon encodes a conserved endonuclease

required for retrotransposition. Cell 87, 905-916.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473, 398-402.

Finnegan, D.J. (1997). Transposable elements: how non-LTR retrotransposons do it. Curr. Biol. 7, R245-248.

Garcia-Perez, J.L., Marchetto, M.C., Muotri, A.R., Coufal, N.G., Gage, F.H., O'Shea, K.S., and Moran, J.V. (2007). LINE-1 retrotransposition in human embryonic stem cells. Hum. Mol. Genet. 16, 1569-1577.

Ginis, I., Luo, Y., Miura, T., Thies, S., Brandenberger, R., Gerecht-Nir, S., Amit, M., Hoke, A., Carpenter, M.K., Itskovitz-Eldor, J., et al. (2004). Differences between human and mouse embryonic stem cells. Dev. Biol. 269, 360-380.

Grover, D., Mukerji, M., Bhatnagar, P., Kannan, K., and Brahmachari, S.K. (2004). Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. Bioinformatics 20, 813-817.

Hancks, D.C., and Kazazian, H.H. (2010). SVA retrotransposons: Evolution and genetic instability. Semin. Cancer Biol. 20, 234-245.

Hancks, D.C., and Kazazian, H.H. (2012). Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev. 22, 191-203.

Hancks, D.C., Goodier, J.L., Mandal, P.K., Cheung, L.E., and Kazazian, H.H. (2011). Retrotransposition of marked SVA elements by human L1s in cultured cells. Hum. Mol. Genet. 20, 3386-3400.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. USA 107, 9222-9227.

Häsler, J., Samuelsson, T., and Strub, K. (2007). Useful 'junk': Alu RNAs in the human transcriptome. Cell Mol. Life Sci. 64, 1793-1800.

Havecker, E.R., Gao, X., and Voytas, D.F. (2004). The diversity of LTR retrotransposons. Genome Biol. 5, 225.

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., et al. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell 6, 479-491.

Hogg, M., Paro, S., Keegan, L.P., and O'Connell, M.A. (2011). RNA editing by mammalian ADARs. Adv. Genet. 73, 87-120.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature 460, 1132-1135.

Huang, C.R., Schneider, A.M., Lu, Y., Niranjan, T., Shen, P., Robinson, M.A., Steranka, J.P., Valle, D., Civin, C.I., Wang, T., et al. (2010). Mobile interspersed repeats are major structural variants in the human genome. Cell 141, 1171-1182.

Huda, A., Mariño-Ramírez, L., and Jordan, I.K. (2010). Epigenetic histone modifications of human transposable elements: genome defense versus exaptation. Mob. DNA 1, 2.

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Närvä, E., Ng, S., Sourour, M., Hämäläinen, R., Olsson, C., et al. (2011). Copy number variation and selection during reprogramming to pluripotency. Nature 471, 58-62.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129-1133.

Jung, Y.W., Hysolli, E., Kim, K.Y., Tanaka, Y., and Park, I.H. (2012). Human induced pluripotent stem cells and neurodegenerative disease: prospects for novel therapies. Curr. Opin. Neurol. 25, 125-130.

Kano, H., Godoy, I., Courtney, C., Vetter, M.R., Gerton, G.L., Ostertag, E.M., and Kazazian, H.H. (2009). L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev. 23, 1303-1312.

Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Izpisúa Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature 460, 1140-1144.

Kazazian, H.H., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and Antonarakis, S.E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332, 164-166.

Kim, K.Y., Hysolli, E., and Park, I.H. (2011). Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. Proc. Natl. Acad. Sci. USA 108, 14169-14174.

Kinomoto, M., Kanno, T., Shimura, M., Ishizaka, Y., Kojima, A., Kurata, T., Sata, T., and Tokunaga, K. (2007). All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Res. 35, 2955-2964.

Kolosha, V.O., and Martin, S.L. (1997). In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. Proc. Natl. Acad. Sci. USA 94, 10155-10160.

Kondo, Y., and Issa, J.P. (2003). Enrichment for histone H3 lysine 9 methylation at Alu repeats in human cells. J. Biol. Chem. 278, 27658-27662.

Krull, M., Brosius, J., and Schmitz, J. (2005). Alu-SINE exonization: en route to protein-coding function. Mol. Biol. Evol. 22, 1702-1711.

Kunarso, G., Chia, N.Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.S., Ng, H.H., and Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. 42, 631-634.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., et al. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. 22, 908-917.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Lefort, N., Feyeux, M., Bas, C., Féraud, O., Bennaceur-Griscelli, A., Tachdjian, G., Peschanski, M., and Perrier, A.L. (2008). Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. Nat. Biotechnol. 26, 1364-1366.

Lev-Maor, G., Ram, O., Kim, E., Sela, N., Goren, A., Levanon, E.Y., and Ast, G. (2008). Intronic Alus influence alternative splicing. PLoS Genet 4, e1000204.

Liang, Q., Conte, N., Skarnes, W.C., and Bradley, A. (2008). Extensive genomic copy number variation in embryonic stem cells. Proc. Natl. Acad. Sci. USA 105, 17453-17456.

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471, 68-73.

Lyon, M.F. (1998). X-chromosome inactivation: a repeat hypothesis. Cytogenet. Cell Genet. 80, 133-137.

Lyon, M.F. (2000). LINE-1 elements and X chromosome inactivation: a function for 'junk' DNA? Proc. Natl. Acad. Sci. USA 97, 6248-6249.

Macia, A., Muñoz-Lopez, M., Cortes, J.L., Hastings, R.K., Morell, S., Lucena-Aguilar, G., Marchal, J.A., Badge, R.M., and Garcia-Perez, J.L. (2011). Epigenetic control of retrotransposon expression in human embryonic stem cells. Mol. Cell. Biol. 31, 300-316.

Marchetto, M.C., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell 143, 527-539.

Martens, J.H., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J. 24, 800-812.

Martin, S.L., and Bushman, F.D. (2001). Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol. Cell. Biol. 21, 467-475.

Martin, S.L., Cruceanu, M., Branciforte, D., Wai-Lun Li, P., Kwok, S.C., Hodges, R.S., and Williams, M.C. (2005). LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. J. Mol. Biol. 348, 549-561.

Mathias, S.L., Scott, A.F., Kazazian, H.H., Boeke, J.D., and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. Science 254, 1808-1810.

Mätlik, K., Redik, K., and Speek, M. (2006). L1 antisense promoter

drives tissue-specific transcription of human genes. J. Biomed. Biotechnol. *2006*, 71753.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766-770.

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. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. Nature *468*, 443-446.

Musova, Z., Hedvicakova, P., Mohrmann, M., Tesarova, M., Krepelova, A., Zeman, J., and Sedlacek, Z. (2006). A novel insertion of a rearranged L1 element in exon 44 of the dystrophin gene: further evidence for possible bias in retroposon integration. Biochem. Biophys. Res. Commun. *347*, 145-149.

Närvä, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., Borghese, L., Itskovitz-Eldor, J., Rasool, O., Dvorak, P., et al. (2010). High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. Nat. Biotechnol. *28*, 371-377.

Nigumann, P., Redik, K., Mätlik, K., and Speek, M. (2002). Many human genes are transcribed from the antisense promoter of L1 retrotransposon. Genomics *79*, 628-634.

Ohnishi, Y., Totoki, Y., Toyoda, A., Watanabe, T., Yamamoto, Y., Tokunaga, K., Sakaki, Y., Sasaki, H., and Hohjoh, H. (2010). Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. Nucleic Acids Res. *38*, 5141-5151.

Osenberg, S., Paz Yaacov, N., Safran, M., Moshkovitz, S., Shtrichman, R., Sherf, O., Jacob-Hirsch, J., Keshet, G., Amariglio, N., Itskovitz-Eldor, J., et al. (2010). Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. PLoS One *5*, e11173.

Ostertag, E.M., Goodier, J.L., Zhang, Y., and Kazazian, H.H. (2003). SVA elements are nonautonomous retrotransposons that cause disease in humans. Am. J. Hum. Genet. *73*, 1444-1451.

Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. Nature *451*, 141-146.

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., et al. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature *473*, 394-397.

Peng, Z., Cheng, Y., Tan, B.C., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., et al. (2012). Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat. Biotechnol. *30*, 253-260.

Pomp, O., Dreesen, O., Leong, D.F., Meller-Pomp, O., Tan, T.T., Zhou, F., and Colman, A. (2011). Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. Cell Stem Cell *9*, 156-165.

Quinlan, A.R., Boland, M.J., Leibowitz, M.L., Shumilina, S., Pehrson, S.M., Baldwin, K.K., and Hall, I.M. (2011). Genome sequencing of mouse induced pluripotent stem cells reveals retroelement stability and infrequent DNA rearrangement during reprogramming. Cell Stem Cell *9*, 366-373.

Richards, M., Tan, S.P., Tan, J.H., Chan, W.K., and Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. Stem Cells *22*, 51-64.

Roy-Engel, A.M., El-Sawy, M., Farooq, L., Odom, G.L., Perepelitsa-Belancio, V., Bruch, H., Oyeniran, O.O., and Deininger, P.L. (2005). Human retroelements may introduce intragenic polyadenylation signals. Cytogenet. Genome Res. *110*, 365-371.

Ruzov, A., Tsenkina, Y., Serio, A., Dudnakova, T., Fletcher, J., Bai, Y., Chebotareva, T., Pells, S., Hannoun, Z., Sullivan, G., et al. (2011). Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res. *21*, 1332-1342.

Speek, M. (2001). Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol. Cell. Biol. *21*, 1973-1985.

Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., Van der Elst, J., Liebaers, I., and Sermon, K. (2008). Recurrent chromosomal abnormalities in human embryonic stem cells. Nat. Biotechnol. *26*, 1361-1363.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol. *12*, R54.

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Han, J.W., Kim, S., Namburi, S., Hermetz, K., Kim, J.J., Rudd, M.K., et al. (2011). Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. PLoS Genet. *7*, e1002154.

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. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tanaka, Y., Yamashita, R., Suzuki, Y., and Nakai, K. (2010). Effects of Alu elements on global nucleosome positioning in the human genome. BMC Genomics *11*, 309.

Taniguchi-Ikeda, M., Kobayashi, K., Kanagawa, M., Yu, C.C., Mori, K., Oda, T., Kuga, A., Kurahashi, H., Akman, H.O., DiMauro, S., et al. (2011). Pathogenic exon-trapping by SVA retrotransposon and rescue in Fukuyama muscular dystrophy. Nature *478*, 127-131.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tucker, B.A., Scheetz, T.E., Mullins, R.F., DeLuca, A.P., Hoffmann, J.M., Johnston, R.M., Jacobson, S.G., Sheffield, V.C., and Stone, E.M. (2011). Exome sequencing and analysis of induced pluripotent stem cells identify the cilia-related gene male germ cell-associated kinase (MAK) as a cause of retinitis pigmentosa. Proc. Natl. Acad. Sci. USA *108*, E569-576.

van den Hurk, J.A., Meij, I.C., Seleme, M.C., Kano, H., Nikopoulos, K., Hoefsloot, L.H., Sistermans, E.A., de Wijs, I.J., Mukhopadhyay, A., Plomp, A.S., et al. (2007). L1 retrotransposition can occur early in human embryonic development. Hum. Mol. Genet. *16*, 1587-1592.

Wahlstedt, H., Daniel, C., Enserö, M., and Ohman, M. (2009). Large-scale mRNA sequencing determines global regulation of RNA editing during brain development. Genome Res. *19*, 978-986.

Wallace, N., Wagstaff, B.J., Deininger, P.L., and Roy-Engel, A.M. (2008). LINE-1 ORF1 protein enhances Alu SINE retrotransposition. Gene *419*, 1-6.

Wang, Q., Khillan, J., Gadue, P., and Nishikura, K. (2000). Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. Science *290*, 1765-1768.

Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J., and Nishikura, K. (2004). Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. J. Biol. Chem. *279*, 4952-4961.

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. *20*, 1732-1743.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343-348.

Wissing, S., Montano, M., Garcia-Perez, J.L., Moran, J.V., and Greene, W.C. (2011). Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. J. Biol. Chem. *286*, 36427-36437.

Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011a). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. *25*, 679-684.

Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K.,

Sun, Y.E., and Zhang, Y. (2011b). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389-393.

Xie, H., Wang, M., Bonaldo, M.e.F., Smith, C., Rajaram, V., Goldman, S., Tomita, T., and Soares, M.B. (2009). High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. Nucleic Acids Res. *37*, 4331-4340.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxy-

lase in mouse embryonic stem cells. Mol. Cell *42*, 451-464.

Yu, F., Zingler, N., Schumann, G., and Strätling, W.H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res. *29*, 4493-4501.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.