

Retrotransposon: A Versatile Player in Human Preimplantation Development and Health

Yuliang Qu^{1,2}, Zsuzsanna Izsvák³, Jichang Wang^{1,2,*}

¹Department of histology and embryology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China

²Key Laboratory for Stem Cells and Tissue Engineering (Sun Yat-sen University), Ministry of Education, Guangzhou 510080, China

³Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin 13092, Germany

*Correspondence: wangjch53@mail.sysu.edu.cn (J.W.)

More than half of human genome is comprised of transposable elements (TEs), most of which are the retrotransposons (or class I TEs). As the name suggests, retrotransposons propagate via an RNA intermediate that is reversely transcribed and integrated into new genomic loci (a so-called 'copy and paste' mechanism). There are three main types of retrotransposons in human genome, including the long terminal repeats (LTRs), non-long terminal repeats (non-LTRs), and SINE-VNTR-Alu (SVA) elements. The LTR elements or endogenous retroviruses (ERVs) are derived from ancient retroviral infections. Over millions of years of evolution, multiple ERV-derived sequences have been fixed in vertebrate genomes, constituting around 8% of human genome. On the other hand, the non-LTRs, including long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs), approximately account for 30% of human genome. The LINE-1 (L1) is an autonomous element with the ability to produce the whole enzymatic machinery for mobilization. In contrast, as one type of non-autonomous hominid-specific retrotransposons, SVA can be potentially mobilized in *trans* by hijacking the retrotransposition machinery produced by autonomous L1 elements. Despite their substantial occupancy of the human genome, most of TEs remain silenced, and have been long referred as 'junk DNA' or 'fossil records' of ancestral TE invasions. Recently, thanks to the advances in multi-omics and genome-editing techniques, accumulating evidence suggests that the 'junk DNA' could occasionally have a critical function for the human development and health.

Retrotransposons in early human development

Following the evolution of the TEs, ERVs exhibit the original features of their ancestral proviruses to variable extents, ranging from full length elements to highly fragmented relics of the proviral genome. However, some elements occasionally get co-opted (domesticated or exapted) for a beneficial function for the host. For example, proteins encoded by HERVs might be repurposed to support human embryonic development. Such well-known examples are the retroviral envelope gene-coded syncytins (e.g. HERVW-1, HERV-FRD1), playing an essential role in placental syncytiotrophoblast formation. Furthermore, Rec, the accessory protein of the phylogenetically youngest HERVK (HML-2, human mouse mammary tumor virus like-2), was suggested to promote antiviral defense response in human preimplantation embryos (Grow et al., 2015). Besides the full-length ERV sequences

with partial coding potential, around 85% of human ERVs (HERVs) exist as solitary LTRs, termed 'solo LTRs', which provide a rich source of *cis*-regulatory elements for gene expression during human embryonic development. Intriguingly, various primate/hominoid-specific retrotransposons are stage-specifically reactivated during human preimplantation development, indicating potential roles of these retrotransposons in rewiring regulatory network for early human development. They usually serve as binding platforms for multiple pluripotency transcription factors, forming a regulatory network for pluripotency (Yu et al., 2022). Unexpectedly, LTR5_Hs elements, the human-specific regulatory sequences of HERVK, have been reported as functional enhancers that contribute to the specification of human primordial germ cells, suggesting the essential role of the phylogenetically youngest retrotransposons in human germline development (Xiang et al., 2022).

Importantly, the stage-specific activation of retrotransposons depends on precise epigenetic regulation during early human development (Xu et al., 2022; Yu et al., 2022). Recently, exploiting low-input ChIP-seq techniques, Xu et al. and Yu et al. independently discovered the significance of the stage-specific TE associated H3K9me3 landscape in human pre-implantation embryos (Xu et al., 2022; Yu et al., 2022). In early human embryos, similar to mice, specific Krüppel-associated box domain zinc finger proteins (KRAB-ZNFs) have an essential role in regulating the transcription of different LTR families via stage-specific deposition of H3K9me3 (Xu et al., 2022). During the transition from 2-cell to the 8-cell (8C) stage of human embryos, the decreased deposition of H3K9me3 on the hominoid-specific SVA_D retrotransposons facilitates the interaction of these elements with zygotic genome activation (ZGA) genes (Yu et al., 2022). Intriguingly, a set of SVA_D elements could function as enhancers to promote the human ZGA by providing binding sites for 8C transcription factors (e.g., DUXA and ZSCAN4) (Yu et al., 2022). Of note, CRISPR-KRAB-mediated silencing of SVA_D elements led to defective ZGA and compromised blastocyst formation, demonstrating the crucial role of the SVA-derived enhancers in early human development (Yu et al., 2022). Moreover, *de novo* deposition of H3K9me3 in trophectoderm prevents binding of pluripotency transcription factors to certain hominoid-specific retrotransposons which function as regulatory elements for ICM-specific genes, suggesting potential roles of these *de novo* H3K9me3 domains in lineage segregation in preimplantation embryos (Xu et

al., 2022; Yu et al., 2022). It will be of interest to explore whether retrotransposons also contribute to species-specific regulation for lineage specification in post-implantation embryos (Figure 1).

Retrotransposons in human diseases

Despite their potential functionalities in development, TEs remain strictly controlled by human genomic surveillance system in somatic cells, but can be dysregulated in pathological conditions or aging. As one human-specific ERV, aberrantly transcribed HERVK (HML-2) can activate the innate immune response to tumors and aging as well as viral infection, indicating complex roles of retrotransposons in pathological conditions, maybe, in a species-specific manner. The role of HERVK (HML-2) in neurodevelopment is particularly intriguing. Using human pluripotent stem cell-based model combined with the CRISPR-VPR system, Padmanabhan Nair et al. recently reported that aberrant reactivation of HERVK (HML-2) might impair cortical neuronal development by inducing premature expression of genes involved in neuronal differentiation (Padmanabhan Nair et al., 2021), implying that aberrant reactivation of HERVK may be implicated in various neurological disorders including neurodegenerative diseases. Most intriguingly, similar to that in human blastocysts (Grow et al., 2015), multiple copies of the phylogenetically young HERV-K (HML-2) have intact open reading frames and hold the potential to produce retrovirus-like particles (RVLPs) in various pathological conditions as well as senescent cells (Liu et al., 2021). However, the function of HERVK-derived RVLPs in tumor progression and ageing remains largely unknown. It is interesting to explore whether and how HERVK-derived RVLPs regulate immune response to cancer and ageing in the future.

LINE-1 (L1) was shown to affect the function of immune cells, which might also play crucial roles in immune response to cancer and ageing. Liang et al. recently revealed that BMAL1, an essential component of the molecular circadian clock, could prevent stem cell senescence via stabilizing heterochromatin and thus repressing activation of the L1-cGAS-STING pathway, highlighting that the aberrantly activated L1 transcription drives cellular senescence by stimulating IFN-I-directed innate immune signaling (Liang et al., 2022). Unexpectedly, the transcripts containing L1-derived sequences might contribute to maintaining the quiescent state of naive CD4⁺ T cells, whereas these transcripts are downregulated upon T cell activation (Marasca et al.,

2022). Expression of 'L1-containing transcripts' was also observed in dysfunctional T cells, suggesting their potential function in governing T cell exhaustion (Marasca et al., 2022).

Importantly, as the only autonomously mobile transposons in human, L1_Hs may induce insertional mutagenesis. Indeed, pan-cancer genomic analyses revealed that nearly half of human cancer cell lines contain somatic retrotransposition events, suggesting the relevant role of L1 insertion in some human oncogenesis by causing megabase-scale deletions (Rodriguez-Martin et al., 2020). However, one recent study shows the potential role of L1 in repressing tumor progression. The methyl H3K9-binding protein MPP8, a component of the HUSH complex, was shown to suppress L1 in myeloid leukemia, whereas the reactivation of L1 impairs leukemia cell growth *in vitro* and in xenotransplants, suggesting tumor-type-specificity of L1 retrotransposition (Gu et al., 2021).

Altogether, these recent studies support the notion that co-option of retrotransposons might contribute to human development. On the other hand, the frequently reported dysregulation of retrotransposons in pathological conditions underlines the significance of the genome surveillance system (Figure 1). In fact, like other eukaryotes, humans have evolved a panel of highly orchestrated molecular mechanisms to control TE activity, including the PIWI/piRNA machinery, KRAB-ZFP-KAP1 complex, DNA methylation, histone modifications, inhibition of splicing and RNA modifications. Given the impact of TEs in human aging, neurodegeneration, inflammation and cancer, it is of particular tempt to investigate the underlying mechanisms that keep them under control. Furthermore, despite the advances, how those TE-derived sequences have been co-opted in the human genome remains to be elucidated. Recent progresses in gene editing technology and long-read sequencing would pave the way to further decipher versatile roles of these elements in human embryonic development and diseases.

Acknowledgements

This work is supported by grants of National Key R&D Program of China (Nos.2018YFA0107003 and 2018YFA0108301), and National Natural Science Foundation of China (No.31871444) to J.W. We apologize to the authors of other high-quality articles that we were unable to cite owing to space limitations.

Author contributions

The manuscript was initiated and written by Y.Q., Z.I., and J.W.

Conflict of interest

The authors declare no competing interests.

Accepted Manuscript

References

- Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., *et al.* (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* 522, 221-225.
- Gu, Z., Liu, Y., Zhang, Y., Cao, H., Lyu, J., Wang, X., Wylie, A., Newkirk, S.J., Jones, A.E., Lee, M., *et al.* (2021). Silencing of LINE-1 retrotransposons is a selective dependency of myeloid leukemia. *Nature Genetics* 53, 672-682.
- Liang, C., Ke, Q., Liu, Z., Ren, J., Zhang, W., Hu, J., Wang, Z., Chen, H., Xia, K., Lai, X., *et al.* (2022). BMAL1 moonlighting as a gatekeeper for LINE1 repression and cellular senescence in primates. *Nucleic Acids Res* 50, 3323-3347.
- Liu, X., Liu, Z., Sun, L., Ren, J., Wu, Z., Jiang, X., Ji, Q., Wang, Q., Fan, Y., Cai, Y., *et al.* (2021). Resurrection of human endogenous retroviruses during aging reinforces senescence. *bioRxiv*, 2021.2002.2022.432260.
- Marasca, F., Sinha, S., Vadalà, R., Polimeni, B., Ranzani, V., Paraboschi, E.M., Burattin, F.V., Ghilotti, M., Crosti, M., Negri, M.L., *et al.* (2022). LINE1 are spliced in non-canonical transcript variants to regulate T cell quiescence and exhaustion. *Nature Genetics* 54, 180-193.
- Padmanabhan Nair, V., Liu, H., Ciceri, G., Jungverdorben, J., Frishman, G., Tchieu, J., Cederquist, G.Y., Rothenaigner, I., Schorpp, K., Klepper, L., *et al.* (2021). Activation of HERV-K(HML-2) disrupts cortical patterning and neuronal differentiation by increasing NTRK3. *Cell Stem Cell* 28, 1566-1581 e1568.
- Rodriguez-Martin, B., Alvarez, E.G., Baez-Ortega, A., Zamora, J., Supek, F., Demeulemeester, J., Santamarina, M., Ju, Y.S., Temes, J., Garcia-Souto, D., *et al.* (2020). Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nat Genet* 52, 306-319.
- Xiang, X., Tao, Y., DiRusso, J., Hsu, F.-M., Zhang, J., Xue, Z., Pontis, J., Trono, D., Liu, W., and Clark, A.T. (2022). Human reproduction is regulated by retrotransposons derived from ancient Hominidae-specific viral infections. *Nature Communications* 13, 463.
- Xu, R., Li, S., Wu, Q., Li, C., Jiang, M., Guo, L., Chen, M., Yang, L., Dong, X., Wang, H., *et al.* (2022). Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos. *Cell Stem Cell* 29, 1051-1066 e1058.

Yu, H., Chen, M., Hu, Y., Ou, S., Yu, X., Liang, S., Li, N., Yang, M., Kong, X., Sun, C., *et al.* (2022). Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during human preimplantation development. *Cell Stem Cell* 29, 1031-1050 e1012.

Accepted Manuscript

Figure 1. Retrotransposons in human development and disease.

Under proper control of the surveillance system, retrotransposons could regulate gene expression in both *cis* and *trans*. Functionally, human development might benefit from retrotransposons that exert multiple important roles in early embryogenesis, including zygote genome activation (ZGA) and lineage segregation in preimplantation embryos. On the other hand, aberrant expression of retrotransposons has been implicated in various pathological processes, such as tumorigenesis, autoimmune diseases and aging. The figure was prepared by Figdraw.

Accepted Manuscript

Figure 1

