이학박사학위논문

RNA 메틸화 변형에 의한 인간 L1 레트로트랜스포존의 진화 기작 연구

Evolutionary mechanism of human L1 retrotransposon by RNA m^{6}A modification

2021 년 8 월

서울대학교 대학원

생명과학부

황성연

RNA 메틸화 변형에 의한 인간 L1 레트로트랜스포존의 진화 기작 연구

Evolutionary mechanism of human L1 retrotransposon by RNA m^{6}A modification

지도교수 안 광 석

이 논문을 이학박사학위논문으로 제출함
2021 년 8 월

서울대학교 대학원
생명과학부
황성연

황성연의 이학박사학위논문을 인준함
2021 년 8 월

위 원 장 _____
부위원장 _____
위 원 _____
위 원 _____
위 원 _____

Evolutionary mechanism of human L1 retrotransposon by RNA m^6A modification

A Dissertation Submitted in Partial Fulfillment of
the Requirement for the Degree of
DOCTOR OF PHILOSOPHY

To the Faculty of
School of Biological Sciences
at
Seoul National University
by
Sung-Yeon Hwang

Date approved
July, 2021

CONTENTS

LIST OF TABLE AND FIGURES

.

LIST OF ABBREVIATIONS

LINE-1, L1: Long interspersed nuclear element 1

SINE: Short interspersed nuclear element

LTR : Long terminal repeat

ORF: Open reading frame

RNP: Ribonucleoprotein

m^6A: N6-methyladenosine

METTL3: Methyltransferase-like enzyme 3

ALKBH5: α-ketoglutarate-dependent dioxygenase AlkB homolog 5

FTO: Fat mass and obesity-associated protein

eIF3: Eukaryotic initiation factor 3

d4T: 2'-3'-didehydro-2'3'-dideoxythymidine

mblastI: blasticidin S deaminase gene

hESCs: human embryonic stem cells

qPCR: Quantitative real-time PCR

RIP: RNA immunoprecipitation

MeRIP: Methylated RNA immunoprecipitation

CLIP: Crosslinking-immunoprecipitation

FISH: fluorescence in situ hybridization

LEAP: L1 element amplification protocol

1. ABSTRACT

Sung-Yeon Hwnag

School of Biological Sciences
The Graduate School
Seoul National University

Long interspersed element-1 (LINE-1; L1) comprises a group of active autonomous retrotransposons in humans. L1s have undergone amplification and divergence over the last ~80 million years of primate evolution, and occupy approximately ~17% of the human genome. Since the mobility of L1 retrotransposons can pose a threat to genome integrity, the host has evolved to restrict L1 replication. However, mechanisms underlying L1 propagation out of the host surveillance remains unclear. Here, I propose a novel survival strategy of L1, which exploits RNA N6-methyladenosine (m^6A) modification. I discovered that m^6A 'writer' METTL3 facilitates L1 retrotransposition, whereas m^6A 'eraser' ALKBH5 suppresses it. The essential m^6A cluster that is located on L1 5′ UTR serves as a docking site for eukaryotic initiation factor 3 (eIF3), enhances translational efficiency and promotes the formation of L1

ribonucleoprotein. Furthermore, I traced a recent episode of L1 5′ UTR evolution by analyzing species-specific L1s in three different primates; human, chimpanzee, and gorilla, and found that the most functional m^6A motif-containing L1s have been positively selected and became a distinctive feature of evolutionarily young L1s. Thus, these results demonstrate that L1 retrotransposons hijack RNA m^6A modification system for its successful replication.

Keyword: Long interspersed element 1, N6-methyladenosine RNA modification, Primate evolution, METTL3, ALKBH5, eIF3

Student ID Nr. 2015-20455

2. INTRODUCTION

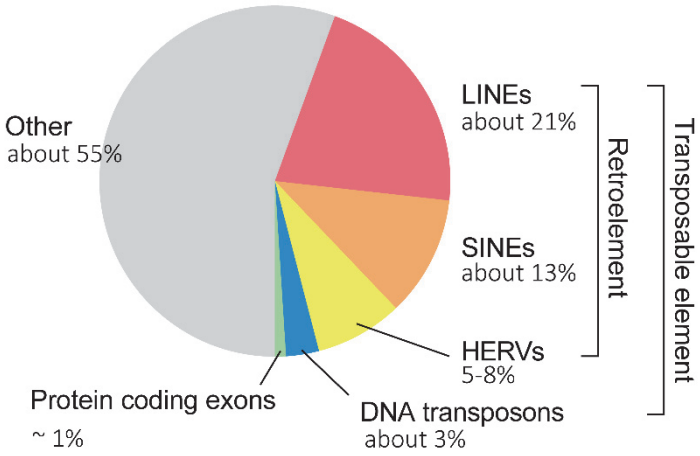
2.1. Transposable elements in human genome

Approximately 45% of the human genome is derived from transposable elements (also known as “jumping genes”) (Lander et al., 2001) (Figure 1A). There are two major groups of transposable elements: DNA transposons and retroelements. DNA transposons comprise about 3% of the human genome (Figure 1A). Currently, there are no active DNA transposons in the human genome (Lander et al., 2001).

Retroelements make up about 40% of the human genome (Figure 1A). They comprise two classes of retrotransposons: long terminal repeat (LTR) retrotransposons (also known as the human endogenous retroviruses, HERV) and non-LTR retrotransposons (Figure 1B). LTR retrotransposons are relics of past rounds of germline infection by retroviruses. Almost all LTR retrotransposons in the human genome lost their activity and seems to have been rendered nearly immobile with in last few million years (Dewannieux et al., 2006; Mager and Stoye, 2015; Payer and Burns, 2019).

Non-LTR retrotransposons are as old as the earliest multi-cellular organisms and their 28 clades have origins in the Precambrian Era of 600 million years ago (Goodier, 2016; Kapitonov et al., 2009; Malik et al., 1999). Two major groups of non-LTR retrotransposons are LINEs (long interspersed elements) and SINEs (short interspersed elements) (Figure 1B). Long interspersed element-1 (LINE-1; L1) is currently an active autonomous retrotransposon. The retrotransposition activity of L1 is mediated by the L1-encoded proteins. SINEs are nonautonomous elements and include Alu elements and SINE-R-VNTR (variable-number tandem repeat)-Alu (SVA) (Figure 1B). SINEs do not encode protein and require L1 proteins for their propagation.

A


Other
about 55%
LINEs
about 21%
SINEs
about 13%
HERVs
5-8%
Retroelement
Transposable element
Protein coding exons
~ 1%
DNA transposons
about 3%

B

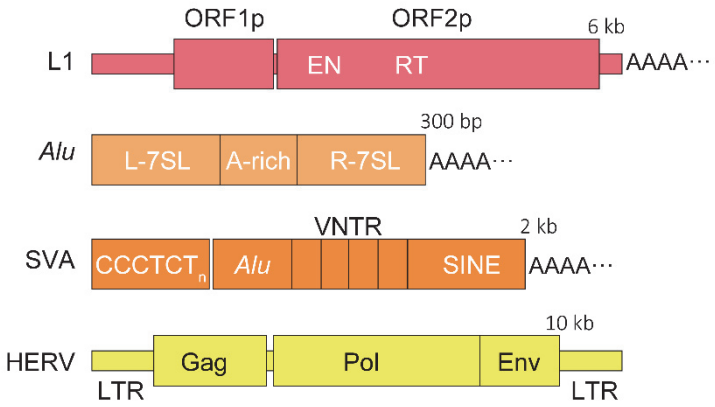

ORF1p
ORF2p
6 kb
L1
EN
RT
AAAA···
300 bp
Alu
L-7SL
A-rich
R-7SL
AAAA···
VNTR
2 kb
SVA
CCCTCT n
Alu
SINE
AAAA···
10 kb
HERV
Gag
Pol
Env
LTR
LTR

Figure 1. Schematic overview of retroelements in human genome

(A) The pie chart shows classes of human transposable elements and protein-coding exons as a percentage of the genome. (B) A schematic of human retroelements with their full-length size denoted. Long interspersed element1 (L1) encodes two open reading frames: ORF1p and ORF2p. ORF2p protein has endonuclease (EN) and reverse transcriptase (RT) domains. The non-autonomous elements SINEs (short interspersed nuclear elements) are non-coding RNAs that exploit L1 machinery for its mobility. SINES include *Alu* and SVA elements. Alu elements are bipartite, with the two arms derived from the signal recognition particle RNA 7SL and are about 280 bp long. SVA is a composite element containing variable number tandem repeats (VNTRs). LTR retrotransposons or human endogenous retroviruses (HERV) are flanked by long terminal repeats (LTR) and encode three retroviral proteins: Gag, Pol, and Env.

2.2. Long interspersed nuclear element 1 (LINE-1)

Long interspersed element-1 (LINE-1; L1) is currently an active autonomous retrotransposon in the primate genomes and utilizes a copy-and-paste mechanism for its expansion. The mobility of L1s has acted as a source of genome evolution and genetic variation (Goodier and Kazazian, 2008). L1s have been evolving and amplifying themselves during at least ~170 million years of mammalian radiation (Khan et al., 2006). As a result of this massive expansion, roughly 500,000 copies of L1s currently constitute a notable portion (~17%) of the human genome (Lander *et al.*, 2001). Most L1s are incapable of mobilization owing to truncation, rearrangement, or mutation (Grimaldi et al., 1984; Ostertag and Kazazian, 2001). However, 80-100 L1 copies remain retrotransposition-competent in the human genome (Beck et al., 2010; Brouha et al., 2003).

A retrotransposition-competent L1 is 6 kb in length and contains a 5′ untranslated region (UTR), two open reading frames (ORF1 and ORF2), and a short 3′ UTR. The 5′ UTR serves as an internal promoter for the transcription of L1 mRNA (Swergold, 1990). The L1-ORF1 encodes a 40-kDa protein (ORF1p) that acts as a nucleic acid chaperon (Martin and Bushman, 2001), while L1-

ORF2 encodes a 150-kDa protein (ORF2p) that acts as an endonuclease and reverse transcriptase (Feng et al., 1996; Mathias et al., 1991). ORF1p and ORF2p associate preferentially with their parental mRNA to form an L1 ribonucleoprotein (RNP) particle (Hohjoh and Singer, 1996; Kulpa and Moran, 2005). The L1 RNP enters the nucleus and then inserts its cDNA into a new genomic location via target-primed reverse transcription (Cost et al., 2002; Jurka, 1997; Luan et al., 1993) (Figure 2).

Since active L1 retrotransposition can pose a threat to genomic integrity, the host cell stringently restricts the L1 replication cycle. Numerous cellular factors suppress L1 retrotransposition by targeting the L1 promoter region or L1 RNP (Goodier, 2016). Under the selective pressures imposed by these factors, L1s contrived to evolve a strategy for evading repression (Castro-Diaz et al., 2014b; Jacobs et al., 2014). Sequence divergence could help L1s acquire or discard regulatory motifs for its adaptive evolution. However, there is limited information about the regulatory sequences that were acquired or lost and the roles of such events in L1 survival and evolution.

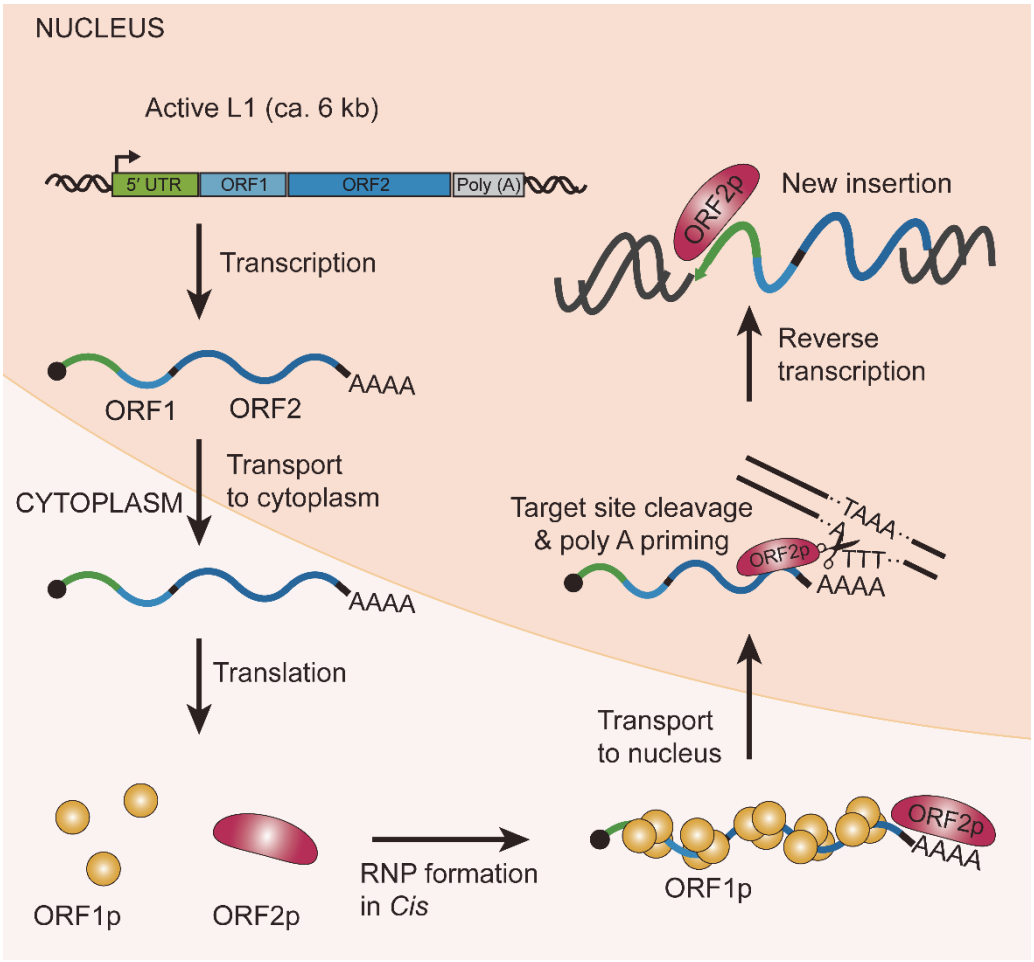

NUCLEUS
Active L1 (ca. 6 kb)
5′ UTR
ORF1
ORF2
Poly (A)
Transcription
AAAA
ORF1
ORF2
Transport
to cytoplasm
CYTOPLASM
AAAA
Translation
ORF1p
ORF2p
RNP formation
in Cis
ORF1p
ORF2p
AAAA
Transport
to nucleus
Target site cleavage
& poly A priming
ORF2p
TAAA
TTT
AAAA
Reverse
transcription
ORF2p
New insertion

Figure 2. L1 replication cycle

Long interspersed nuclear element 1 (L1) is the only autonomous mobile element that are active in humans. Full-length L1 is 6 kb in length and contains a 5′ untranslated region (UTR), two open reading frames (ORF1 and ORF2), and a short 3′ UTR. 5′ UTR of L1 contains an internal bidirectional DNA polymerase II (Pol II) promoter (green). ORF1 encodes an RNA-binding protein (L1 ORF1p, yellow) and the protein encoded by ORF2 (L1 ORF2p, magenta) has an endonuclease (EN) domain and reverse transcriptase (RT) domain. Once L1 RNA is transcribed and translated, the RNA interacts with its ORF proteins, forming a ribonucleoprotein (RNP) complex. RNP complex of L1 then moves into the nucleus, catalyzes a nick in the DNA and reverse transcribes its parental L1 mRNA for insertion.

2.3. The role of m^6A modification in the biological functions

N6-methyladenosine (m^6A) is the most prevalent internal modification in eukaryotic mRNAs, which determines RNA function and fate (Fu et al., 2014; Meyer and Jaffrey, 2014). Several enzymes dynamically process the m^6A modification of mRNA. The methyltransferase-like enzyme METTL3, which is the catalytic subunit of the RNA methyltransferase complex, adds m^6A at the consensus motif DRAmCH (where D = G/A/U, R = G/A, and H = U/C/A) (Dominissini et al., 2012; Liu et al., 2014; Meyer et al., 2012). Conversely, m^6A is removed by the demethylases α-ketoglutarate-dependent dioxygenase AlkB homolog 5 (ALKBH5) or fat mass and obesity-associated protein (Jia et al., 2011; Zheng et al., 2013). m^6A can be recognized specific reader proteins such as YTH-domain containing proteins, the mammalian IGF2 mRNA-binding protein family (IGF2BP proteins), heterogeneous nuclear ribonucleoprotein A2B1, and eukaryotic initiation factor 3 (eIF3). These proteins regulate splicing, stability, translation and/or localization of mRNAs (Figure 3).

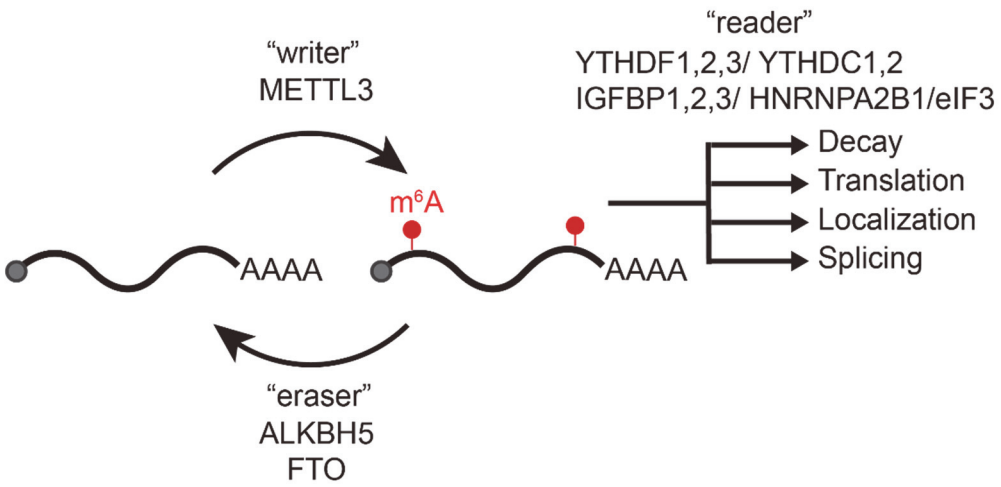


Figure 3. RNA m^{6}A modification determines the fate of mRNA

The m^{6}A 'writer' complex, which comprises the core methyltransferase-like protein 3 (METTL3) installs m^{6}A co-transcriptionally. The m^{6}A 'erasers', ALKBH5 and FTO are demethylases that convert m^{6}A into A. m^{6}A affects the fate of mRNA by recruiting m^{6}A-binding proteins. The YTH domain-containing proteins (YTHDF1,2,3 and YTHDC1,2), The mammalian IGF2 mRNA-binding protein family (IGF2BP proteins), heterogeneous nuclear ribonucleoprotein A2B1, and eukaryotic initiation factor 3 (eIF3) bind to m^{6}A and affect the stability, translation, splicing, and/or localization of mRNAs.

2.4. m^{6}A modification and L1 retrotransposon RNA

m^{6}A modification regulates a variety of physiological events, including stem cell differentiation, circadian rhythms, fertility, and microRNA biogenesis (Alarcon et al., 2015; Batista et al., 2014; Fustin et al., 2013; Wang et al., 2014; Zheng *et al.*, 2013). Emerging studies have revealed that m6A modification in viral transcripts affects the gene expression and replication of viruses such as HIV-1 and HCV (Gokhale et al., 2016; Lichinchi et al., 2016). Despite the critical role of m6A in pathogenic viral transcripts, it remains unclear whether m6A participates in the regulation of the endogenous parasites, L1 retrotransposons.

Here, I demonstrate that m^{6}A metabolism is crucial for L1 mobility through siRNA-mediated silencing or the ectopic expression of the RNA m^{6}A machinery. To verify the role of m^{6}A in the L1 replication cycle, I mapped m^{6}A RNA modifications in L1 RNA and identified the m^{6}A cluster consisting of three m^{6}A sites in L1 5′ UTR. Notably, the loss of these m^{6}A sites abrogates L1 retrotransposition, which suggests the essential role of m^{6}A in L1 mobility. I further revealed the functional role of m^{6}A in enhancing L1 translational efficiency through the recruitment of the known m^{6}A binding protein,

eukaryotic initiation factor 3 (eIF3). The upregulated translation by m^6A promotes the formation of L1 RNP particles. Lastly, I traced the evolutionary history of the L1 5′ UTR by scrutinizing the species-specific lineage of L1 subfamilies from human, chimpanzee, and gorilla genomes. The A332 m^6A site as the primary binding site of eIF3 appeared in the L1PA3 subfamily lineage ~12 million years ago. Since then, it has become the distinctive feature of evolutionarily young L1s in all three primates. my findings suggest that the m^6A modification of L1 5′ UTR is essential for L1 retrotransposition and is one of the driving forces for successful L1 expansion in primate genomes.

3. MATERIALS AND METHODS

3.1. Cells

HeLa cells were cultured in DMEM supplemented with 10% (v/v) fetal bovine serum (FBS, HyClone) and 1% (v/v) GlutaMAXI (Gibco). Human embryonic carcinoma PA-1 cells were cultured in RPMI 1640 supplemented with 10% (v/v) FBS (HyClone) and 1% (v/v) GlutaMAXI (Gibco). hESCs (H9, Wicell Research) were cultured in defined hESC culture medium (Stem Cell Technology) on hESC-qualified extracellular matrix (Corning)-coated culture dishes (Corning) or on tissue culture wall plate (Falcon). The cultures were incubated at 37°C in 5% CO_2.

3.2. Antibodies

Antibodies used in the experiments are listed below: Anti LINE-1 ORF1p Antibody, clone 4H1 (Merck, MABC1152); Anti METTL3 antibody (Abcam, ab195352); Anti ALKBH5 antibody (Novus, NBP1-82188); Anti FTO antibody (Abcam, ab124892); Anti-N6-methyladenosine (m^6A) Antibody (Merck, ABE572); Anti Vinculin antibody (Sigma, V9131); Anti HSP70 antibody

(Stressgen, SPA-810); Anti-FLAG® M2 antibody (Sigma, F3165); Anti-T7 tag® antibody (Abcam, ab9138); Anti eIF3b antibody (Bethyl, A301-761A); Anti-pan Ago Antibody, clone 2A8 (Merck, MABE56); Anti GAPDH antibody (AbFrontier, LF-PA0212); Anti GFP antibody (Santa Cruz Biotechnology, SC-9996); Anti HA antibody (Cell signaling, 3724); Peroxidase AffiniPure Goat Anti-Mouse IgG (Jackson ImmunoResearch Laboratories, 115-035-062); Peroxidase AffiniPure Goat Anti-Rabbit IgG (Jackson ImmunoResearch Laboratories, 111-035-003); Fluorescein (FITC) AffiniPure Rabbit Anti-Goat IgG (Jackson ImmunoResearch Laboratories, 305-095-047)

3.3. Plasmids

The FLAG-HA-pcDNA3.1-derived plasmids used in this study were named using FH as a prefix with the respective protein names specified. AcGFP, ALKBH5, and FTO cDNA were cloned into FLAG-HA-pcDNA3.1 (Addgene, 52535) for overexpressing N-terminally FLAG-HA-tagged protein. FH-based plasmids were generated by restriction enzyme cloning using *Xba*I and *Pme*I (NEB). The site-directed mutagenesis of FH-ALKBH5 to catalytically inactive mutant (H204A) construct was performed using the Phusion High-Fidelity polymerase kit (Thermo Fisher Scientific).

pJJ101-L1-dn6 2.2, which is referred to as pL1Hs in this study, is a pCEP4-based plasmid that contains an active human L1(L1-dn6) and was generously provided by J. L. Garica-Perez (Klawitter et al., 2016). For the mutagenesis of L1 5′ UTR m^6A sites, the L1-dn6 5′ UTR and ORF1 region containing *Not*I and *Age*I restriction sites was recloned into plasmid pCMV14. Using site-directed mutagenesis PCR, the following mutants of pCMV14 L15′ UTR ORF1 plasmids were prepared: Δ5′ UTR, A332T, A495T, A600T, A332/600T, and A332/495/600T (named as m^6A mut). Next, *Not*I-*Age*I fragments of pCMV14 L15′ UTR ORF1 mutant constructs were amplified and subcloned into pL1Hs. To generate pL1PA2$^{5′\ UTR}$, I synthesized the L1PA2 5′ UTR region based on reported consensus sequences using gene synthesis (Cosmogenetech). I then replaced the L1Hs 5′ UTR of pL1Hs with L1PA2 5′ UTR, as described above.

L1-firefly luciferase-tagged plasmids pYX014 and pYX015 were gifts from W. An (Xie et al., 2011). pYX014 encodes L1 constructs under the L1 native 5′ UTR promoter. pYX015 carries a retrotransposition-defective mutation in L1 ORF1. pYX014 and pYX015 plasmids contain a *Renilla* luciferase cassette to normalize transfection efficiency levels. To generate pYX014 L1Hs and m^6A mut constructs, *Not*I-*Pml*I fragments that were 2166 bp in length, including those spanning from the L1 5′ UTR to the forepart of ORF2 in pL1Hs and pL1

m^{6}A mut, were subcloned into pYX014 via restriction enzyme cloning.

pAD3TE1 is an L1.3 plasmid containing the T7 gene 10 epitope tag on the carboxyl-terminus of ORF1p, TAP tag on the carboxyl-terminus of ORF2p, and 24 copies of the MS2 loop repeat in the 3′ UTR (Doucet et al., 2010). pAD3TE1 was gift from Aurélien J. Doucet. The generation of L1 5′ UTR m^{6}A mutant constructs of pAD3TE1 was performed according to the method for pL1Hs mutant construct generation.

L1-neo-TET, a codon-optimized synthetic L1 construct, was generously provided by Astrid Roy-Engel (Addgene, 51284). The L1-neo-TET lacks a 5′ UTR. To generate a 5′ UTR-containing L1-neo-TET construct, the 5′ UTR of pL1Hs was amplified using PCR and the amplicon was inserted downstream of the CMV promoter of L1-neo-TET.

pFR-L1Hs 5′ UTR plasmids were generated by restriction enzyme cloning. The L1 5′ UTR of pL1Hs and firefly luciferase of pGL3-Basic (Promega) were cloned into pCMV14 downstream of the CMV promoter. Thereafter, the neomycin-resistant gene located downstream of the SV40 promoter was substituted with *Renilla* luciferase gene encoded by pYX014. Site-directed mutagenesis was used to generate the following m^{6}A motif-abrogating mutants of pFR-based plasmids: A332T, A495T, A569T, A600T, A679T, A758T, A839T,

A332/600T, and A332/495/600T.

pDEST HA-derived plasmids were named using HA as a prefix with the respective proteins, YTHDF1 and YTHDF2. YTHDF1 and YTHDF2 cDNA were cloned into pDEST HA vector using pENTR/D-TOPO vector (Invitrogen) and Gateway® LR Clonase® II Enzyme mix (Invitrogen). HA tag sequence is located in 5′ end of insert for overexpressing N-terminally HA-tagged protein.

Lentiviral short-hairpin RNA (shRNA) constructs for LacZ, METTL3, and ALKBH5 were generated by cloning into the pLKO.1 vector (Addgene, 10878) as per the protocol mentioned in Addgene. The following oligonucleotides were used for generating lentiviral shRNA constructs:

shLacZ	5′ CCGGGCGATCGTAATCACCCGAGTGCTCGAGCACTC GGGTGATTACGATCGCTTTTTG 3′
shMETTL3	5′ CCGGGCCAAGGAACAATCCATTGTTCTCGAGAACA ATGGATTGTTCCTTGGCTTTTTG 3′
shALKBH5	5′ CCGGCCTCAGGAAGACAAGATTAGACTCGAGTCTA ATCTTGTCTTCCTGAGGTTTTTG 3′

3.4. Immunoblotting

The cells were lysed in RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 1 mM DTT) supplemented with cOmplete protease inhibitor cocktail (Roche) for 15 min on ice. The lysates were cleared by centrifugation and mixed with Laemmli sample buffer. The mixture was then boiled at 98°C for 10 min, separated by SDS-PAGE in 10% gels, and transferred onto nitrocellulose blotting membranes (Amersham). The membranes were blocked by incubating with 5% skim milk in Tris-Buffered Saline Tween-20 (TBST) for 30 min and incubated overnight at 4°C with the respective primary antibodies at 1:1,000 dilution, except for anti-EIF3b antibody at 1:2,000. Subsequently, the membranes were washed thrice with TBST and incubated with HRP-conjugated secondary antibodies (Peroxidase AffiniPure Goat Anti-Mouse IgG, 115-035-062, Jackson ImmunoResearch Laboratories or Peroxidase AffiniPure Goat Anti-Rabbit IgG, 111-035-003, Jackson ImmunoResearch Laboratories) at 1:5,000 dilution in 5% skim milk/TBST. After washing thrice with TBST, the immunocomplexes were imaged using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific).

3.5. RNA interference

siRNAs directed against METTL3 (L-005170-02), ALKBH5 (L-004281-01), FTO (L-004159-01), or non-targeting siRNAs (D-001210-01-50) were purchased from Dharmacon. All siRNA transfections were performed using the DharmaFECT 1 transfection reagent (Dharmacon) according to the manufacturer's instructions.

3.6. RNA extraction and RT-qPCR

Total RNA was extracted using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. For the removal of genomic or plasmid DNA, total RNA was treated with recombinant DNase I (Takara) for 1 h at 37°C, followed by purification using the NucleoSpin RNA Clean-up kit (Macherey-Nagel). cDNA synthesis was performed using the ReverTra Ace qPCR RT Kit (Toyobo) according to the recommended protocol. TOPreal^{TM} qPCR 2X PreMIX (Enzynomics) was used for subsequent qPCR reactions. The qPCR primers used in this study are listed in Table 1.

3.7. L1 *mblastI* retrotransposition assay

HeLa cells were plated at 8×10^4 cells per well in 12-well plates. After 18 h, the cells were transfected with L1 plasmid (pJJ101-L1 dn6 2.2; pL1Hs) at 800 ng per well using Lipofectamine 3000 (Invitrogen) following manufacturer's instructions. Two days later, the medium was exchanged with a medium supplemented with 200 μg/ml hygromycin B (Invitrogen) to select the transfected cells. Cell selection continued for 4 days, and the hygromycin B-resistant cells were re-seeded at 2.5×10^4 per well in a 6-well plate. The next day, blasticidin S (Invitrogen) was added to a final concentration of 8 μg/ml and the cells were cultured for 7-9 days in its presence. The colonies were stained with crystal violet and counted using Colony, version 1.1 (Fujifilm). Retrotransposition assays were performed using RNA interference targeted toward METTL3, ALKBH5, and FTO with slight modifications in the process described above. For this, 6×10^4 HeLa cells were seeded into 12-well plates with 40 nM siRNA-Dharmafect1 (Dharmacon) mixture. After 24 h, the cultures were divided equally and plated into 2 wells in 12-well plates. The next day, the cells were transfected with pL1Hs at 500 ng per well. Four days after transfection, the cells were plated at 6×10^4 cells per well in 6-well plates and selected using 8 μg/ml of blasticidin S. Retrotransposition assays with the

overexpression of AcGFP, ALKBH5, and FTO were performed as described above. Briefly, HeLa cells were transfected 500 ng FH-plasmid and 700 ng of L1 plasmid, and four days after transfection, 6×10^4 cells were re-seeded into a well in a 6-well plate and selected after treatment with blasticidin S for 7-9 days.

3.8. Luciferase assay

HeLa cells were plated at 8×10^4 cells per well in 12-well plates. The next day, the cells were transfected with 800 ng per well of the pFR vector (pCMV-L15′ UTR-firefly luciferase) using Lipofectamine 3000 (Invitrogen). Two days later, the transfected cells were harvested and luminescence was measured using the Dual-Luciferase Reporter Assay System (Promega) according to manufacturer's instruction. Briefly, 250 μl of passive lysis buffer was used to lyse cells in each well in 12-well plates. Next, 20 μl of the lysate was mixed with 100 μl of the Luciferase Assay Reagent II, and the luminescence of firefly luciferase was measured using a microplate luminometer (BERTHOLD). *Renilla* luciferase activity was subsequently measured after administering 100 μl Stop & Glo Reagent.

3.9. L1 luciferase retrotransposition assay

Initially, 8×10^4 HeLa cells were plated into 12-well plates. After 18 h, the cells were transfected with 800 ng of L1 plasmid (pYX014) per well using Lipofectamine 3000 (Invitrogen). After 48 h, the medium was exchanged with a medium supplemented with 1 μg/ml of puromycin (Invitrogen). The cells were further selected for 2 days. Luminescence measurement was performed, as described in the section on luciferase assay. In experiments using HeLa cells devoid of eIF3b, cells were not selected with puromycin but were cultivated for 5 days.

3.10. Quantification of transfected pL1 plasmid

Genomic DNA and transfected DNA were purified using the QIAamp DNA Blood Mini Kit (Qiagen) following the manufacturer's instructions. For qPCR analyses of purified DNA, Equivalent amounts of DNA (50 - 100 ng) from each sample were subject to qPCR reaction. Data were normalized to *MDM2*, and L1 plasmid were detected using reporter L1-specific primer. qPCR primers are listed in Table 1.

3.11. Annexin V-APC staining assay

Cells were labeled with 5 μl of APC-conjugated Annexin V (BioLegend, 640920) according to the manufacturer's instructions. After incubation in the dark for 20 min at RT, the cells were analyzed by flow cytometry using Flow-Activated Cell Sorter Canto II (BD Bioscience) and Flowjo software (version 10, Flowjo). The AnnexinV-APC-negative population was regarded as alive, whereas the Annexin V-APC–positive populations were taken as measurements of apoptotic/necrotic cells.

3.12. RNA stability assay

HeLa cells were plated at 1.5×10^5 cells per well in a 6-well plate. The following day, the cells were transfected with 1.5 ug of pL1Hs or pL1 m^6A mut using Lipofectamine 3000 (Invitrogen). After 24 h, the cells were re-seeded into 4 wells of a 12-well plate. After 24 h, the cells were treated with 10 μg/ml actinomycin D added at 6, 4, 2, and 0 h before RNA extraction. RNA extraction and RT-qPCR were performed as described in the method section.

3.13. Nuclear/Cytosolic fractionation

HeLa cells were plated at 1.5 × 10^5 cells per well in a 6-well plate. The following day, the cells were transfected with 1 ug of pL1Hs and pL1 m^6A mut using Lipofectamine 3000 (Invitrogen). After 48 h, the cells were fractionated using PARIS kit (Invitrogen) according to the manufacturer's protocol for all steps except the RNA extraction step. RNA was purified using TRIzol reagent and treated with rDNaseI, as described in RNA extraction section.

3.14. Polysome fractionation

Ten milliliter of 10-50% linear sucrose gradients in base solution (100 mM NaCl, 20 mM Tris pH 7.5, 10 mM $MgCl_2$, 100 μg/ml of cycloheximide) were prepared a day before polysome fractionation. For polysome fixation, 7-10 × 10^6 cells were incubated for 10 min in a media containing 100 μg/ml of cycloheximide at 37°C and were collected by scrapping with PBS containing 100 μg/ml of cycloheximide. After centrifugation at 1200 × *g*, 4°C for 5 min, the cell pellets were lysed in 100 μl of polysome extraction buffer (20 mM Tris pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 0.5% NP-40) supplemented with RNase inhibitor (Enzynomics), protease/phosphatase inhibitor cocktail (Cell signaling),

and 1 mM DTT. The cells were incubated in the buffer for 10 min at 4°C and centrifuged at 12,000 × *g*, 4°C for 10 min to remove debris and nuclei. The protein concentration was measured using the Pierce BCA Protein Assay kit (Thermo Fisher Scientific). Five hundred to six hundred microgram of the lysate was introduced at the top of the linear sucrose gradient and centrifuged at 222,000 × *g*, 4°C for 2 h using the SW41Ti rotor of the Beckman ultracentrifuge. Fifty microgram of the lysate was saved as input RNA. After centrifugation, 1 ml fractions were collected from the top to the bottom of the gradient using the BioLogic LP system and fraction collector (BioRad) with UV absorbance at 260 nm. Next, 250 µl of each fraction was mixed with 750 µl of TRIzol LS reagent (Invitrogen) and 20 ng of spike-in RNA (synthesized firefly luciferase mRNA). RNA extraction and qPCR were performed as described above. The levels of RNA in each fraction were normalized to those of spike-in RNA and input RNA.

3.15. YTHDFs RNA-immunoprecipitation

HeLa cells were plated at 1×10^6 cells in 100 mm dishes. The following day, the cells were transfected with 3.5 µg of pL1Hs or m^6A mut and 3.5 µg of HA-YTHDF1 or 2. After 48 h, cells were harvested and resuspended in RIP lysis

buffer (150 mM KCl, 10 mM HEPES (pH 7.6), 2 mM EDTA, 0.5% NP40, 1 mM DTT, cOmplete protease inhibitor cocktail (Roche), 400 unit/ ml RNase inhibitor (Enzynomics)) for 10 min at 4°C. Lysates were cleared by centrifugation (4°C, 15,000 × *g* for 15 min) and filtrated by passing through a 0.45-μm membrane syringe filter. Input samples for RNA extraction and immunoblot assay were saved respectively (10% of lysates). For antibody-bead preparation, 4 μl of HA antibody (Cell signaling, 3724) was diluted in 80 μl of NT2 buffer (200 mM NaCl, 50 mM HEPES (pH7.6), 2 mM EDTA, 0.05% NP40) and incubated with Dynabead protein G for 30 min with rotation at room temperature. After incubation, antibody-bead complex were washed and resuspended in a volume of NT2 buffer equivalent to lysate. Lysates were then added into bead-containing tube and incubated overnight on a rotating wheel at 4°C. Afterwards, the beads were subject to wash with 5 × 1 ml portions of ice-cold NT2 buffer. The 20% of beads were resuspended in 40 μl of RIP lysis buffer and analyzed by western blotting with the input samples. The rest 80% of beads as well as saved input RNA samples were mixed with 1 ml of TRIzol (Invitrogen), supplemented with 50 ng of spike-in RNA. RNA was extracted using 1 ml of TRIzol supplemented with 50 ng of spike-in RNA.

3.16. Northern blot

HeLa cells were seeded in 100 mm dishes at 1 × 10^6 cells. Next day, 7 μg of reporter-deleted pL1 construct, pL1Hs$^{\Delta BLA}$ or pL1m^{6}A mut$^{\Delta BLA}$ were transfected into cells. At 3 d post transfection, poly (A) + RNA was purified using TRIzol reagent and the Poly (A) purist Mag kit (Invitrogen). The eluted RNA was purified by ethanol precipitation, and dissolved in formaldehyde load dye (Invitrogen). 2 μg of poly (A) + RNA was separated in 1% formaldehyde agarose gels in MOPS gel running buffer (NorthernMax kit, Invitrogen) according to the manufacturer's protocol. After electrophoresis, the lane containing the size markers were cut off and stained separately with EtBr. RNA was then transferred to BrightStar Nylon membranes (Invitrogen) through capillary transfer for 4 h. RNA-bound mebranes were cross-linked using UV light (120 mJ/cm^2) and preincuabated in PerfectHyb hybrydization buffer (NorthernMax kit, Invitrogen) for 30 min at 68°C. Radioisotopes-labeled RNA probes specific to L1 5′ UTR, Hygromycin^R, or beta-actin (final concentration ~1×10^6 cps / ml) were mixed with PerfectHyb hybrydization buffer, and incubated with membrane overnight at 68°C with constant rotation. Membranes were then washed two times with low stringency wash solution (2× SSC, 0.1% SDS) and two times with high stringency wash solution (0.1× SSC, 0.1% SDS).

The blots were exposured to Fuji 32P screens, and scanned by Typhoon FLA7000 (version 1.2).

3.17. Northern blot probe production

RNA probe templates containing T7 promoter sequence were synthesized by PCR. Strand-specific RNA probes were generated and labeled with α-^{32}P UTP using the EZ™ T7 High Yield In Vitro Transcription kit (Enzynomics) according to the manufacturer's instructions. Briefly, 500-800 ng of templates were subject to in vitro transcription reaction (0.5 mM ATP, CTP, GTP, 1 mCi/ml α-^{32}P UTP, 1 mM DTT and 1 unit/μl RNase inhibitor) and incubated for 6 h. After reaction, rDNase I (Takara) was treated to remove template DNA. All probes were purified using NucleoSpin RNA clean-up (MACHEREY-NAGEL). Probe template sequence of L1 5′ UTR, Hygromycin^{R}, or *beta-actin* are listed in below. T7 promoter sequences are in bold.

L1 5′ UTR (1-232, 232bp)

5′ **TAATACGACTCACTATAGGG**CTCGGAAAGGGAACTCCCTGACCC
CTTGCGCTTCCCAGGTGAGGCAATGCCTCGCCCTGCTTCGGCTCGCG
CACGGTGCGCACACACACTGGCCTGCGCCCACTGTCTGGCACTCCCT
AGTGAGATGAACCCGGTACCTCAGATGGAAATGCAGAAATCACCGT
CTTCTGCGTCGCTCACGCAGGGAGCTGTAGACCGGAGCTGTTCCTAT
TCGGCCATCTTGGCTCCTCCC 3′

Hygromycin^{R} (181-326, 146bp)

5′ **TAATACGACTCACTATAGGG**CGTTATGTTTATCGGCACTTTGCA
TCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAG
CGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGT
TGCAAGACCTGCCTGAAACCGAACTGCC 3′

beta-actin (1464-1588, 125bp)

5′ **TAATACGACTCACTATAGGG**CCAAATATGAGATGCGTTGTTACA
GGAAGTCCCTTGCCATCCTAAAAGCCACCCCACTTCTCTCTAAGGAG
AATGGCCCAGTCCTCTCCCAAGTCCACACAGGGGAGGTGATAGCAT
TGCTTTCG 3′

3.18. Crosslinking immunoprecipitation and qPCR (CLIP-qPCR)

eIF3-RNA CLIP-qPCR was performed as described previously (Meyer et al., 2015) with some modifications. For each experiment, 1.2×10^6 HeLa cells were plated on two 100 mm dishes each. The next day, the cells were transfected with 6 μg of L1 plasmid per dish using Lipofectamine 3000 (Invitrogen). Two days later, the cells were washed twice with cold PBS, and allowed to form UV crosslinks on ice under 150 kJ/cm^2 of UV 254 nm light (XL-1500, Spectrolinker). The cells were scraped and transferred to PBS and pelleted by centrifugation at 1000 × *g*, 4°C for 3 min. The pellets were resuspended in 1 ml of lysis buffer (50 mM Tris pH 7.5, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5%

sodium deoxycholate, 1X cOmplete protease inhibitor cocktail, 1 mM DTT, 80 unit/ml RNase inhibitor). The lysate was passed through a 21G needle ten times and shock-frozen using liquid nitrogen. The lysate was thawed on ice and centrifuged at 15,000 × *g* for 15 min. The supernatant was further cleared by filtering through a 0.22 μm membrane. From each lysate, 5% was retained as input. For immunoprecipitation, 10 μl of Dynabeads Protein G (Invitrogen) was washed twice with lysis buffer and incubated with 3 μg of eIF3b antibody (A301-761A, Bethyl) on a rotating wheel at room temperature for 1 h. The cell lysates were mixed with the antibody-bead complex and rotated overnight at 4°C. The beads were washed five times in high-salt buffer (50 mM Tris pH 7.5, 500 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, 1 mM EDTA, 1 mM DTT, 80 unit/ml RNase inhibitor). The antibody-lysate mixture and the conserved input lysates were resuspended in 100 μl of 1X Proteinase K buffer (100 mM Tris pH 7.5, 50 mM NaCl, 10 mM EDTA, 1% SDS). Next, 1 mg Proteinase K (Macherey-Nagel) was added into the suspensions. Protein digestion was conducted at 50°C for 2 h in a shaking incubator. After incubation, 100 μl of 7 M Urea (w/v)-1X Proteinase K buffer was added into the immunoprecipitation samples, and the samples were re-incubated at 50°C for 2 h in a shaking incubator. RNA was extracted using TRIzol LS supplemented with 20 ng of spike-in RNA.

3.19. Methyl-RNA immunoprecipitation (MeRIP)-sequencing

MeRIP was performed as described earlier (Dominissini et al., 2013) with some modifications. HeLa cells were plated on two 100 mm dishes at 1.2×10^6 cells per dish. After 18 h, the cells were transfected with 8 μg of pL1Hs per dish. After 48 h, poly (A)+ RNA was extracted using the Poly (A) purist Mag kit (Invitrogen). The poly (A)+ RNA was mixed with RNA fragmentation reagents (Invitrogen) and fragmented into oligonucleotide that was 50-150 nt in length by heating to 75°C for 5 min. Fragmented RNA was purified by ethanol precipitation. Next, 6 μg of fragmented RNA was incubated with 4 μg of anti-m^6A antibody (Merck, ABE572) in MeRIP buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and 0.1% NP-40) on a rotating wheel for 2 h at 4°C. After that, the immunoprecipitation mixtures were mixed with Dynabead protein A (Invitrogen) and incubated overnight on a rotating wheel at 4°C. After washing five times with MeRIP buffer, RNA was eluted twice by incubating in elution buffer on a rotating wheel for 1 h at 4°C (6.7 mM m^6A sodium salt and 200 unit/ml RNase inhibitor-containing MeRIP buffer). The eluted RNA was purified by ethanol precipitation. cDNA libraries were prepared as previously described (Kim and Kim, 2019). Briefly, RNA was dephosphorylated using calf intestinal alkaline phosphatase (NEB) and labeled with γ-^{32}P-ATP using T4

polynucleotide kinase (Takara). RNA was separated by 10% urea-PAGE and purified from the excised gel corresponding to 50-150 nt RNA fragments. The extracted RNA was ligated to a 3′ adapter using T4 RNA ligase 2, truncated KQ (NEB). The RNA was then purified from free 3′ adapters by repeated gel excision. The 3′ adapter-ligated RNA was ligated to a 5′ adapter using T4 RNA ligase 1 (NEB) and subsequently reverse transcribed using SuperScript III reverse transcriptase (Invitrogen). The cDNA library was amplified by PCR using Phusion HF polymerase (Thermo Fisher Scientific), separated by 6% acrylamide gel electrophoresis, and purified by gel excision. The libraries were sequenced to 2 × 100 base-pair reads on the Illumina HiSeq 2500. The sequence of the 3′ and 5′ adapters, reverse transcription primer, and 5′ and 3′ PCR primers are listed in Table 2.

For MeRIP-seq analysis, the adapters were trimmed using Cutadapt (Martin, 2011) (cutadapt -g TACAGTCCGACGATC -A TGGAATTCTCGGGTGCCAAGG). The 3′ and 5′ adapter sequences in the first and second read in a pair (owing to the short insert size) were further trimmed and the read pairs with either reads < 18 bp were discarded. The remaining reads were then aligned to the combined human genome (hg19), and reporter L1 (pL1Hs) sequence using Spliced Transcripts Alignment to a

Reference (STAR) (Dobin et al., 2012) and peak calling was performed using MACS2 (Zhang et al., 2008). For analyzing the m^6A modifications in endogenous L1, the sequence reads from human embryonic stem cells were retrieved (Batista *et al.*, 2014) (accession code: GSE52600) and were aligned against L1Hs consensus sequence using STAR. The codes are available from https://github.com/hastj7373/merip-seq.

3.20. RNA FISH and Immunofluorescence

The L1 MS2-stem-loop constructs pAD3TE1 L1Hs and pAD3TE1 L1 m^6A mut were transfected into HeLa cells. The following day, the cells were re-seeded on sterile coverslips where 200 μg/ml hygromycin B was added for selection. After 3 days, the cells were fixed with 3.7% formaldehyde in PBS for 10 min, and the fixation was quenched by adding 0.1 M glycine in PBS for 10 min. The fixed cells were permeabilized in 70% ethanol for at least 3 h to 1 week at 4ºC. The cells were then rehydrated with PBS for 30 min and incubated in a pre-hybridization solution (10% formamide, 2X SSC solution) for 30 min at 37ºC. Hybridization was performed overnight at 37ºC in 50 μL of hybridization solution containing 10% formamide, 2X SSC, 10% dextran sulfate, 50 μg yeast tRNA, 0.2% BSA, 0.1 M DTT, 50 units RNase inhibitor (Enzynomics), and 10

ng of MS2-Q670 probe (generously provided by Hye Yoon Park (Lionnet et al., 2011); listed below) at 37ºC. Next, the cells were washed twice with a pre-hybridization solution for 30 min. For the immunofluorescence experiment, the hybridized cells were incubated in the blocking solution (10% formamide, 2X SSC, 0.2% BSA) for 1 h, followed by incubation with anti-T7 primary antibody (ab9138, Abcam) diluted in the blocking solution (1:200) for 2 h. The cells were washed twice with the pre-hybridization solution for 15 min and incubated with FITC-conjugated anti-goat secondary antibody (305-095-047, Jackson immunoresearch) diluted in blocking solution (1:200) for 1 h. The cells were washed twice as described above, and the coverslips were mounted on slide glasses using the Vectashield antifade medium with DAPI (Vector Laboratories). The samples were imaged using an inverted microscope Nikon Eclipse Ti2 equipped with a 1.45 numerical aperture Plan apo λ 100× oil objective and a sCMOS camera (Photometrics prime 95B 25 mm). For each field of view, stacks of images of 6 μm were captured at every 0.3 μm in the DAPI395, GFP488, and Alexa647 channels using the NIS-Elements software.

The sequence of the RNA FISH probes are: MS2LK20 (5′ TTTCTAGAGTCG ACCTGCAG 3′), MS2 LK51-1 (5′ CTAGGCAATTAGGTACCTTAG 3′), and MS2 LK51-2 (5′ CTAATGAACCCGGGAATACTG 3′). Each probe was

labeled with two Quasar 670 dyes at both ends. The mixture of the three probes were used for RNA FISH of L1 RNA tagged with the MS2 loops.

3.21. Co-localization analysis of RNA FISH and IFA microscope image

Binary masks of cells were generated using the ROI manager in ImageJ (Schneider et al., 2012). Protein and mRNA particles from z-stack images were detected using the TrackNTrace software (Stein and Thiart, 2016). After the detection of particles, the protein-mRNA pairs with an intermolecular distance of 330 nm (3 pixels) were considered as co-localizing pairs. The intensities of proteins co-localizing with mRNA were determined based on the amplitude of the fitted 2D Gaussian function from the TrackNTrace software.

3.22. Lentivirus production and viral transduction

Lentiviruses were produced by co-transfection of the pLKO.1-shRNA plasmid with packaging plasmid psPAX2(Addgene, 12260) and envelope plasmid pCMV-VSVG (Addgene, 8454) into HEK293T cells. After 48 h of post-transfection, viral supernatants were collected and filtrated through a 0.45 μM filter. Recombinant lentivirus was concentrated using Lenti-X concentrator (Takara) and stored at -80°C until use.

For shRNA-mediated knockdown, virus particles were added with 8 μg/ml of polybrene. The transduced cells were selected with 1 μg/ml puromycin after 72 h from viral infection.

3.23. Heat shock

HeLa cells were plated at 1.5×10^5 cells per well in a 6-well plate. The following day, the cells were transfected with 1 ug of pL1Hs and pL1 m^6A mut using Lipofectamine 3000 (Invitrogen). After 48 h, cells were incubated at 42°C for 1 h, and then were harvested at 9 h post heat shock stress. Cell lysates were subjected to western blot assay.

3.24. L1 element amplification protocol (LEAP)

The LEAP assay was performed as described previously (Kulpa and Moran, 2006). Briefly, HeLa cells were plated (4×10^6 cells in 60 mm dishes); the following day, the cells were transfected with 3 μg of L1 plasmid (pL1Hs) using Lipofectamine 3000 (Invitrogen). After 48 h, 200 μg/ml hygromycin B was added to the media to select the cells carrying the L1 plasmid. After 2 days of selection, the cells were lysed with CHAPS lysis buffer (10 mM Tris pH 7.5, 0.5% CHAPS (w/v), 1 mM $MgCl_2$, 1 mM EGTA, and 10% glycerol) supplemented with 1 mM DTT and the cOmplete protease inhibitor cocktail (Roche) and cleared by centrifugation (4°C, 20,000 × *g* for 15 min). The cleared lysates were loaded on a sucrose cushion (20 mM Tris pH 7.5, 80 mM NaCl, 8 mM $MgCl_2$, 1 mM DTT, 1X protease inhibitor, 4 ml of 8.5% sucrose (from the top) and 6 ml of 17% sucrose (from the bottom) solutions) in 13.2 ml Ultra-Clear tubes (Beckman Coulter) and centrifuged at 17,8000 × *g*, 4°C for 2 h in a SW41Ti rotor of Beckman ultracentrifuge. The colorless pellets were suspended by pipetting in 100 μL of RNase-free water supplemented with 1X protease inhibitors. Pierce BCA Protein Assay (Thermo fisher Scientific) was conducted to determine the protein concentration. Three microgram of the RNP samples were retained and used later in RNA isolation and immunoblotting experiments.

Seven hundred and fifty nanogram of each RNP sample was mixed with the LEAP assay reaction buffer (50 mM Tris pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, 0.05% Tween-20, 0.2 mM dNTPs, 1 mM DTT, 0.4 μM 3′ RACE adapter (5′ - GCG AGC ACA GAA TTA ATA CGA CTC ACT ATA GGT TTT TTT TTT TTV N -3′), 40 units of RNase inhibitor (Enzynomics), total reaction volume: 50 μl) and incubated at 37°C for 1 h. One microliter of LEAP cDNA products were subsequently amplified using 0.4 μM of L1 LEAP primer with the Phusion High-Fidelity polymerase kit (Thermo Fisher Scientific). PCR amplicons were separated and visualized in EtBr-stained 2% agarose gel.

3.25. eIF3 PAR-CLIP analysis

I utilized previously published PAR-CLIP data for eIF3a, b, d, and g (Lee et al., 2015). Briefly, the authors immunoprecipitated eIF3b from 4-thiouridine-and-UV-treated 293T cells to capture the eIF3-RNA complex. After high-salt washing and RNase digestion, they separated individual eIF3-RNA complexes through denaturing gel electrophoresis. eIF3a, b, d, and g were identified from four separate bands using mass-spectrometry and the interacting RNAs were purified and sequenced. Although three replicates were generated for each protein, only the first replicate was used for each. After retrieving the raw

sequence files from NCBI (accession code: GSE65004), reads with low basecall qualities were excluded using the fastq_quality_filter from FASTX Toolkit (-q 25 -p 80; version 0.0.13.2; http://hannonlab.cshl.edu/fastx_toolkit/). PCR duplicates were also excluded using fastx_collapser. Moreover, I excluded the reads that were shorter than 10-nt after trimming primer IDs and 3′ adapters from further analysis (cutadapt -a TGGAATTCTCGGGTGCCAAGG -u 12 -m 10; version 2.3; (Martin, 2011)). The remaining reads were mapped to the L1Hs consensus sequence, wherein upto three mismatches were allowed using bowtie2 (Langmead and Salzberg, 2012) (--local –norc –score-min L,-18,2; version 2.2.4). For mean coverage analysis of 5′ UTR, ORF1, ORF2, and 3′ UTR, the number of reads that begin and end within each region were counted and the number was divided by the length of the corresponding region. The codes used for analyzing PAR-CLIP and mapping data are available from https://github.com/schanbaek/eif3_par-clip.

3.26. Comparison analysis of species-specific m^{6}A site

To identify the species-specific full-length L1s in human, chimpanzee, and gorilla genome, I used BLAT-based and liftOver-based methods (Tang et al., 2018) with a computational approach. Only L1s of which insertion sites and

two flanking regions are supported to be unique to the human, chimpanzee or gorilla genome by BLAT and liftOver were included in the further analyses. Then, I eliminated certain ambiguous elements containing gap sequence in the reference genome data and those that were less than 5.5 kb. The flanking sequences (2 kb, both upstream and downstream) of each species-specific L1 candidate were manually compared to the orthologous loci in human (GRCh37/hg19; Feb. 2009), chimpanzee (CSAC Pan_troglodytes-3.0/panTro5; May. 2016), gorilla (GSMRT3/gorGor5; May. 2016), and orangutan (Susie_PABv2/ponAbe3; Jan. 2018) genomes. The flanking sequences were used to identify the orthologous positions in the other genomes using BLAST-Like Alignment Tool (BLAT). I collected and retrieved the species-specific full-length L1s. I then classified the L1 subfamilies (L1Hs, L1PA2~L1PA5) using RepeatMasker utility (Smit, 2013-2015). Multiple sequence alignment of species-specific full-length L1s in each genome was performed using MUSCLE (MUltiple Sequence Comparison by Log- Expectation) under the default option (Edgar, 2004). The conserved sequence motifs at the three sites (A332, A495, and A600) were visualized using the program Weblogo (Crooks et al., 2004).

3.27. Statistical analysis

GraphPad Prism 7.00 was used for statistical analysis. Two-sided student's t test was used for unpaired data. Two-sided Kolmogorov-Smirnov test was used to assess the quantification of number and intensity of co-localizing puncta in Figure 34C and D. For multiple comparisons, one-way ANOVA with Tukey's or Dunnett's multiple comparison test were used. *P* values <0.05 were considered significant.

Table 1. Primers used for qPCR or RT-PCR.

HPRT1	
F	TGA CAC TGG CAA AAC AAT GCA
R	GGT CCT TTT CAC CAG CAA GCT

SON	
F	TGA CAG ATT TGG ATA AGG CTC A
R	GCT CCT CCT GAC TTT TTA GCA A

CREBBP	
F	CAT GGC CAA GAT GGG AAT AA
R	TGC ATC TGA GAC ATA TTT GGC

L1 5′ UTR	
F	AGC CTA ACT GGG AGG CA
R	TTC CCC ATC TTT GTG GT

L1 ORF1	
F	GCA AGG CAG GCC AAC GTT CA
R	CCT TTC TCT CTG GCT GCC CT

L1 ORF2	
F	CAG GGC AAT CAG GCA GGA GA
R	TTG GGC TGA GAC GAT GGG GT

L1 reporter (pJJ101-L1; pL1Hs)	
F	CAC CCT AAC TGA CAC ACA TTC CAC
R	CAT GTC TGG ATC CGG CCT CCC

Hygromycin^{R}	
F	CCT GAA CTC ACC GCG ACG TC
R	GTC AAG CAC TTC CGG AAT CGG

GAPDH	
F	GCA AAT TCC ATG GCA CCG T
R	TCG CCC CAC TTG ATT TTG G

L1 3′ UTR	
F	TGGGTATATACCCAAATGAG
R	TGCGCTGCACCCACTAATGT

pYX014 Renilla	
F	CTT AGG CAG ATC GTC GCT GG
R	CTC CCA AGC AAG ATC ATG CGG

L1 reporter (pYX014; L1-luc)	
F	GGA TGA TCT GGT TGC CGA AG
R	CCT TCG TGA CTT CCC ATT TGC C

L1 LEAP	
F	GGA ATT CGA TGG GCA ATG TGC ACA TGT AC
R	GCG AGC ACA GAA TTA ATA CGA CT

MALAT1	
F	CAT TCG CTT AGT TGG TCT AC
R	TTC TAC CGT TTT TAG CTT C

cMYC	
F	CCT ACC CTC TCA ACG ACA GCA G
R	CTT GTT CCT CCT CAG AGT CGC TG

cJUN	
F	GAT AAT CCA GTC CAG CAA CGG
R	GTT CTG GCT GTG CAG TTC G

PSMB6	
F	CAA CCA CTG GGT CCT ACA TCG C
R	GTT CAA TGC TGT GGA AAC CGA GC

spike in (firefly luciferase)	
F	AAG GTT GTG GAT CTG GAT AC
R	GAT TGT TTA CAT AAC CGG AC

MDM2	
F	GGT TGA CTC AGC TTT TCC TCT TG
R	GGA AAA TGC ATG GTT TAA ATA GCC

Table 2. Oligonucleotides for the construction of MeRIP-seq cDNA library

5' adapter	5′-GUU CAG AGU UCU ACA GUC CGA CGA UCN NNN-3′
3' adapter	5′-rApp NN NNT GGA ATT CTC GGG TGC CAA GG/3ddC/-3′
Reverse transcription primer	5′-GCC TTG GCA CCC GAG AAT TCC A-3′
PCR forward primer (RP1)	5′-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACG TTC AGA GTT CTA CAG TCC GA-3′
PCR index reverse primer (RPI X)	5′-CAA GCA GAA GAC GGC ATA CGA GAT **NNN NNN** GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA-3′

RPI X	6 nt index
RPI1	CGTGAT
RPI10	AAGCTA
RPI11	GTAGCC
RPI12	TACAAG

4. RESULTS

4.1. METTL3 and ALKBH5 regulate L1 retrotransposition.

To determine whether RNA m^{6}A modification affects L1 retrotransposition, I evaluated the effects of the RNA m^{6}A machinery on L1 retrotransposition using a cell-based engineered L1-reporter assay (Moran et al., 1996). For the assay, I used the pJJ101 L1 dn6 2.2 construct (hereafter referred to as pL1Hs) that contains a blasticidin S deaminase gene (*mblastI*) within the 3′ UTR antisense to the SV40 promoter (Klawitter *et al.*, 2016; Morrish et al., 2002) (Figure 4). When L1 is successfully integrated into the host chromosome, the cells acquire resistance to blasticidin (Figure 4).

I depleted the m^{6}A methyltransferase METTL3, RNA demethylase ALKBH5, and FTO using small-interfering RNA (siRNAs) in HeLa cells and transfected pL1Hs vector. In METTL3-depleted cells, the number of blasticidin S-resistant colonies, which represent successful L1 retrotransposition, was reduced by >2-fold compared to that of control siRNA (Figure 5A). Conversely, the silencing of ALKBH5 increased L1 mobility, while the silencing of FTO did not affect L1 retrotransposition (Figure 5A). The depletion of the m^{6}A machinery did not vitiate cell viability (Figure 6A). In a reciprocal experiment, I performed an L1

retrotransposition assay with the ectopic expression of RNA m^6A demethylase ALKBH5 or FTO. Notably, the overexpression of ALKBH5 inhibited L1 mobility by ~4-fold, whereas FTO overexpression did not affect L1 mobility compared to that in AcGFP-expressing negative control cells (Figure 5B). I hypothesized that ALKBH5 may function as an L1 restriction factor by removing essential m^6A for L1 mobility. To examine whether the enzymatic function of ALKBH5 is critical for L1 mobility suppression, I performed L1 assays using the plasmid-encoding catalytically inactive mutant of ALKBH5 ($ALKBH5^{H204A}$). As anticipated, ALKBH5 could successfully restrained L1 mobility to levels that were comparable to that suppressed by a reverse transcription inhibitor (stavudine; d4T), whereas $ALKBH5^{H204A}$ overexpression did not result in the restriction of L1 mobility (Figure 5C). The viability of transfected cells remained unaffected (Figure 6B).

The pL1Hs plasmid encodes reporter L1 downstream of the CMV promoter and L1 5′ UTR promoter. Since the presence of the CMV promoter might affect L1-associated m^6A modification, I used a pYX014 L1-luciferase vector driven only by the L1 5′ UTR promoter. Using pYX014, the firefly luciferase reporter within the 3′ UTR allowed me to assess L1 mobility by measuring luminescence as previously reported (Xie *et al.*, 2011) (Figure 7A).

Overexpression of ALKBH5 impaired L1 retrotransposition, regardless of the presence of the CMV promoter (Figure 7B). In line with this result, depletion of METTL3 or ALKBH5 regulates L1 mobility, whereas FTO knockdown did not affect (Figure 7C). These results indicate that ALKBH5-specific m^6A substrates are necessary for L1 expansion. To summarize, my data support the functional role of the m^6A machinery in regulating L1 retrotransposition.

RNA m^6A metabolism regulates gene expression at post-transcriptional levels. Therefore, I speculated that the m^6A machinery would influence the protein expression of L1. Immunoblot analysis of HeLa cells devoid of m^6A enzymes revealed that m^6A enzymes regulate the expression of ORF1p (Figure 8A). Overexpression of ALKBH5 inhibited ORF1p expression, while the ectopic overexpression of FTO and $ALKBH5^{H204A}$ did not affect the ORF1p expression (Figure 8B, C). In each condition, the transfection efficiency of pL1Hs was not affected by siRNA or plasmids transfection (Figure 9A, B). Furthermore, neither the depletion of RNA m^6A machinery nor the overexpression of ALKBH5 altered the levels of expression of the control EGFP (Figure 9C, D), which indicates that m^6A enzymes do not affect transfection efficiency. These

results suggest that m^{6}A-mediated L1 regulation affects both retrotransposition and L1 protein expression.

Next, I have investigated whether de novo insertions of endogenous L1 are regulated by METTL3 or ALKBH5. For this experiments, I used PA-1 embryonic carcinoma cell line, which highly expresses endogenous L1. To measure de novo L1 integration, I generated METTL3 and ALKBH5-depleted PA-1 cell lines through shRNA lentiviral transduction (Figure 10A, B). Then I cultivated and cryopreserved these cells in passage 1, 6, and 12. These frozen cell stocks were used for genomic DNA (gDNA) extraction. Through gDNA-qPCR, I measured the amount of L1 DNA in gDNA using L1 ORF1-specific primers and 3′ UTR-specific primers. Remarkably, depletion of ALKBH5 increased the enrichment level of L1 3′ UTR, but not that of L1 ORF1 (Figure 11). This result can be attributed to the frequent 5′-truncation of the L1 cDNA during insertion11. However, METTL3 depletion, which decreases m6A modification, did not affect the enrichment of genomic L1 (Fig 11). This result may be due to the tendency of L1-regulating m^{6}A to remain demethylated in the basal status of PA-1 cells.

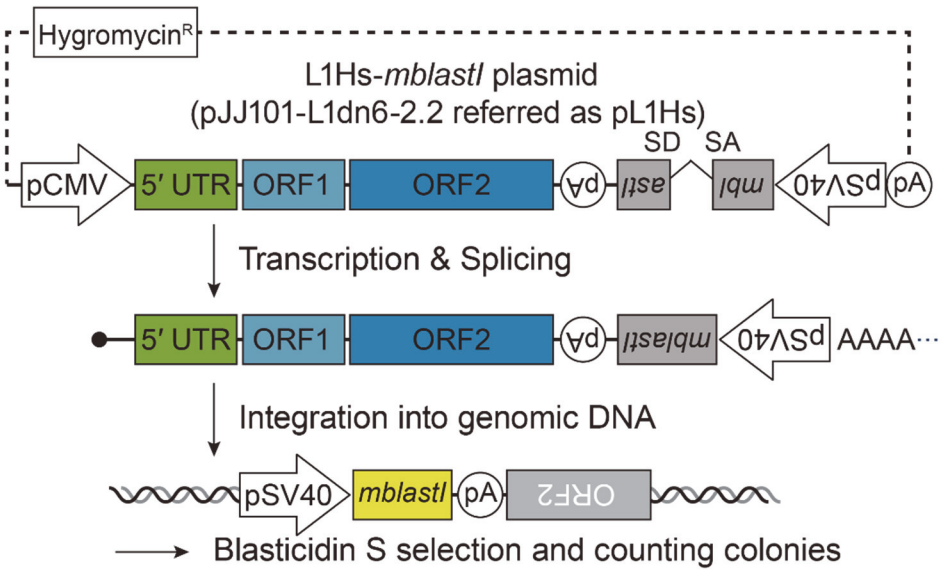

Hygromycin^R
L1Hs-*mblastI* plasmid
(pJJ101-L1dn6-2.2 referred as pL1Hs)
SD
SA
pCMV
5′ UTR
ORF1
ORF2
pA
pSV40
pA
Transcription & Splicing
5′ UTR
ORF1
ORF2
pA
pSV40
AAAA···
Integration into genomic DNA
pSV40
mblastI
pA
ORF2
Blasticidin S selection and counting colonies

Figure 4. A schematic of the L1 construct and an overview of the L1 retrotransposition assay using engineered human L1 construct

The pL1Hs expression cassette is a complete retrotransposition-competent L1 element that encodes L1 ORF1p and L1 ORF2p driven by a CMV promoter. The L1 construct carries a retrotransposition indicator cassette near its 3′-untranslated region (UTR). The cassette contains a blasticidin S deaminase gene (*mblastI*) interrupted by human gamma globin intron. During transcription, the intron is spliced out of the full-length L1 RNA transcript. The spliced L1 RNA is reverse-transcribed and the resulting cDNA is integrated into the genome. Retrotransposition of the resulting RNA leads to expression of the indicator gene, conferring blasticidin-resistance to host cells.

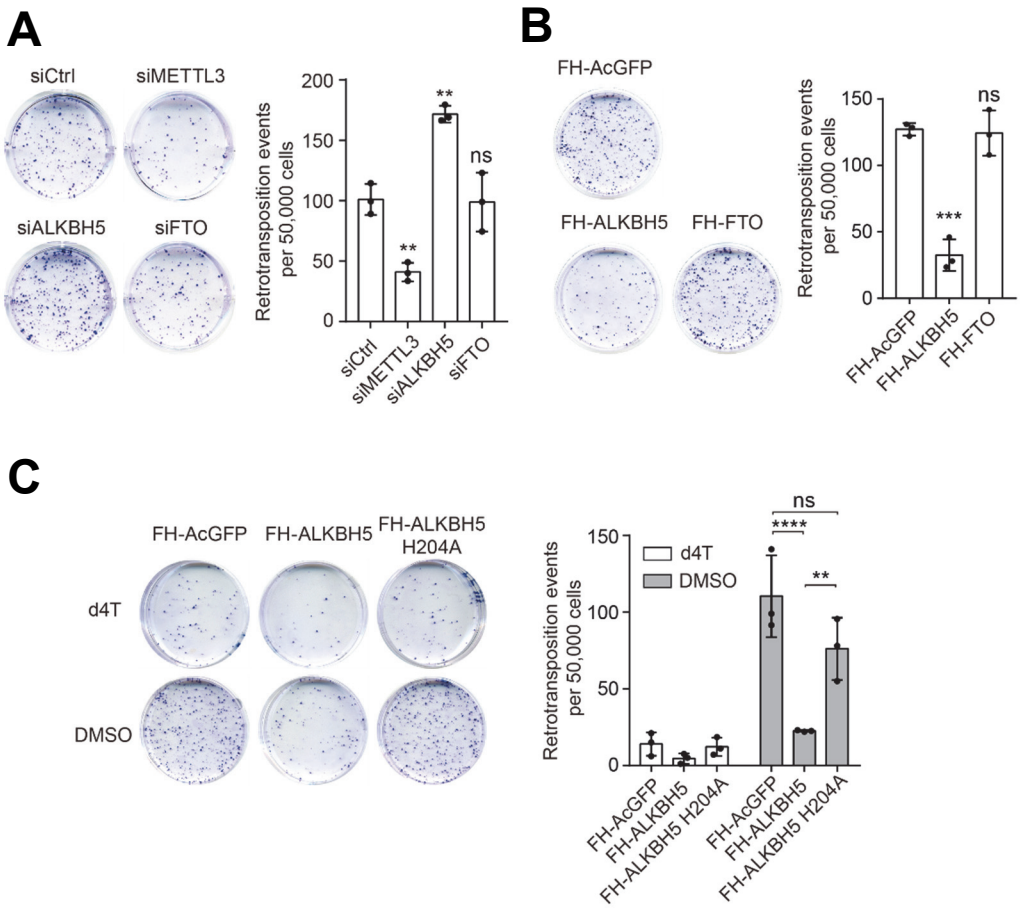

A
siCtrl
siMETTL3
siALKBH5
siFTO
Retrotransposition events per 50,000 cells
200
150
100
50
0
**
**
ns
siCtrl
siMETTL3
siALKBH5
siFTO
B
FH-AcGFP
FH-ALKBH5
FH-FTO
Retrotransposition events per 50,000 cells
150
100
50
0
***
ns
FH-AcGFP
FH-ALKBH5
FH-FTO
C
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A
d4T
DMSO
Retrotransposition events per 50,000 cells
150
100
50
0
d4T
DMSO
ns
****
**
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A

Figure 5. L1 retrotransposition assay in m^{6}A machinery-depleted or -overexpressed HeLa cells

(A) Retrotransposition assay in HeLa cells treated with siRNA that targets METTL3, ALKBH5, or FTO. A non-targeting siRNA (siCtrl) was used as a control. (B) Retrotransposition assays performed by co-transfecting the pL1Hs expression cassette with the indicated m^{6}A enzyme-expressing vectors into HeLa cells. (C) L1 retrotransposition assays were performed in ALKBH5, ALKBH5 catalytically inactive mutant (H204A), or AcGFP(control)-overexpressing cells. Cells treated with 50 μM stavudine (d4T) served as a reverse transcription negative control. (n = 3 independent samples, mean ± s.d., one-way ANOVA and Tukey's multiple comparisons test; ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, in comparison to control, ns: not significant).

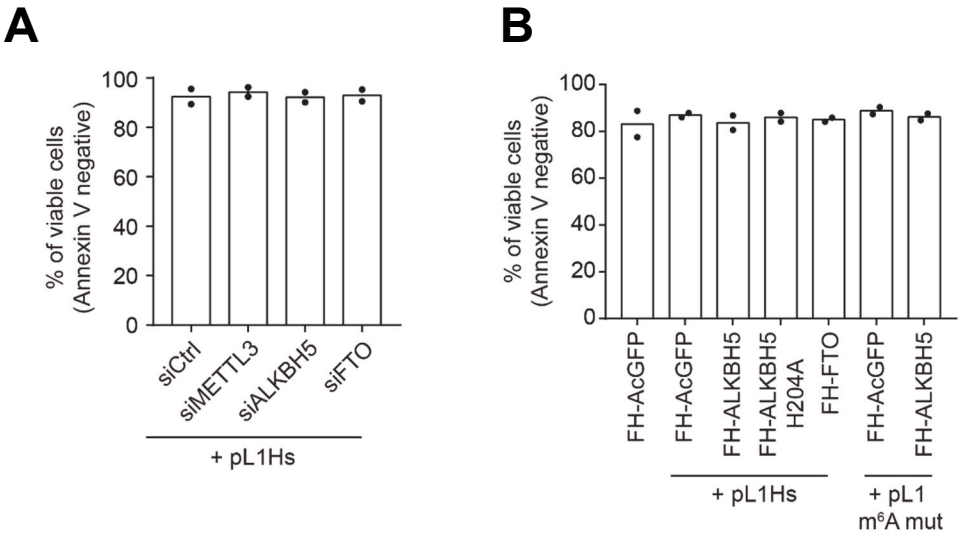

A
% of viable cells (Annexin V negative)
100
80
60
40
20
0
siCtrl
siMETTL3
siALKBH5
siFTO
+ pL1Hs
B
% of viable cells (Annexin V negative)
100
80
60
40
20
0
FH-AcGFP
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A
FH-FTO
FH-AcGFP
FH-ALKBH5
+ pL1Hs
+ pL1 m6A mut

Figure 6. Viability test of m^6A machinery-depleted or -overexpressed HeLa cells

(A) Annexin V assay of pL1Hs-expressing HeLa cells treated with siRNA that targets METTL3, ALKBH5, or FTO. (n = 2 independent samples, mean of two replicates) (B) Annexin V assay of pL1-expressing HeLa cells which are co-transfected with plasmids-encoding indicated genes. (n = 2 independent samples, mean of two replicates)

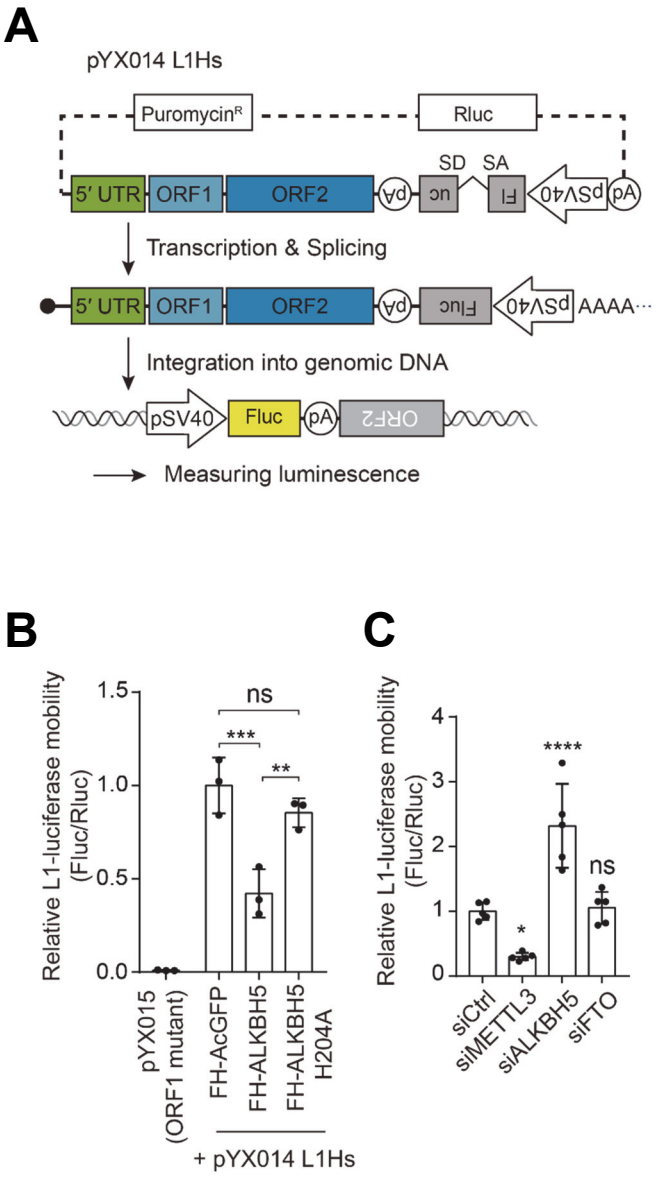

A
pYX014 L1Hs
Puromycin^R
Rluc
5′ UTR
ORF1
ORF2
pA
SD
SA
pSV40
Transcription & Splicing
AAAA···
Integration into genomic DNA
pSV40
Fluc
ORF2
Measuring luminescence
B
Relative L1-luciferase mobility (Fluc/Rluc)
1.5
1.0
0.5
0.0
ns
***
**
pYX015 (ORF1 mutant)
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A
+ pYX014 L1Hs
C
Relative L1-luciferase mobility (Fluc/Rluc)
4
3
2
1
0
****
*
ns
siCtrl
siMETTL3
siALKBH5
siFTO

Figure 7. METTL3 and ALKBH5 regulate L1 retrotransposition, regardless of the presence of the CMV promoter.

(A) A schematic of the L1-luciferase reporter construct (pYX014) used in this study. This construct drives L1-luciferase expression using its promoter in the 5′ UTR. Firefly luciferase acts as the L1 retrotransposition reporter, while Renilla luciferase is used for the normalization of transfection efficiency. (B) L1-luciferase retrotransposition assays performed in ALKBH5 or ALKBH5 H204A –overexpressing HeLa cells. In the dual-luciferase assay, the levels of firefly luciferase were measured using the luminescence of Renilla luciferase. The pYX015 L1 construct carries retrotransposition-defective ORF1 mutations. The ratio of the luminescence of firefly and Renilla luciferase (Fluc/Rluc) was normalized to that in AcGFP-overexpressing cells, which served as a control. (C) L1-luciferase retrotransposition assay using m^6A enzyme-depleted HeLa cells. L1 mobility assessment was performed as in (B). (three (B) or five (C) independent samples, mean ± s.d., one-way ANOVA with Tukey's (B) or Dunnett's multiple comparisons test (C); ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$ and *$p < 0.05$, in comparison to control, ns = not significant).

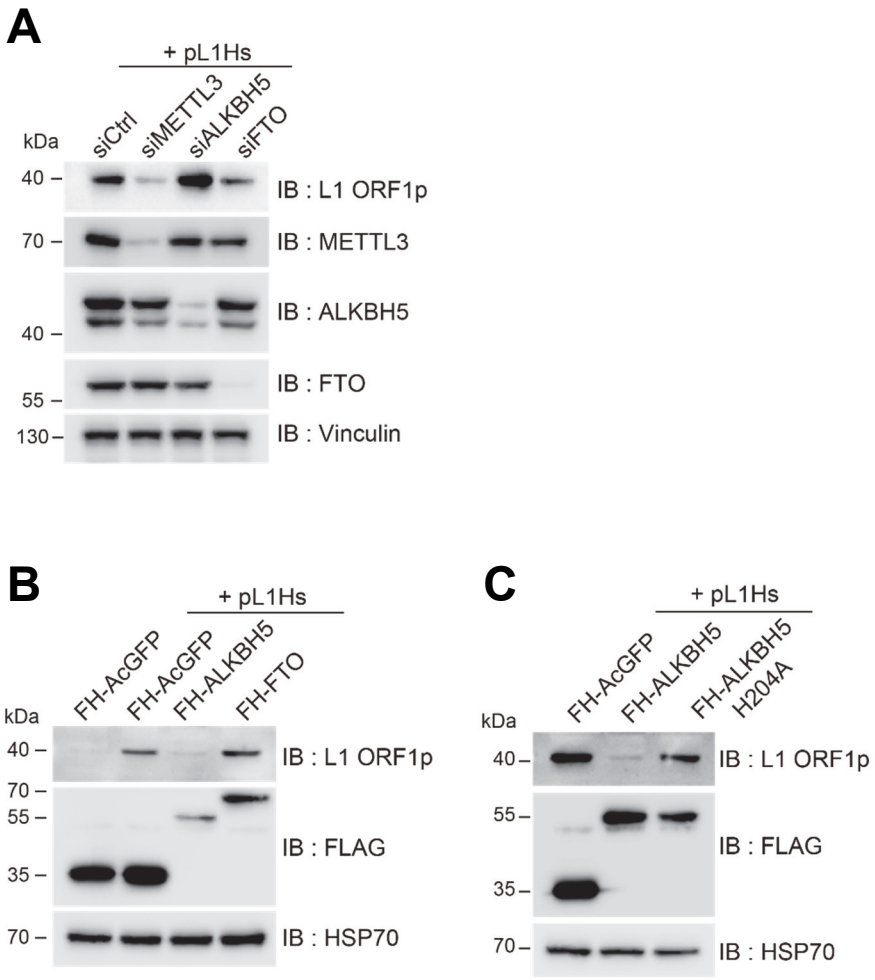

A
+ pL1Hs
siCtrl
siMETTL3
siALKBH5
siFTO
kDa
40 –
IB : L1 ORF1p
70 –
IB : METTL3
IB : ALKBH5
40 –
IB : FTO
55 –
130 –
IB : Vinculin
B
+ pL1Hs
FH-AcGFP
FH-AcGFP
FH-ALKBH5
FH-FTO
kDa
40 –
IB : L1 ORF1p
70 –
55 –
IB : FLAG
35 –
70 –
IB : HSP70
C
+ pL1Hs
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A
kDa
40 –
IB : L1 ORF1p
55 –
IB : FLAG
35 –
70 –
IB : HSP70

Figure 8. L1 ORF1p expression in m^{6}A machinery-depleted or -overexpressed HeLa cells.

(A) Immunoblot assay of lysates from pL1Hs-transfected HeLa cells treated with indicated siRNAs that target m^{6}A enzymes. Vinculin served as a loading control. (B and C) Immunoblot assay using pL1Hs-expressing HeLa cells. AcGFP, ALKBH5, FTO, or ALKBH5H204A overexpression plasmids were co-transfected with pL1Hs. FH-AcGFP served as transfection control. HSP70 served as a loading control. The predicted molecular weight of FLAG-HA tagged proteins are 34 kDa for FH-AcGFP, 51 kDa for FH-ALKBH5, and 65 kDa for FH-FTO. The immunoblot images (A-C) are representative of three independent experiments.

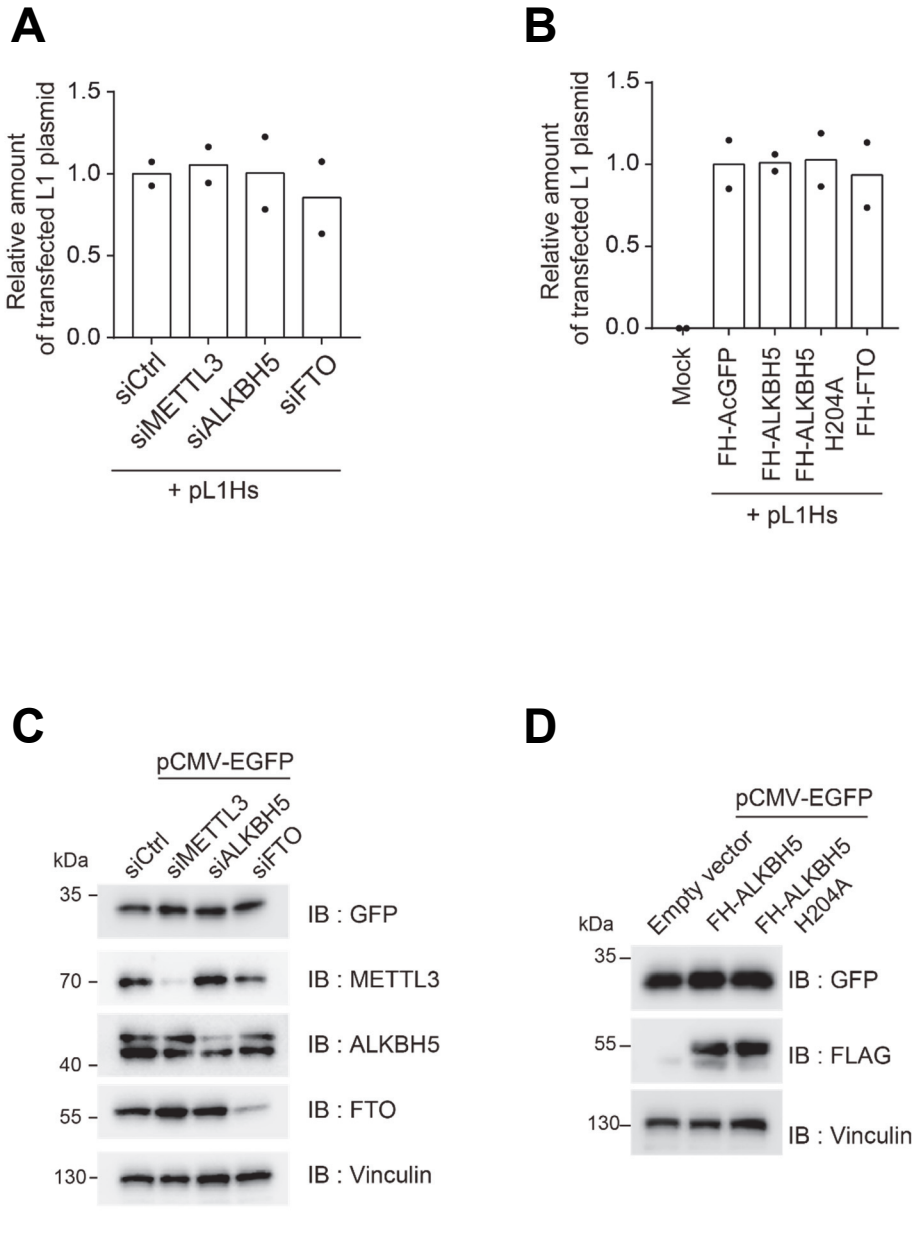
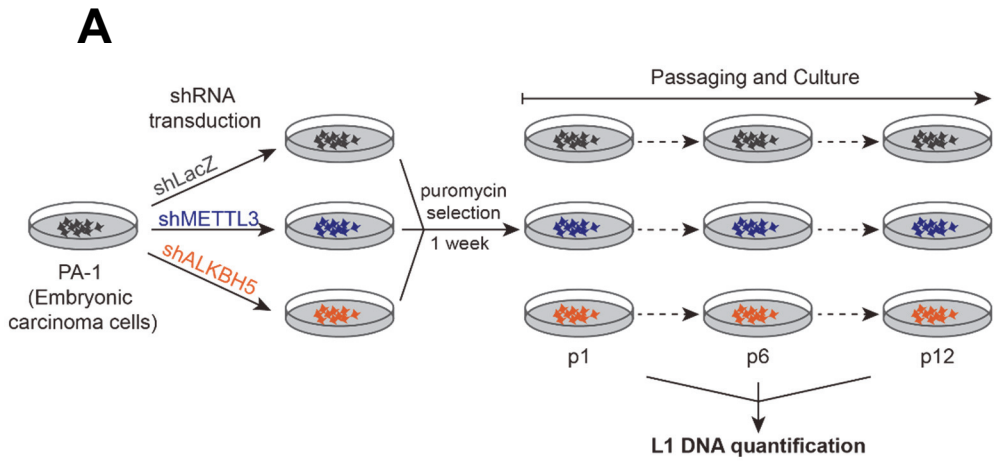

A
Relative amount of transfected L1 plasmid
1.5
1.0
0.5
0.0
siCtrl
siMETTL3
siALKBH5
siFTO
+ pL1Hs
B
Relative amount of transfected L1 plasmid
1.5
1.0
0.5
0.0
Mock
FH-AcGFP
FH-ALKBH5
FH-ALKBH5 H204A
FH-FTO
+ pL1Hs
C
pCMV-EGFP
kDa
siCtrl
siMETTL3
siALKBH5
siFTO
35
IB : GFP
70
IB : METTL3
40
IB : ALKBH5
55
IB : FTO
130
IB : Vinculin
D
pCMV-EGFP
kDa
Empty vector
FH-ALKBH5
FH-ALKBH5 H204A
35
IB : GFP
55
IB : FLAG
130
IB : Vinculin

Figure 9. Modulation of RNA m^{6}A enzymes expression did not affect transfection efficiency.

(A and B) Quantification of transfected pL1 amount in HeLa cells transfected with indicated siRNAs that target m^{6}A enzymes (A) or with plasmids-encoding indicated RNA demethylases (B). Extracted gDNA and L1 plasmids were analyzed by qPCR. The enrichment of transfected pL1 was normalized to the levels of MDM2. (n = 2 independent samples, mean of two replicates)

(C and D) EGFP expression test after indicated siRNA treatment (C) or co-transfection with indicated plasmids (D). EGFP expression was measured for determining the transfection efficiency of pCMV-EGFP and the translation of exogenous gene. The protein ladder is marked with the corresponding molecular weight. Vinculin served as a loading control. The immunoblot images (C and D) are representative of two independent experiments.


A
shRNA
transduction
shLacZ
shMETTL3
shALKBH5
PA-1
(Embryonic
carcinoma cells)
puromycin
selection
1 week
Passaging and Culture
p1
p6
p12
L1 DNA quantification

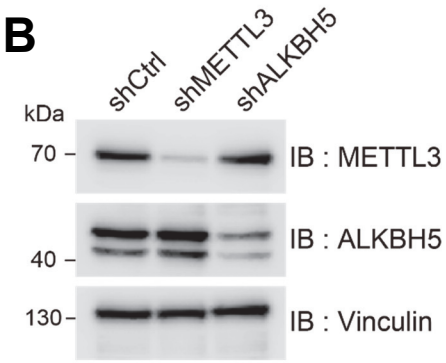

B
shCtrl
shMETTL3
shALKBH5
kDa
70 –
IB : METTL3
40 –
IB : ALKBH5
130 –
IB : Vinculin

Figure 10. Scheme of measuring endogenous L1 retrotransposition in m^{6}A machinery-depleted PA-1 cells

(A) Scheme of generation and cultivation of METTL3- or ALKBH5-depleted cells. (B) Immunoblot assay showing knockdown efficiency of target proteins in shRNA-transduced PA-1 cells. The protein ladder is marked with the corresponding molecular weight. Vinculin served as a loading control.

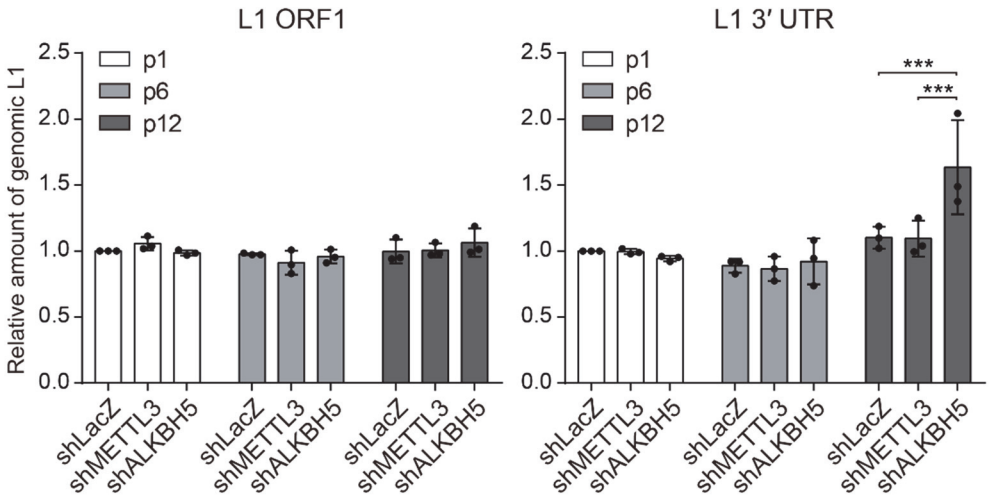

L1 ORF1
L1 3′ UTR
Relative amount of genomic L1
p1
p6
p12
2.5
2.0
1.5
1.0
0.5
0.0
***
***
shLacZ
shMETTL3
shALKBH5

Figure 11. ALKBH5 suppresses endogenous L1 retrotransposition in PA-1 cells.

Quantification of genomic L1s by qPCR using METTL3-or ALKBH5-depleted cells (right). Cell passaging numbers are indicated. Extracted gDNA was analyzed by qPCR using primers specific for L1 ORF1 and 3 ′UTR. The enrichment of genomic L1 was normalized to the levels of MDM2. (n = 3 independent samples, mean ± s.d., two-way ANOVA and Tukey's multiple comparisons test; ***$p < 0.001$)

4.2. L1 RNA is modified by m^{6}A.

Although the possibility of L1 m^{6}A modification was demonstrated in recent studies (Abakir et al., 2020; Liu et al., 2020), it remains unclear whether m^{6}A modification occurs in retrotransposition-competent full-length L1, and if so, which region of the L1 transcript is modified by m^{6}A. To validate whether m^{6}A modifies L1 RNA, I performed methyl-RNA immunoprecipitation (MeRIP) using human embryonic stem cells (H9 hESCs) that express endogenous L1 at sufficient levels (Garcia-Perez et al., 2007; Macia et al., 2011). Through qRT-PCR analysis of the MeRIP eluates, I detected the enrichment of L1 RNA at a level comparable to that for known m^{6}A-modified *SON* and *CREBBP* mRNA, but much more than negative control *HPRT1* mRNA (Figure 12A). To minimize bias resulting from primers in L1 RNA detection, I used three different primer sets that targeted the 5′ UTR, ORF1, and ORF2 regions and did not observe significant differences in the results obtained for these primers (Figure 12A). Similar to the endogenously expressed L1 RNA in hESCs, MeRIP analysis using the m^{6}A antibody clearly demonstrated that the L1 RNA exogenously expressed in HeLa cells undergoes m^{6}A modification (Figure 12B). I then evaluated if the silencing of METTL3 or ALKBH5 would alter the extent of m^{6}A modification of the L1 RNA. Indeed, MeRIP-qPCR with METTL3-

depleted cells revealed lower enrichment of the m^{6}A-modified L1 than of siCtrl-treated cells, whereas ALKBH5 knockdown augmented the levels of m^{6}A-positive L1 (Figure 12C). These results indicate that METTL3 can install m^{6}A modification in L1 transcripts, while ALKBH5 plays a role in removing the modification.

To examine the m^{6}A-modified regions in the L1 transcripts, I analyzed the m^{6}A transcriptome of hESCs reported previously (Batista *et al.*, 2014) and mapped reads to the consensus sequence of L1Hs, the youngest L1 (Khan *et al.*, 2006). I identified 18 peaks across the L1Hs sequence using two biological replicates (Figure 13). Given that the reads from L1s may yield false-positive results, I narrowed down and selected the peaks that are likely to contain m^{6}A motifs from 18 peaks through the m^{6}A prediction score algorithm (SRAMP) (Zhou et al., 2016). SRAMP analysis revealed that the 9 peaks found in the ORF1, ORF2 and 3′ UTR regions do not contain m^{6}A motifs, and that only the 6 peaks located at 5′ UTR have potential m^{6}A motifs (Table 3A).

Next, I mapped the sites of m^{6}A modifications in reporter L1-transfected HeLa cells using MeRIP-seq. Consistent with findings from previous studies (Dominissini *et al.*, 2012; Meyer *et al.*, 2012), my results indicated that the transcriptome-wide distribution of m^{6}A peaks were preferentially found in 3′

UTR and CDS, but not in the 5′ UTR (Figure 14A). By mapping the reads on reporter L1, I obtained five candidate peaks (Figure 14B), and further sorted according to the approach based on the m^6A prediction as described above. All five peaks were classified as m^6A-putative regions with high scores (Table 3B). Two of the featured peaks were located in the 5′ UTR, the other two were located in the ORF1, and another was in ORF2. The m^6A modification sites commonly detected in endogenous and exogenous L1 RNA are A332 and A839, both located in the 5′ UTR (Table 3). This is a notable phenomenon since m^6A modification typically occurs near the stop codon and at the 3′ UTR, and this gives rise to the possibility that the L1 5′ UTR acts as the regulatory hub for L1 mobility via m^6A modification.

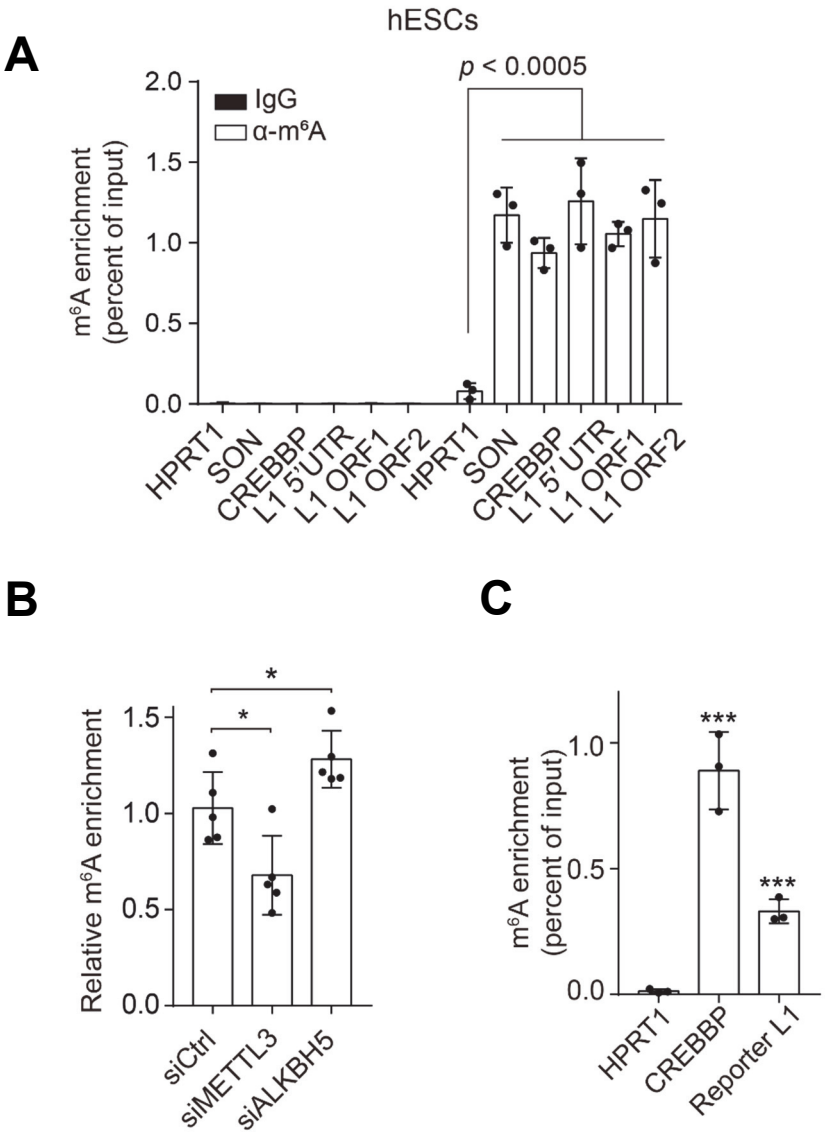

A
hESCs
p < 0.0005
IgG
α-m⁶A
m⁶A enrichment (percent of input)
2.0
1.5
1.0
0.5
0.0
HPRT1
SON
CREBBP
L1 5'UTR
L1 ORF1
L1 ORF2
HPRT1
SON
CREBBP
L1 5'UTR
L1 ORF1
L1 ORF2
B
Relative m⁶A enrichment
1.5
1.0
0.5
0.0
*
*
siCtrl
siMETTL3
siALKBH5
C
m⁶A enrichment (percent of input)
1.0
0.5
0.0
***
***
HPRT1
CREBBP
Reporter L1

Figure 12. L1 RNA is modified by m^{6}A

(A) MeRIP-qPCR analysis of mRNA from H9 hESCs. Eluates from IgG immunoprecipitation served as negative control. Eluted RNA was quantified to determine the percentage of input. (n = 3 independent samples, mean ± s.e.m., one-way ANOVA and Dunnett's multiple comparisons test). (B) MeRIP-qPCR analysis of pL1Hs-transfected HeLa cells with m^{6}A machinery knockdown. Eluted RNAs were quantified using primers specific for reporter L1. The enrichment of RNA was normalized to that of the control. (n = 5 independent samples, mean ± s.e.m., unpaired two-tailed t test, *p < 0.05). (C) MeRIP-qPCR assay of pL1Hs-expressing HeLa cells. The enrichment of the m^{6}A antibody-bound RNA was calculated as a percentage of the input. (n = 3 independent samples, mean ± s.d., unpaired two-tailed t test; ***p < 0.001).

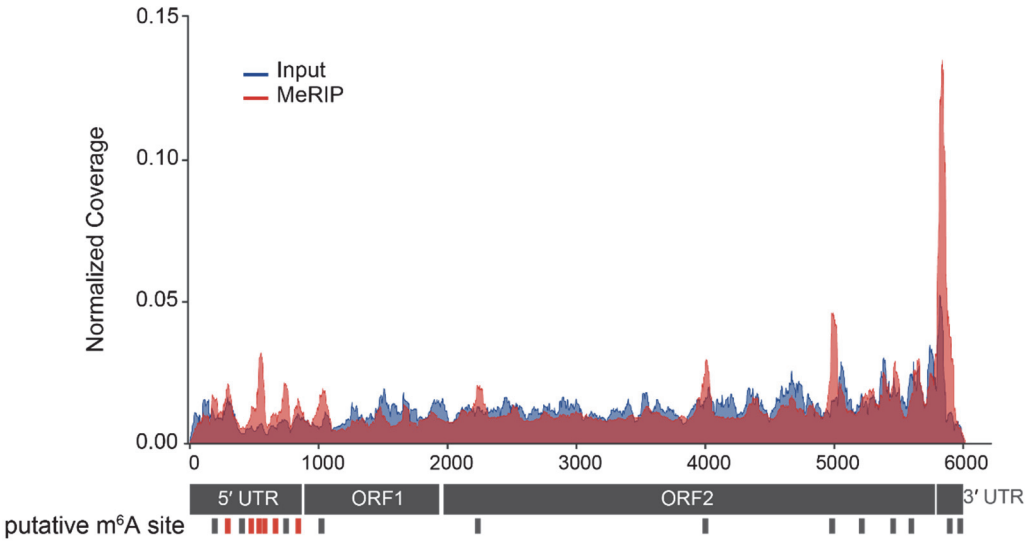
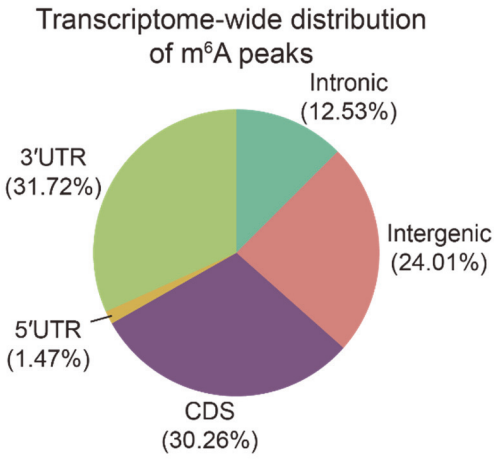

0.15
0.10
0.05
0.00
Normalized Coverage
Input
MeRIP
0
1000
2000
3000
4000
5000
6000
5′ UTR
ORF1
ORF2
3′ UTR
putative m6A site

Figure 13. m^{6}A distribution of L1 RNA from human embryonic stem cells

Map of m^{6}A modification sites in full-length L1Hs from previously reported MeRIP-seq data for H1 hESCs (GSE52600). Read coverage was normalized to the total number of reads mapped to the L1Hs consensus sequence. The plot presents data from MeRIP-seq in red and input RNA-seq in blue. Bars (in red or black) indicate the m^{6}A peaks identified by manual inspection in two replicates. m^{6}A peaks in red correspond to peaks containing high score m^{6}A-prediction sites.

A


Transcriptome-wide distribution
of m6A peaks
Intronic
(12.53%)
3′UTR
(31.72%)
Intergenic
(24.01%)
5′UTR
(1.47%)
CDS
(30.26%)

B

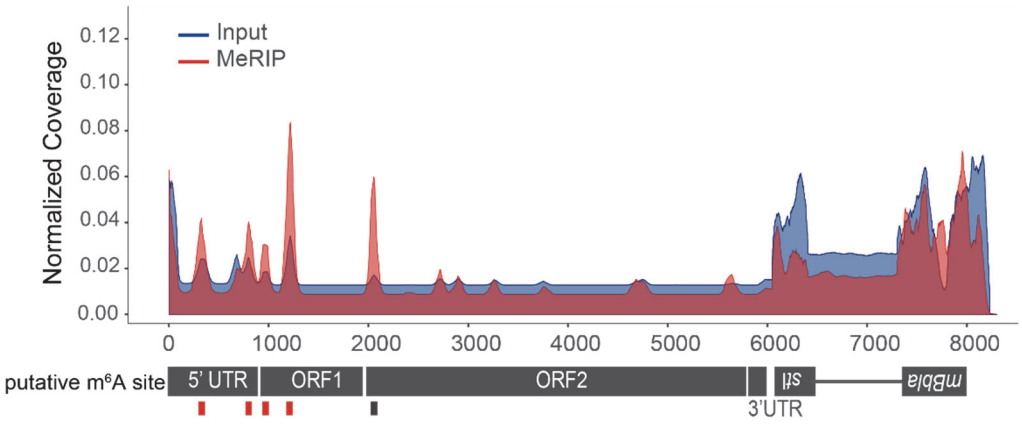

Input
MeRIP
Normalized Coverage
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
putative m6A site
5' UTR
ORF1
ORF2
sil
mBbla
3'UTR

Figure 14. m^{6}A distribution of L1 RNA from pL1-expressing HeLa cells

(A) Transcriptome-wide distribution of m^{6}A peaks from MeRIP-seq data of pL1Hs-expressing HeLa cells. Pie charts indicate the percentage of m^{6}A peaks in the marked region. m^{6}A is highly enriched in 3′ UTR and CDS compared to the distribution of reads in the input samples. (B) Methylation peaks in the full-length pL1Hs containing a retrotransposition reporter. The pL1Hs construct is presented below. *mblastI*, the reporter, is antisense and the gamma-globin intron is inserted in it, as described in Figure 4. Fragmented poly(A) RNA from pL1Hs-expressing HeLa cells was subject to RNA-seq and MeRIP-seq, and analyzed as described previously1. The plot in blue corresponds to the mapping distribution of the input RNA-seq data, while the plot in red corresponds to the MeRIP-seq data.

A

hESC MeRIP-seq

#	m^{6}A peak location	Feature	SRAMP prediction score	
			high score position	moderate score position
1	171-241	5'UTR	none	none
2	266-336	5'UTR	332	none
3	379-449	5'UTR	none	410
4	459-529	5'UTR	495	none
5	532-602	5'UTR	569	581
6	570-640	5'UTR	600	none
7	631-701	5'UTR	679	703
8	713-783	5'UTR	none	758
9	806-876	5'UTR	839	none
10	974-1044	ORF1	none	none
11	2228-2298	ORF2	none	2234
12	3972-4042	ORF2	none	none
13	4954-5024	ORF2	none	none
14	5225-5295	ORF2	none	none
15	5437-5507	ORF2	none	none
16	5601-5671	ORF2	none	none
17	5829-5899	3'UTR	none	none
18	5983-6018	3'UTR	none	none

B

pL1Hs-transfected HeLa MeRIP-seq

#	m^{6}A peak location	Feature	SRAMP prediction score	
			high score position	moderate score position
1	293-363	5'UTR	332	none
2	800-870	5'UTR	839	758
3	899-969	ORF1	931	none
4	1206-1276	ORF1	1203	none
5	1998-2068	ORF2	2037, 2064	none

Table 3. Identification of putative m^{6}A peaks through MeRIP-seq

(A) Identification of endogenous L1 m^{6}A peaks by hESC MeRIP-seq. The peaks were identified by manual inspection. The ratio of read coverage (MeRIP/input) from two replicates was measured. The sections with average ratio > 1.3 were selected, and the ± 35 nt regions from the maximum value were assumed to represent m^{6}A peaks. The 18 identified peaks and their respective features are indicated. The SRAMP m^{6}A tool was used to detect putative m^{6}A sites in the peaks. Both high and moderate score positions in the corresponding peaks were noted. Peaks with high score prediction sites were marked in apricot. (B) Identification of m^{6}A peaks in reporter L1 through MeRIP-seq of pL1Hs-expressing HeLa. Peaks were identified as in (A), although the cut-off ratio was > 1.4. SRAMP was also conducted to determine the putative m^{6}A regions, as in (A). Peaks in apricot indicate high score prediction.

4.3. 5′ UTR m^{6}A cluster is critical for L1 activity.

Given that the L1 5′ UTR has a potential m^{6}A cluster, I next examined whether the L1 5′ UTR is necessary for m^{6}A-dependent L1 regulation. I transfected the 5′ UTR-deleted pL1Hs (pL1Hs Δ5′ UTR) into HeLa cells treated with m^{6}A machinery-targeting siRNA and monitored ORF1p expression. Intriguingly, the knockdown of m^{6}A enzymes did not affect ORF1p expression in the absence of 5′ UTR (Figure 15A). Furthermore, using the codon-optimized synthetic L1 construct that encoded ORFs with the same amino acids yet different nucleotide sequences, I examined whether alterations in ORF1 and ORF2 nucleotide sequences could affect m^{6}A machinery-mediated L1 regulation. Remarkably, silencing of METTL3 or ALKBH5 regulates L1 ORF1p expression only when 5′ UTR is contained in synthetic L1, which indicates that m^{6}A machinery regulates L1 expression in a 5′ UTR-dependent manner (Figure 15B). These results suggest that the L1 5′ UTR contains functional m^{6}A motifs for successful ORF1p expression.

To identify the site of functional m^{6}A in L1 5′ UTR, I selected six adenosine candidates of m^{6}A modification (332, 495, 569, 600, 679, and 839, numbering based on L1PA1 consensus sequence (Khan *et al.*, 2006)) through MeRIP-seq analysis in either hESCs or L1 reporter-expressing HeLa cells (Table 3). I

generated a set of firefly luciferase reporter plasmids encoding L1 5′ UTR or its m^6A-silencing A to T mutants (Figure 16A). To quantify the effect of L1 5′ UTR m^6A mutation without the bias from transfection efficiency, I normalized the firefly luciferase activity to that of *Renilla* luciferase. The dual-luciferase reporter assay revealed that a single A332T, A495T, or A600T mutation reduced the expression of firefly luciferase, compared to that of native 5′ UTR (Figure 16B). However, the weak effect of these single mutants led me to hypothesize that multiple m^6A modifications may function synergistically. Indeed, the double mutation of A332/600T and the triple mutation of A332/495/600T exerted significantly more synergistic and potent effects (Figure 16C).

I next performed the L1 retrotransposition assay using the 5′ UTR m^6A mutants of the pL1Hs construct. Mutations at each m^6A motif of A332, A495, and A600 showed a marginal effect on L1 retrotransposition, whereas A332/A495/A600 triple mutation (hereinafter referred to as pL1 m^6A mut) markedly inhibited L1 mobility (Figure 17A, B, C). I validated the effect of the m^6A cluster using the L1-luciferase reporter construct pYX014. Indeed, the triple m^6A mutant of the L1-luciferase construct (pYX014 L1 m^6A mut) induced approximately 50% decline in L1 mobility compared to that induced by the wild-type L1 (Figure 17D).

To assess the effect of the triple mutation in the m^{6}A modification level of L1, I performed MeRIP-qPCR for comparing m^{6}A enrichments between cells that expressed pL1Hs and pL1 m^{6}A mut. Surprisingly, the triple mutation reduced the enrichment of m^{6}A-modified L1 by approximately 50%, while it did not affect the m^{6}A levels of the endogenous controls *SON* and *CREBBP* (Figure 18A, B). These results indicate that A332, A495, and A600 are the essential adenosines for L1 mobility and serve as m^{6}A modification sites.

Based on my finding that ALKBH5 inhibits L1 mobility, I attempted to determine whether ALKBH5 could restrict the mobility of the L1 m^{6}A mutant. L1 assays with co-transfection of pL1 vectors and FH-ALKBH5 revealed that the ectopic expression of ALKBH5 impaired the retrotransposition of pL1Hs (Figure 19A and Figure 6B). However, ALKBH5 overexpression caused only marginal effects in pL1 m^{6}A mut-expressing cells (Figure 19A). Moreover, silencing the triple m^{6}A modification led to the suppression of L1 mobility in AcGFP-expressing cells, but not in ALKBH5-expressing cells (Figure 19A). In a reciprocal experiment, I measured the L1 retrotransposition frequency of pL1Hs and pL1 m^{6}A mut in ALKBH5-lacking cells. Notably, ALKBH5 knockdown led to the enhancement of L1 mobility in pL1Hs-expressing cells, whereas no measurable changes were observed in pL1 m^{6}A mut-expressing

cells (Figure 19B, C). Consistent with this result, ALKBH5 was not able to suppress L1 ORF1p expression in the absence of the m^{6}A cluster (Figure 20A, B). In summary, I demonstrated that ALKBH5 suppresses L1 expression in the 5′ UTR m^{6}A cluster-dependent manner, which suggests that the L1 5′ UTR m^{6}As serve as the substrates for ALKBH5 demethylation.

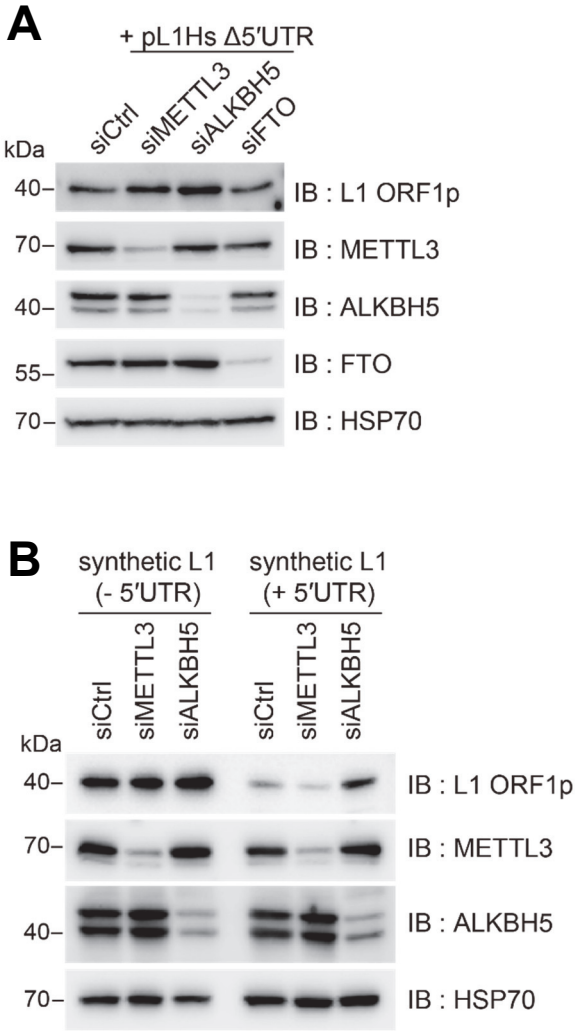
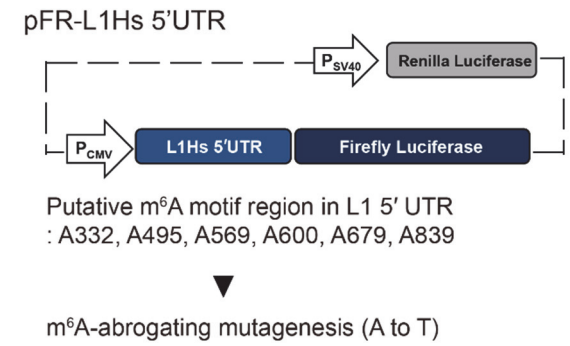

A
+ pL1Hs Δ5′UTR
siCtrl
siMETTL3
siALKBH5
siFTO
kDa
40–
IB : L1 ORF1p
70–
IB : METTL3
40–
IB : ALKBH5
55–
IB : FTO
70–
IB : HSP70
B
synthetic L1
(- 5′UTR)
synthetic L1
(+ 5′UTR)
siCtrl
siMETTL3
siALKBH5
siCtrl
siMETTL3
siALKBH5
kDa
40–
IB : L1 ORF1p
70–
IB : METTL3
40–
IB : ALKBH5
70–
IB : HSP70

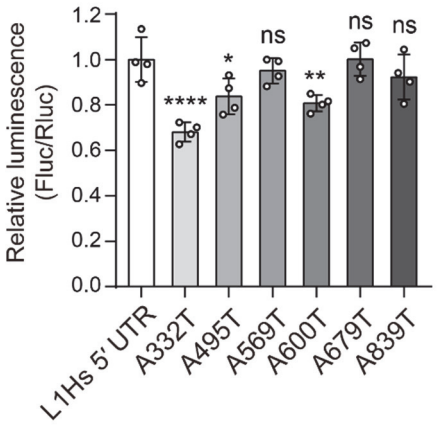
Figure 15. L1 5′ UTR regulates L1 protein expression.

(A) Immunoblot assay of 5′ UTR deletion L1 construct (pL1Hs Δ5′ UTR) in HeLa cells treated with indicated siRNAs. (B) Immunoblot assay depicting the effect of 5′ UTR using synthetic L1 constructs (L1-neo-TET). siRNA and plasmids transfection were performed as in (A). HSP70 served as a loading control. The immunoblot images (A and B) are representative of two independent experiments.

A


pFR-L1Hs 5'UTR
P_{SV40}
Renilla Luciferase
P_{CMV}
L1Hs 5'UTR
Firefly Luciferase
Putative m⁶A motif region in L1 5′ UTR
: A332, A495, A569, A600, A679, A839
m⁶A-abrogating mutagenesis (A to T)

B


Relative luminescence (Fluc/Rluc)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
****
*
ns
**
ns
ns
L1Hs 5′ UTR
A332T
A495T
A569T
A600T
A679T
A839T

C

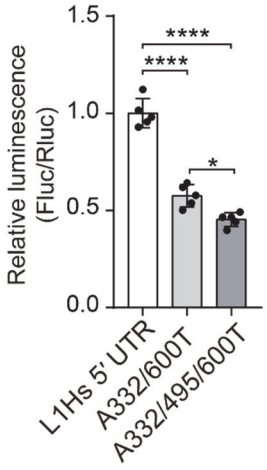

Relative luminescence (Fluc/Rluc)
1.5
1.0
0.5
0.0
****
****
*
L1Hs 5′ UTR
A332/600T
A332/495/600T

Figure 16. Identification of L1 5′ UTR m^{6}A modification sites

(A) Schematic of the dual-luciferase plasmid carrying L1 5′ UTR upstream of the firefly luciferase gene (pFR-L1Hs 5′ UTR). Firefly luciferase luminescence reflected the effect of 5′ UTR and of its mutations. (B and C) Dual-luciferase assay using HeLa cells transfected with pFR-L1Hs 5′ UTR or its A to T m^{6}A-abrogating mutant. The ratio of the luminescence of firefly and Renilla luciferase (Fluc/Rluc) was normalized to pFR-L1Hs 5′ UTR-expressing cells (mean ± s.d., four (B) or five (C) independent samples). Statistical significance was calculated by one-way ANOVA with Dunnett's (B), Tukey's multiple comparisons test (C) (****$p < 0.0001$, and **$p < 0.01$, *$p < 0.05$, ns: not significant).

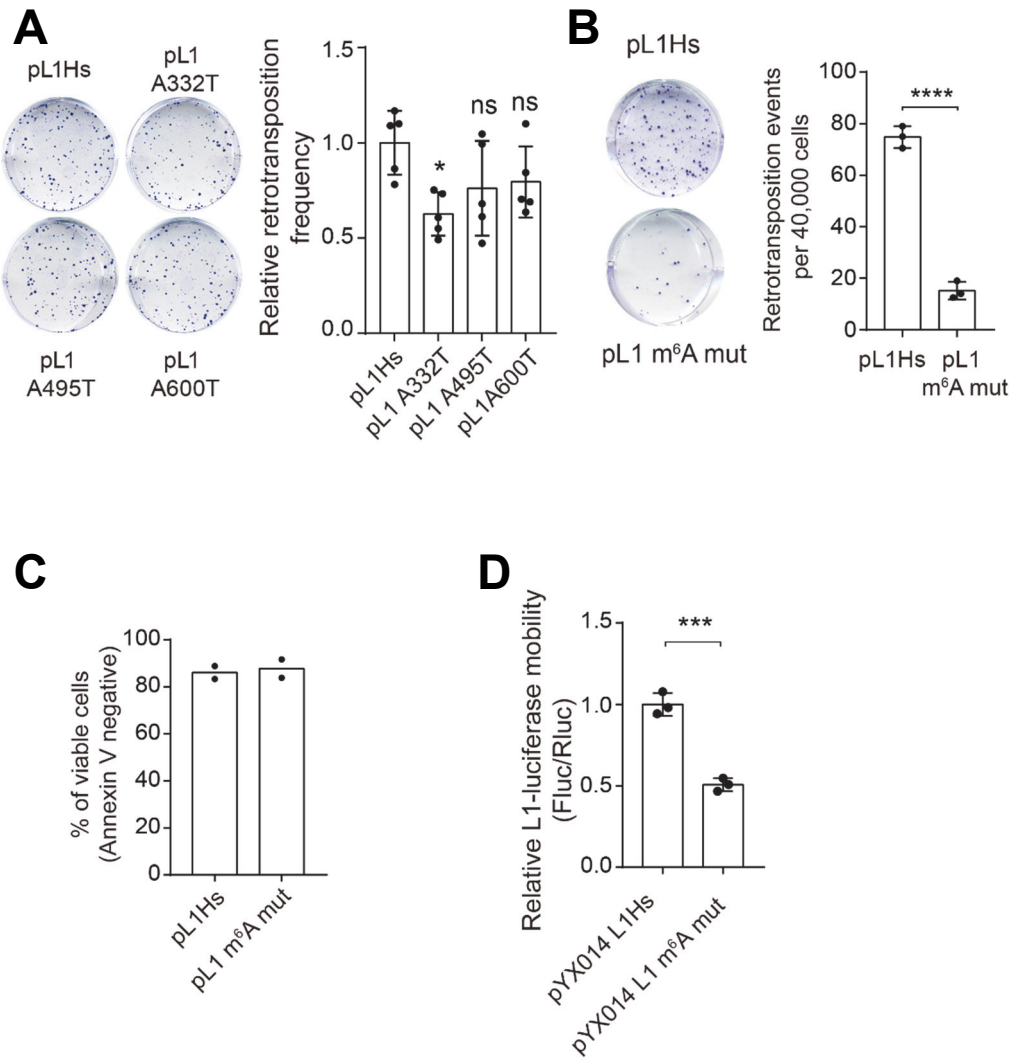

A
pL1Hs
pL1 A332T
pL1 A495T
pL1 A600T
Relative retrotransposition frequency
1.5
1.0
0.5
0.0
*
ns
ns
pL1Hs
pL1 A332T
pL1 A495T
pL1A600T
B
pL1Hs
pL1 m6A mut
Retrotransposition events per 40,000 cells
100
80
60
40
20
0
****
pL1Hs
pL1 m6A mut
C
% of viable cells (Annexin V negative)
100
80
60
40
20
0
pL1Hs
pL1 m6A mut
D
Relative L1-luciferase mobility (Fluc/Rluc)
1.5
1.0
0.5
0.0
***
pYX014 L1Hs
pYX014 L1 m6A mut

Figure 17. L1 5′ UTR m^{6}A cluster promotes L1 activity.

(A) Retrotransposition assay with a pL1Hs construct with a single m^{6}A mutation. The retrotransposition frequency was normalized to that of pL1Hs. Representative images of retrotransposition-positive HeLa foci are shown to the right of the bar graph (n = 5 independent samples, mean ± s.d., one-way ANOVA and Dunnett's multiple comparisons test; *p < 0.05, in comparison to control, ns = not significant). (B) L1 assays using the triple m^{6}A mutated L1 construct (pL1 m^{6}A mut) in HeLa cells (n = 3 independent samples, mean ± s.d., unpaired two-tailed t test; ****p < 0.0001). (C) Viability test of pL1Hs- or pL1 m^{6}A mut-expressing HeLa cells through Annexin V assay (n = 2 independent samples, mean of two replicates). (D) L1-luciferase retrotransposition assay using pYX014 L1Hs- or its m^{6}A mut. Ratios of luminescence (Fluc/Rluc) were normalized to those of pYX014 L1Hs. (n = 3 independent samples, mean ± s.d., unpaired two-tailed t test; ***p < 0.001)

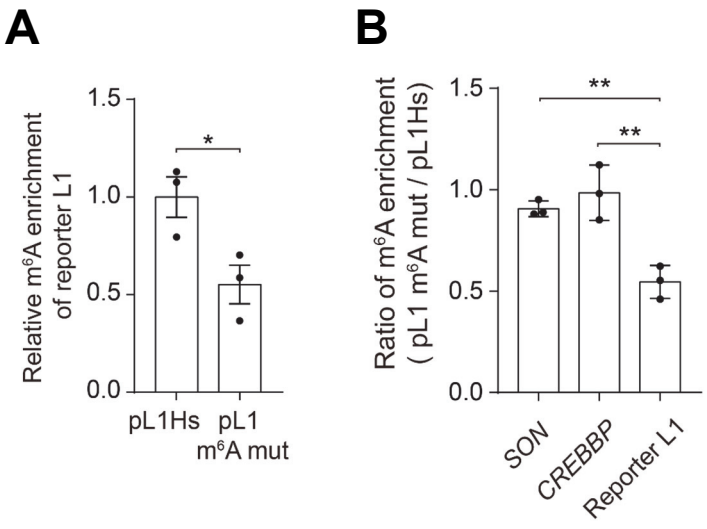

A
B
Relative m6A enrichment of reporter L1
1.5
1.0
0.5
0.0
*
pL1Hs
pL1 m6A mut
Ratio of m6A enrichment (pL1 m6A mut / pL1Hs)
**
**
SON
CREBBP
Reporter L1

Figure 18. The triple m^{6}A site mutations in L1 5 ′UTR reduce the extent of m^{6}A-modified L1 RNA

(A) MeRIP-qPCR analysis for evaluating the effect of the triple m^{6}A mutation construct (pL1 m^{6}A mut). m^{6}A antibody-bound L1 RNA was normalized to that of pL1Hs-transfected cells. (n = 3 independent samples, mean ± s.e.m., unpaired two-tailed t test; *p < 0.05) (B) The ratio of m^{6}A enrichment was calculated using MeRIP-qPCR with pL1Hs- or pL1m^{6}A mut-expressing HeLa cells. SON and CREBBP served as a transfection-independent control. (n = 3 independent samples, mean ± s.d., one-way ANOVA and Tukey's multiple comparisons test; **p < 0.01, in comparison to reporter L1)

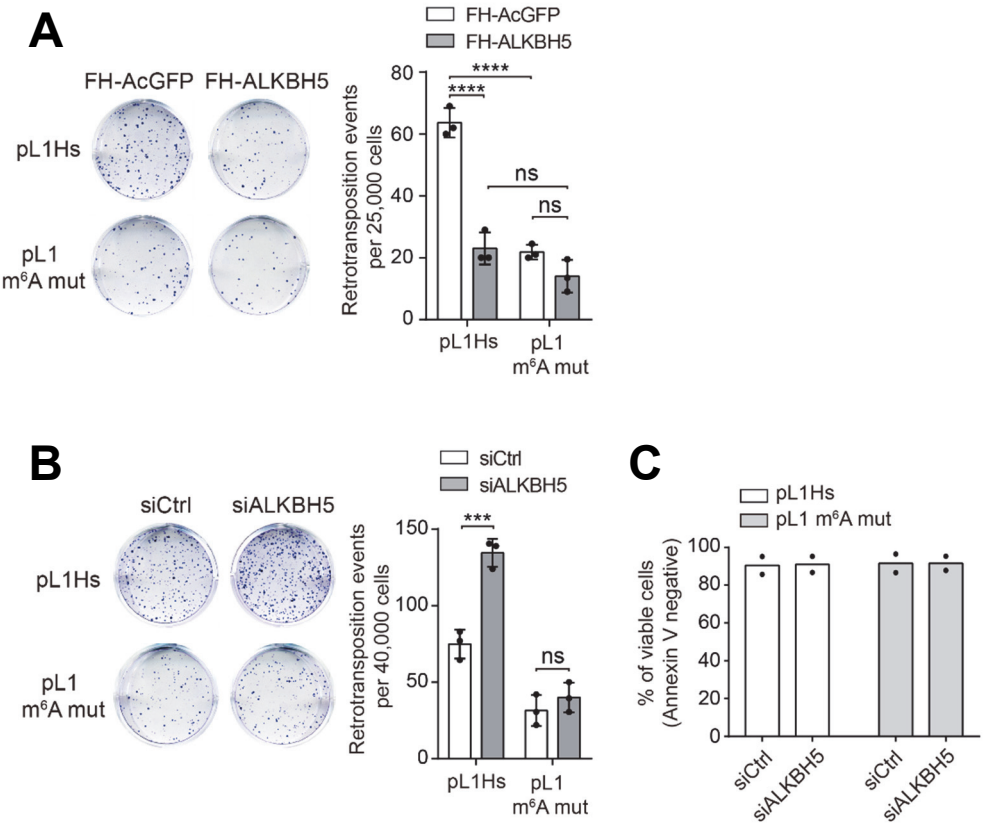

A
FH-AcGFP
FH-ALKBH5
FH-AcGFP FH-ALKBH5
pL1Hs
pL1
m6A mut
Retrotransposition events
per 25,000 cells
80
60
40
20
0
****
****
ns
ns
pL1Hs
pL1
m6A mut
B
siCtrl
siALKBH5
siCtrl siALKBH5
pL1Hs
pL1
m6A mut
Retrotransposition events
per 40,000 cells
150
100
50
0
***
ns
pL1Hs
pL1
m6A mut
C
pL1Hs
pL1 m6A mut
% of viable cells
(Annexin V negative)
100
80
60
40
20
0
siCtrl
siALKBH5
siCtrl
siALKBH5

Figure 19. m^{6}A cluster of L1 5′ UTR is a substrate for ALKBH5

(A and B) Retrotransposition assay using pL1Hs-or pL1 m^{6}A mut-expressing HeLa cells ALKBH5 overexpression (A) or silencing (B) (n = 3 independent samples, mean ± s.d., Tukey's multiple comparisons test; ****p < 0.0001, ***p < 0.001, ns: not significant). (C) Annexin V viability test of pL1Hs- or pL1 m^{6}A mut-expressing HeLa cells which is treated with indicated siRNAs. (n = 2 independent samples, mean of two replicates)

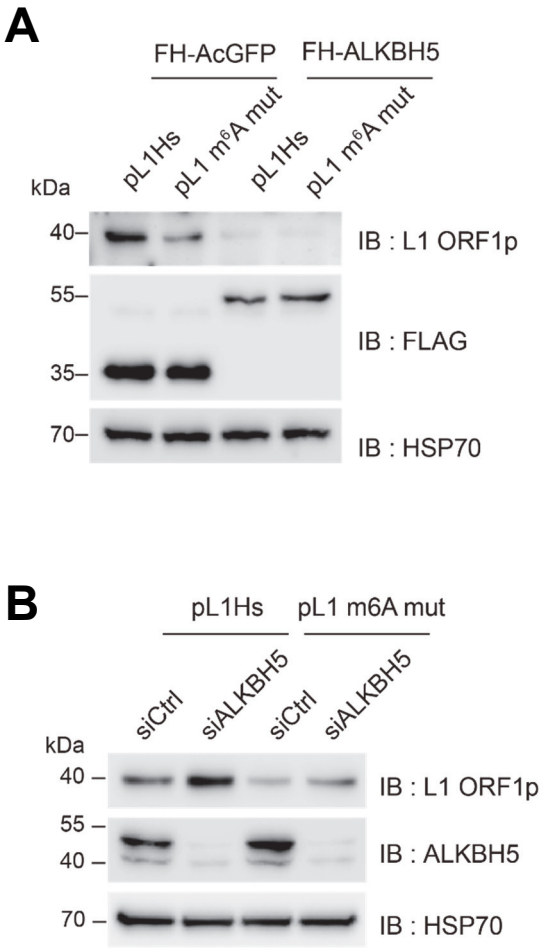

A
FH-AcGFP
FH-ALKBH5
pL1Hs
pL1 m6A mut
pL1Hs
pL1 m6A mut
kDa
40–
55–
35–
70–
IB : L1 ORF1p
IB : FLAG
IB : HSP70
B
pL1Hs
pL1 m6A mut
siCtrl
siALKBH5
siCtrl
siALKBH5
kDa
40 –
55 –
40 –
70 –
IB : L1 ORF1p
IB : ALKBH5
IB : HSP70

Figure 20. ALKBH5 suppresses L1 ORF1p expression by targeting L1 5 ′UTR m^{6}A cluster.

(A) Immunoblot assay for ORF1p quantification using HeLa cells co-transfected with pL1 construct and AcGFP- or ALKBH5-encoding plasmids. (The predicted molecular weight; 34 kDa for FH-AcGFP and 51 kDa for FH-ALKBH5) (B) Immunoblot assay depicting L1 ORF1p expression of indicated pL1-expressing HeLa cells with knockdown of ALKBH5. The non-targeting siRNA siCtrl served as a control. HSP70 served as a loading control. The immunoblot images (A and B) are representative of three independent experiments.

4.4. m^{6}A modification promotes the translational efficiency of L1 RNA.

Given that m^{6}A regulates L1 ORF1p expression, I investigated the stages in the L1 replication cycle that are regulated by m^{6}A modification. First, I quantified L1 RNA expression in the presence or absence of the 5′ UTR m^{6}A cluster using two different plasmids, pL1Hs and pYX014. Irrespective of the vectors used, L1 m^{6}A mutation did not influence the levels of L1 RNA expression through northern blot and qRT-PCR (Figure 21A, B, C). I next assessed the stability of reporter L1 mRNAs with or without the 5′ UTR m^{6}A mutation using the transcription inhibitor, actinomycin D. L1 RNA was more stable in both pL1Hs- and pL1 m^{6}A mut-expressing HeLa cells when compared to positive control, c*MYC* mRNA (Figure 22A). I did not observe any significant difference in L1 RNA stability by m^{6}A mutation (Figure 22A). I next examined the distribution of reporter L1 mRNAs in the nuclear and cytoplasmic fractions. In comparison to that of *GAPDH* (abundant in the cytoplasm) and *MALAT1* (abundant in the nucleus), over 80% of the L1 mRNA was present in the cytoplasmic fraction and the m^{6}A-deficient mutation did not affect the cellular localization of L1 RNA (Figure 22B).

Several recent studies have linked 5′ UTR m^{6}A modification to translational

efficiency in the context of cellular stress (Coots et al., 2017; Meyer *et al.*, 2015; Zhou et al., 2015). Besides, a previous study raised the possibility that the presence of the L1 5′ UTR determines the quality of L1 RNA (An et al., 2011). Therefore, I reasoned that the L1 5′ UTR m^6A cluster could modulate the translation of L1 RNA. To test this hypothesis, I performed an immunoblot assay in HeLa cells that expressed a single to triple m^6A mutant of the pL1 construct. The expression levels of ORF1p gradually decreased as the number of mutations increased at similar transfection efficiencies (Figure 23A, B). In addition, through polysome profiling, I captured polysome-bound RNA to assess the translational efficiency of L1 RNA. The deletion of the m^6A cluster significantly reduced the enrichment of polysome-bound L1 RNA compared to that of pL1Hs (Figure 24A, B). To validate these results, I investigated whether m^6A regulates the translational efficiency of endogenous L1 mRNAs in PA-1 human embryonic carcinoma cells. Consistent with the effects of m^6A machinery depletion in pL1Hs-expressing HeLa cells (Figure 8A), ALKBH5 knockdown augmented the production of endogenous ORF1p while METTL3 knockdown reduced ORF1p synthesis (Figure 25A). The comparable levels of L1 mRNA in PA-1 cells with or without ALKBH5 depletion suggests that the enhanced production of ORF1p is a consequence of translational upregulation (Figure 25B). Consistent with this result, the levels of polysome-associated L1

RNA substantially increased in ALKBH5-depleted PA-1 cells in comparison to the control cells (Figure 26A, B), which indicates that ALKBH5 regulates L1 retrotransposition by suppressing the efficiency of L1 RNA translation.

eIF3 is an m^6A-binding protein and promotes the selective translation of mRNAs that bear m^6A in 5′ UTR (Meyer *et al.*, 2015). These characteristics of eIF3 lead me to hypothesize that the L1 5′ UTR m^6A cluster serves as a docking site for eIF3 to promote translation. To define the functional relationship between eIF3 and the L1 m^6A cluster, I analyzed previously reported data from photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation sequencing (PAR-CLIP seq) of eIF3 subunits a, b, d, and g (Lee *et al.*, 2015). By mapping the reads from PAR-CLIP of the eIF3 subunits along the endogenous L1Hs, I revealed that eIF3 exhibits preferential binding to the L1 5′ UTR (Figure 27A). Furthermore, the PAR-CLIP clusters were significantly enriched in the A332 m^6A region in all four eIF3 subunits, while the A495 m^6A region contained PAR-CLIP clusters of three eIF3 subunits: eIF3a, d, and g (Figure 27B). The comparable eIF3-binding sites in the A600 m^6A region were not observed (Figure 27B). To verify the interaction between eIF3 and the L1 m^6A cluster, I transfected pL1Hs or pL1 m^6A mut into HeLa cells and performed UV crosslinking immunoprecipitation using eIF3b

antibody (Figure 28A). Through RT-qPCR analysis of the immunoprecipitated eluates, I observed the enrichment of L1 RNA comparable to *c-JUN,* a known eIF3-bound mRNA, in pL1Hs-expressing cells (Figure 28B). *PSMB6* and eluates from IgG immunoprecipitation served as negative controls. Remarkably, the silencing of the m^6A cluster reduced the quantity of eIF3-bound L1 RNA by approximately 70%, which indicates that the L1 5′ UTR m^6A cluster bears the eIF3 docking site (Figure 28B). Indeed, eIF3 knockdown suppressed endogenous L1 ORF1p expression in PA-1 cells and L1 retrotransposition in HeLa cells (Figure 29A, B).

Under cellular stress like heat shock, eIF3 promotes cap-independent translation of mRNAs which bear m^6As in 5′ UTR. Heat shock stress suppresses most cap-dependent translation (Holcik and Sonenberg, 2005), but 5′ UTR m^6A of HSP70 mRNA was reported to enable efficient protein synthesis via cap-independent translation (Meyer *et al.*, 2015; Zhou *et al.*, 2015). However, the translation of HSP70 mRNA also requires 5′ end m7GpppG-capping (Song et al., 1995), which suggests that cap-independent translation by 5′ UTR m^6A could be turned on during certain stress condition. Therefore, I speculated that L1 ORF1p translation might be regulated in similar way of HSP70. Indeed, through immunoblot assays, I revealed that heat shock stress induced L1

ORF1p synthesis in pL1Hs-expressing HeLa cells, but not in pL1 m^{6}A mut-expressing HeLa cells (Figure 29C). This result indicates that heat shock upregulates 5′ UTR m^{6}A of L1 RNA not only of HSP70 mRNA, also suggests that eIF3-mediated cap-independent translation is involved in the regulation of L1 ORF1p translation.

Since another m^{6}A binding protein, YTHDF1, regulates translation efficiency of m^{6}A-modified RNA and interacts with eIF3 (Wang et al., 2015), I tested whether YTHDF1 also binds to L1 5′ UTR m^{6}A cluster. Through immunoprecipitation and qPCR analysis, I confirmed that YTHDF1 and another YTH protein, YTHDF2 interacts with eIF3 and L1 RNA. (Figure 30A, B, C). However, 5′ UTR m^{6}A cluster mutation did not impair interaction between YTHDFs and L1 RNA (Figure 30C). These data suggest that YTHDFs bind to L1 RNA via m^{6}A in region other than 5′ UTR. Collectively, the 5′ UTR m^{6}A cluster specifically recruits eIF3 for the efficient translation of L1 RNA.

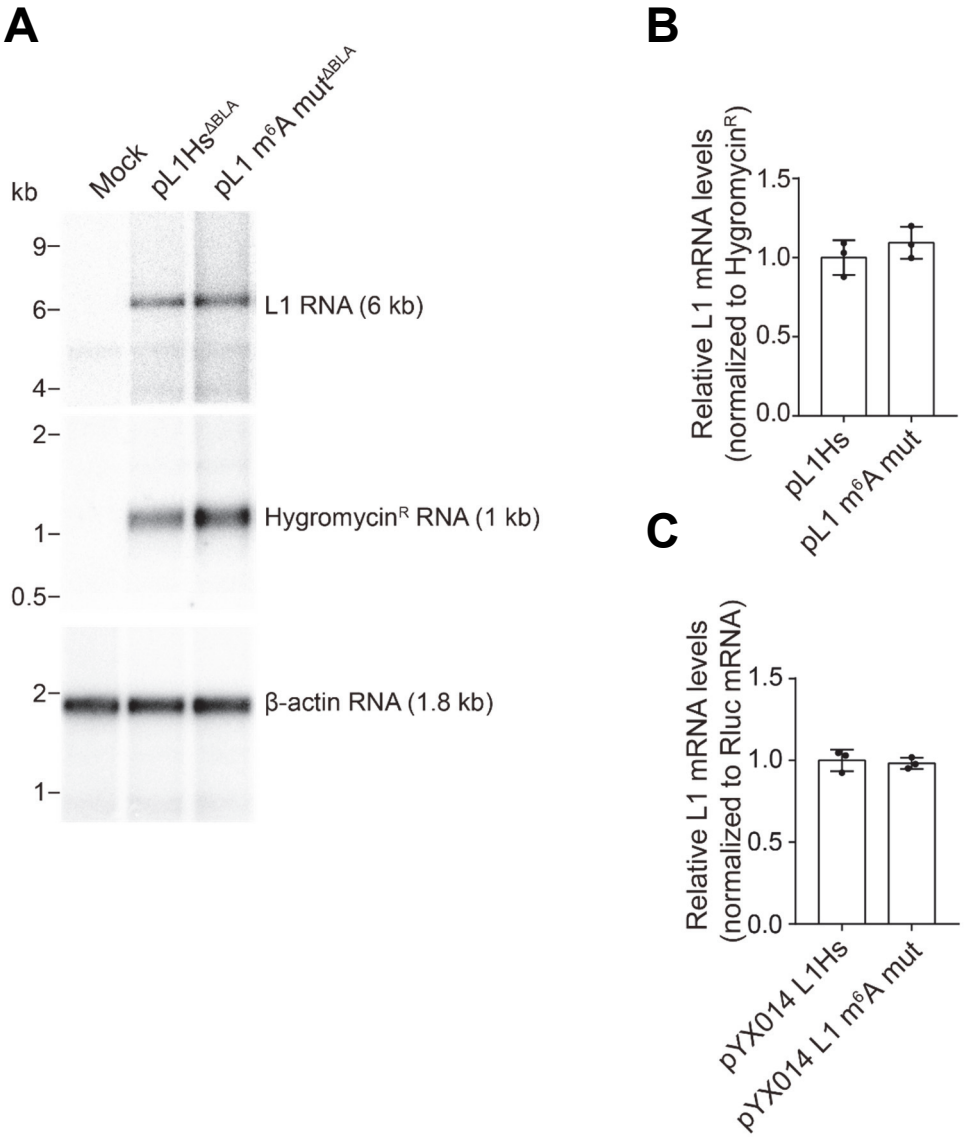

A
Mock
pL1Hs^ΔBLA
pL1 m6A mut^ΔBLA
kb
9
6
4
2
1
0.5
L1 RNA (6 kb)
HygromycinR RNA (1 kb)
2
1
β-actin RNA (1.8 kb)
B
Relative L1 mRNA levels (normalized to HygromycinR)
1.5
1.0
0.5
0.0
pL1Hs
pL1 m6A mut
C
Relative L1 mRNA levels (normalized to Rluc mRNA)
1.5
1.0
0.5
0.0
pYX014 L1Hs
pYX014 L1 m6A mut

Figure 21. L1 5′ UTR m^{6}A cluster does not affect RNA expression.

(A) Northern blot of full-length L1 mRNAs expressed from a plasmid encoding a full-length L1Hs or L1 m^{6}A mut lacking a reporter cassette (pL1Hs$^{\Delta BLA}$ or pL1m^{6}A mut$^{\Delta BLA}$). Hygromycin^R and β-actin served as loading controls. Marks on the left indicate positions of the RNA reference. Northern blot images are representative of two independent experiments. (B and C) The levels of RNA expression of reporter L1. HeLa cells were transfected with pL1Hs and its m^{6}A mutant construct (b), or pYX014 L1Hs and its mutant (c). The relative levels of reporter L1 transcripts are normalized to those of pL1Hs-encoded hygromycin-resistant gene (b) or pYX014-encoded Renilla luciferase mRNA (c). (n = 3 independent samples, mean ± s.d., unpaired two-tailed t test; p = 0.3403 for (b) and p = 0.7003 for (c))

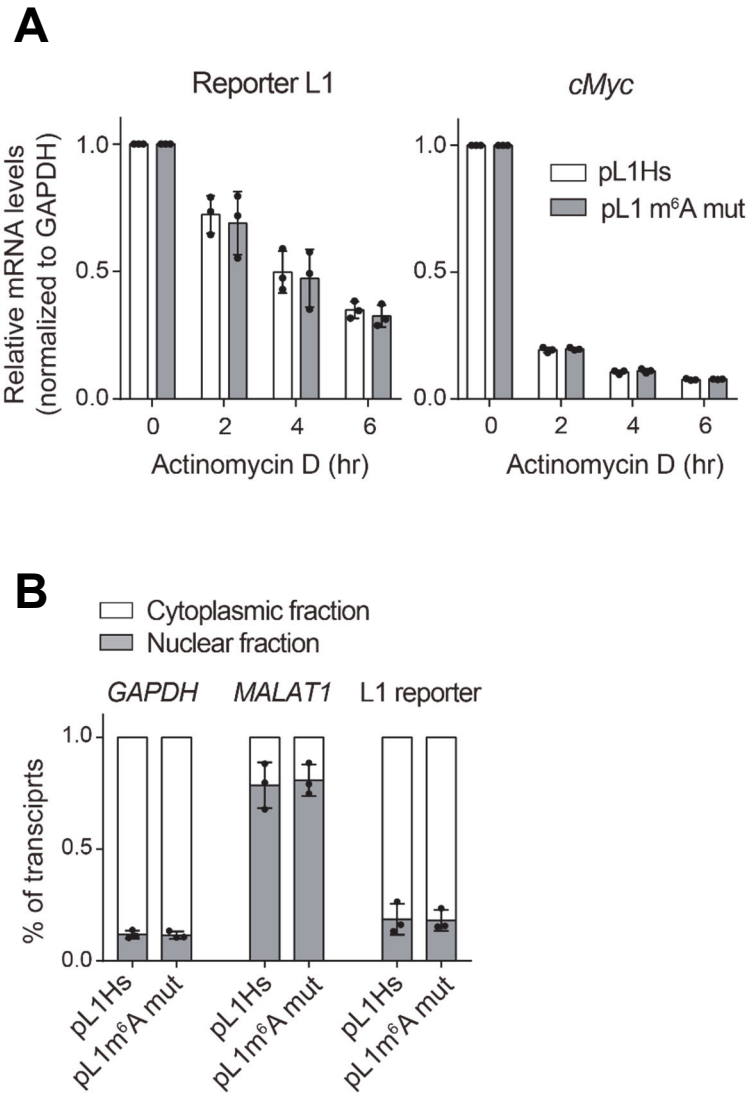
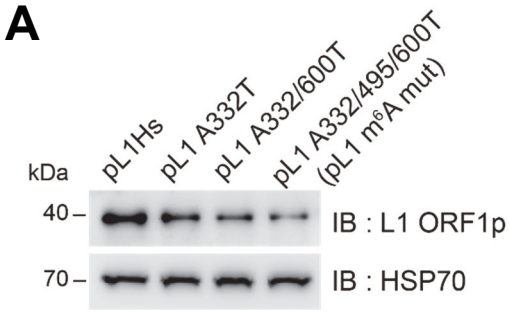

A
Reporter L1
cMyc
Relative mRNA levels (normalized to GAPDH)
1.0
0.5
0.0
0
2
4
6
Actinomycin D (hr)
pL1Hs
pL1 m6A mut
B
Cytoplasmic fraction
Nuclear fraction
GAPDH
MALAT1
L1 reporter
% of transciprts
pL1Hs
pL1m6A mut

Figure 22. L1 RNA stability and cellular localization are not regulated by 5′ UTR m^{6}A cluster.

(d) L1 RNA decay assay using pL1Hs-or pL1 m^{6}A mut-expressing HeLa cells. The cells were harvested at 0, 2, 4, and 6 h after actinomycin D treatment. mRNA levels were normalized to those of GAPDH. cMyc served as a positive control for this assay. (n = 3 independent samples, mean ± s.d., two-way ANOVA; p = 0.5003 for reporter L1 and p = 0.2826 for cMyc) (e) Distribution of reporter L1 RNA in the cytoplasmic and nuclear fraction of pL1-expressing HeLa cells. The percentage of transcripts was estimated assuming that the sum of the percentages of cytoplasmic and nuclear transcripts is 100%. (n = 3 independent samples, mean ± s.d., unpaired two-tailed t test; p = 0.8794 for GAPDH, p = 0.7721 for MALAT1, p = 0.9235 for reporter L1).


A
kDa
pL1Hs
pL1 A332T
pL1 A332/600T
pL1 A332/495/600T
(pL1 m6A mut)
40
70
IB : L1 ORF1p
IB : HSP70

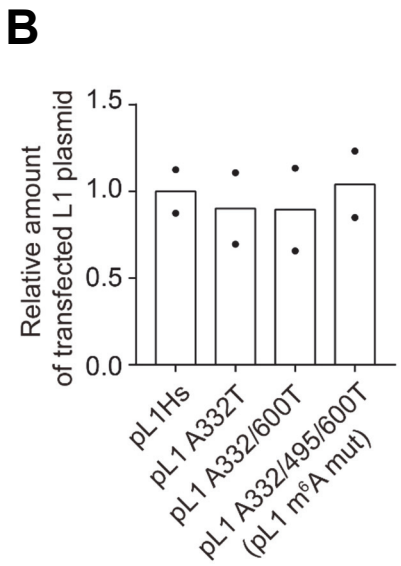

B
Relative amount
of transfected L1 plasmid
1.5
1.0
0.5
0.0
pL1Hs
pL1 A332T
pL1 A332/600T
pL1 A332/495/600T
(pL1 m6A mut)

Figure 23. 5′ UTR m^{6}A cluster is crucial for L1 ORF1p expression.

(A) Immunoblot analysis for assessing the effect of m^{6}A mutation in L1 ORF1p levels. HSP70 served as a loading control. The immunoblot images are representative of three independent experiments. (B) Quantification of transfected pL1 amount in HeLa cell expressing indicated pL1 constructs. (n = 2 independent samples, mean of two replicates)

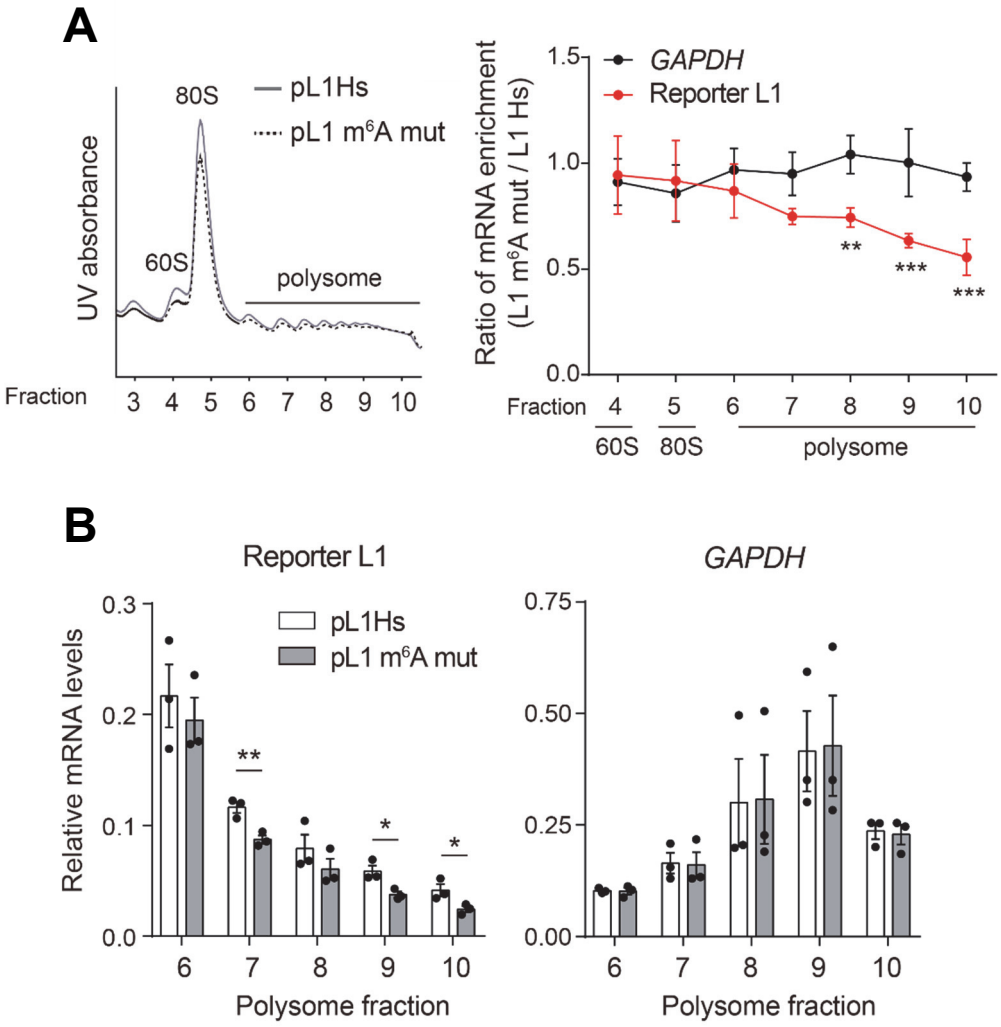

A
80S
60S
pL1Hs
pL1 m6A mut
UV absorbance
polysome
Fraction
3 4 5 6 7 8 9 10
Ratio of mRNA enrichment
(L1 m6A mut / L1 Hs)
GAPDH
Reporter L1
1.5
1.0
0.5
0.0
**
***
***
Fraction 4 5 6 7 8 9 10
60S 80S polysome
B
Reporter L1
GAPDH
pL1Hs
pL1 m6A mut
Relative mRNA levels
0.3
0.2
0.1
0.0
0.75
0.50
0.25
0.00
**
*
*
6 7 8 9 10
Polysome fraction
6 7 8 9 10
Polysome fraction

Figure 24. 5′ UTR m^{6}A cluster enhances the translational efficiency.

(A) Polysome profiling of pL1Hs- or pL1 m^{6}A mut-expressing HeLa cells (left panel). Ratio of the polysome-bound mRNA levels in pL1 m^{6}A mut-expressing cells to those in pL1Hs-expressing cells (right panel). The levels of RNA in each polysome fraction were normalized to the spike-in control and to the levels of input RNA. (n = 4 independent samples, mean ± s.d., two-way ANOVA and Bonferroni's multiple comparisons test; ***p < 0.001, **p < 0.01, and *p < 0.05 in comparison to the enrichment ratio of GAPDH in each fraction)

(B) Relative mRNA levels of polysome-bound reporter L1 and GAPDH. Polysome profiling of pL1-expressing HeLa cells was performed by sucrose gradient sedimentation. Specific mRNA levels were measured using RT-qPCR. The values are normalized to those of spike-in RNA and then to those of input RNA. (n = 3 independent samples, mean ± s.e.m., unpaired two-tailed t test; **p < 0.01, *p < 0.05)

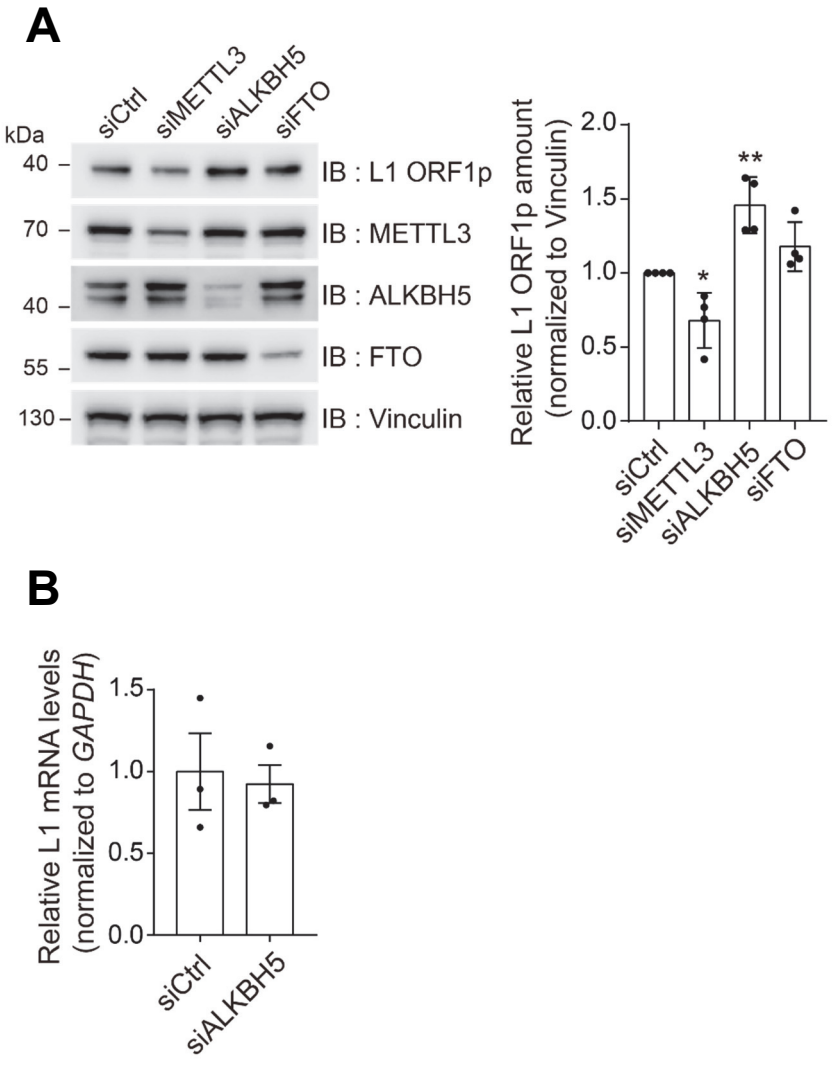

A
siCtrl
siMETTL3
siALKBH5
siFTO
kDa
40
70
40
55
130
IB : L1 ORF1p
IB : METTL3
IB : ALKBH5
IB : FTO
IB : Vinculin
Relative L1 ORF1p amount
(normalized to Vinculin)
2.0
1.5
1.0
0.5
0.0
*
**
siCtrl
siMETTL3
siALKBH5
siFTO
B
Relative L1 mRNA levels
(normalized to GAPDH)
1.5
1.0
0.5
0.0
siCtrl
siALKBH5

Figure 25. m^{6}A enzymes regulate the translation of endogenous L1 mRNA.

(A) Immunoblot assay for determining endogenous L1 ORF1p levels in PA-1 cells treated with indicated siRNAs (left). Vinculin served as a loading control. The immunoblot images are representative of four independent experiments. Quantification of L1 ORF1p levels is shown as values normalized to those of Vinculin (right). L1 ORF1p amounts are normalized to those of Vinculin. Band intensity quantification were performed using ImageJ (n = 4 independent samples, mean ± s.d., Dunnett's multiple comparisons test; **p < 0.01, *p < 0.05). (B) Endogenous L1 mRNA expression levels in ALKBH5-depleted PA-1 cells. The RNA levels were estimated using RT-qPCR with the specific primer for L1 5′ UTR and GAPDH. (n = 3 independent samples, mean ± s.e.m., unpaired two-tailed t test; p = 0.7852)

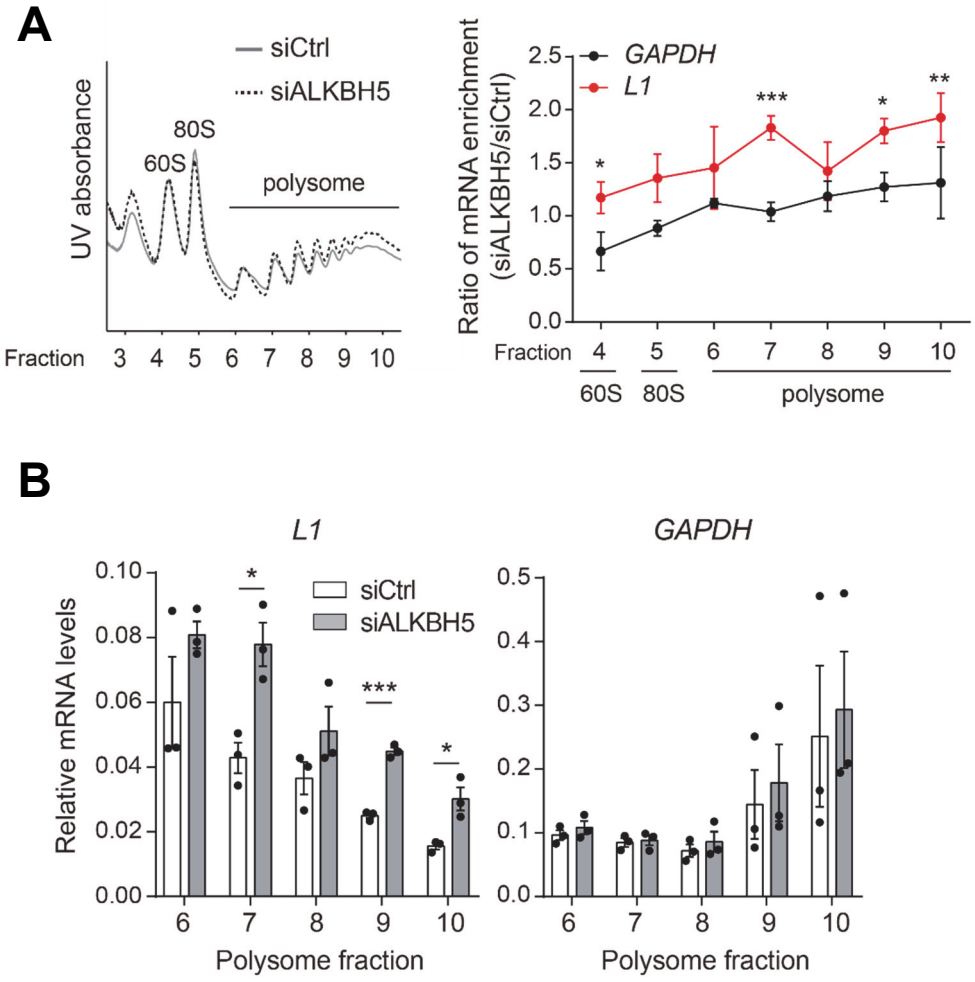
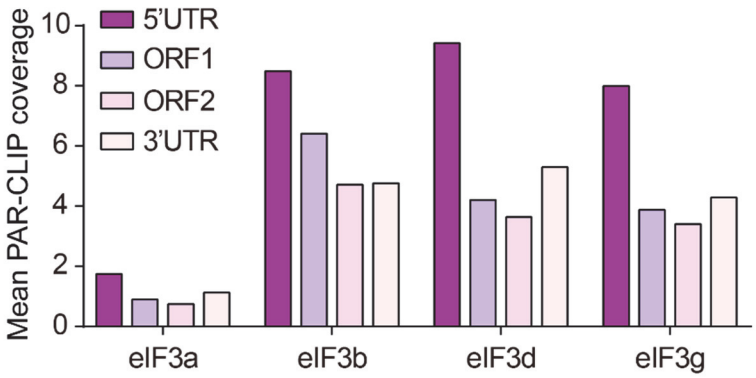

A
siCtrl
siALKBH5
80S
60S
polysome
UV absorbance
Fraction 3 4 5 6 7 8 9 10
Ratio of mRNA enrichment (siALKBH5/siCtrl)
GAPDH
L1
2.5
2.0
1.5
1.0
0.5
0.0
*
***
*
**
Fraction 4 5 6 7 8 9 10
60S 80S polysome
B
L1
GAPDH
Relative mRNA levels
0.10
0.08
0.06
0.04
0.02
0.00
siCtrl
siALKBH5
*
***
*
6 7 8 9 10
Polysome fraction
0.5
0.4
0.3
0.2
0.1
0.0
6 7 8 9 10
Polysome fraction

Figure 26. ALKBH5 suppresses the translational efficiency of L1 mRNA.

(A) Polysome profiling of PA-1 cells lacking ALKBH5 compared to siCtrl (left panels). The levels of RNA in each polysome fraction were normalized to the spike-in control and to the levels of input RNA. (n = 3 independent samples, mean ± s.d., two-way ANOVA and Bonferroni's multiple comparisons test; ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$ in comparison to the enrichment ratio of GAPDH in each fraction) (B) Relative mRNA levels of polysome-bound endogenous L1 RNA from ALKBH5-depleted PA-1 cells. The polysome-bound RNA was quantified using RT-qPCR. L1 5′ UTR-specific primer was used to detect the endogenous L1 RNA. RNA expression levels were first normalized to those of spike-in RNA and then to those of input RNA (n = 3 independent samples, mean ± s.e.m., unpaired two-tailed t test; ***$p < 0.001$, *$p < 0.05$).

A


Mean PAR-CLIP coverage
10
8
6
4
2
0
5'UTR
ORF1
ORF2
3'UTR
eIF3a
eIF3b
eIF3d
eIF3g

B

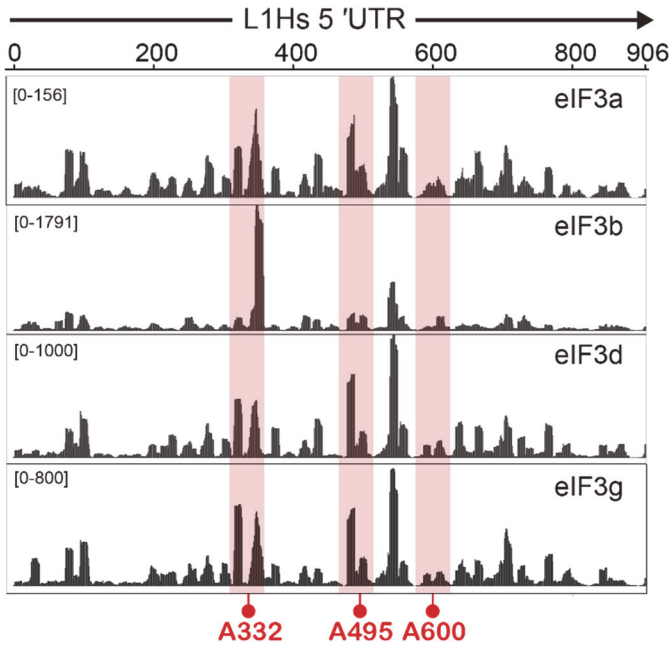

L1Hs 5 'UTR
0
200
400
600
800
906
[0-156]
eIF3a
[0-1791]
eIF3b
[0-1000]
eIF3d
[0-800]
eIF3g
A332
A495
A600

Figure 27. eIF3 binding sites in L1 5′ UTR

(A) Distribution of eIF3 subunit crosslinking sites along the L1Hs consensus sequence. (B) Identification of eIF3 binding sites in L1Hs 5′ UTR. The red boxes indicate the m^6A sites-containing region. (A and B) Previously reported data (GSE65004) was used for the analysis (Lee *et al.*, 2015).

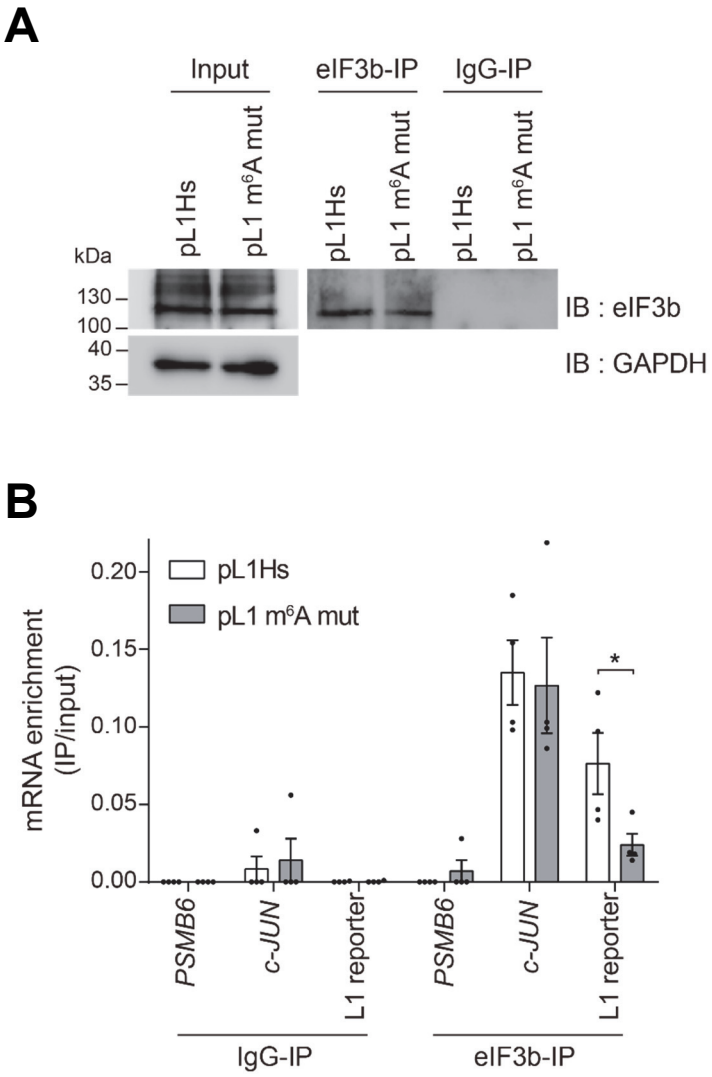

A
Input
eIF3b-IP
IgG-IP
pL1Hs
pL1 m6A mut
pL1Hs
pL1 m6A mut
pL1Hs
pL1 m6A mut
kDa
130
100
40
35
IB : eIF3b
IB : GAPDH
B
pL1Hs
pL1 m6A mut
mRNA enrichment (IP/input)
0.20
0.15
0.10
0.05
0.00
*
PSMB6
c-JUN
L1 reporter
PSMB6
c-JUN
L1 reporter
IgG-IP
eIF3b-IP

Figure 28. eIF3 interacts with L1 5′ UTR m^{6}A cluster.

(A) Immunoblot assay of eIF3b-immunoprecipitation samples. HeLa cells were transfected with pL1 constructs. Two days after transfection, the cells were subject to UV crosslinking and eIF3b immunoprecipitation. The samples were boiled after adding 10 nM MgCl2 to fragment the crosslinked RNA. GAPDH served as the loading control for the input lysates. The immunoblot images are representative of three independent experiments. (B) eIF3 UV-CLIP-qPCR using pL1Hs- or pL1 m^{6}A mut-expressing HeLa cells. IgG-IP and PSMB6 served as negative controls (n = 4 independent samples, mean ± s.e.m., unpaired two-tailed t test; *$p < 0.05$).

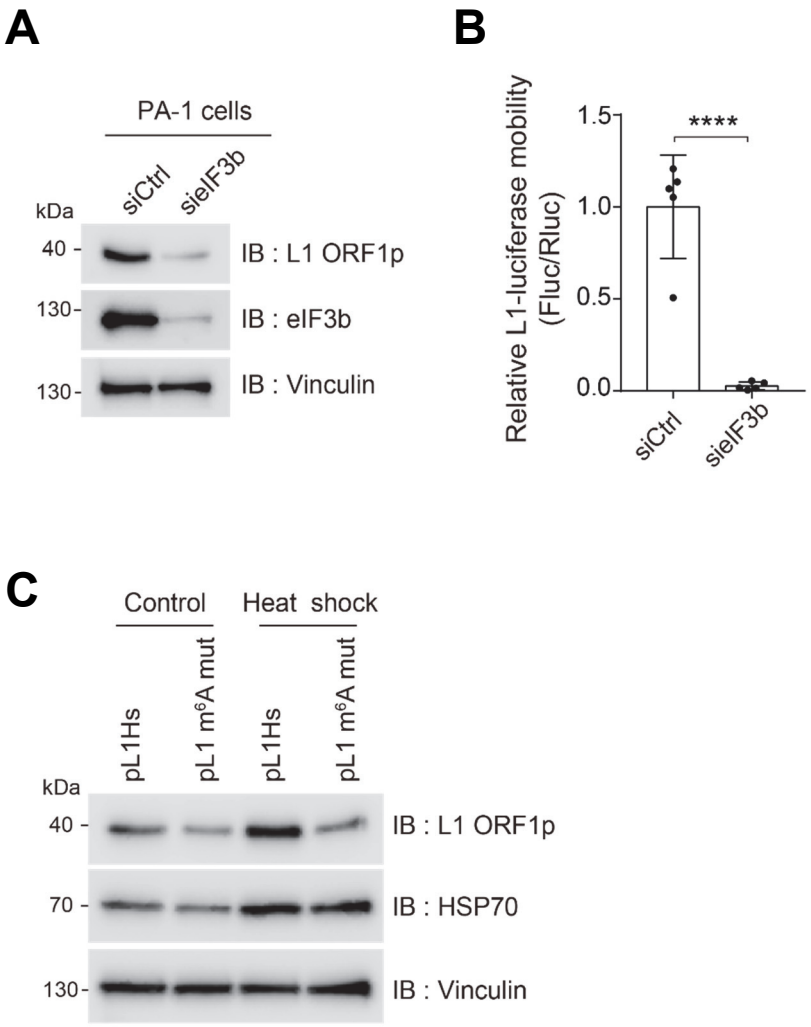

A
PA-1 cells
siCtrl
sieIF3b
kDa
40
130
130
IB : L1 ORF1p
IB : eIF3b
IB : Vinculin
B
Relative L1-luciferase mobility
(Fluc/Rluc)
1.5
1.0
0.5
0.0
****
siCtrl
sieIF3b
C
Control
Heat shock
pL1Hs
pL1 m6A mut
pL1Hs
pL1 m6A mut
kDa
40
70
130
IB : L1 ORF1p
IB : HSP70
IB : Vinculin

Figure 29. eIF3 recruitment is crucial for L1 activity.

(A) Immunoblot assay of PA-1 cells devoid of eIF3b by using siRNA. Vinculin served as a loading control. The immunoblot images are representative of two independent experiments. (B) L1-luciferase retrotransposition assay using pYX014 L1Hs constructs. pYX014 L1Hs was transfected into sieIF3b-treated HeLa cells. After 5 d from transfection, luminescence of firefly luciferase (Fluc) and of renilla luciferase (Rluc) were measured. Ratios of luminescence (Fluc/Rluc) were normalized to those of siCtrl (n = 5 independent samples, mean ± s.d., unpaired two-tailed t test; ****$p < 0.0001$). (C) Heat shock stress promotes L1 ORF1p translation. pL1-transfected cells were incubated at 42°C for 1 h, and then were harvested at 9 h post heat shock stress. Cell lysates were subjected to western blot assay.

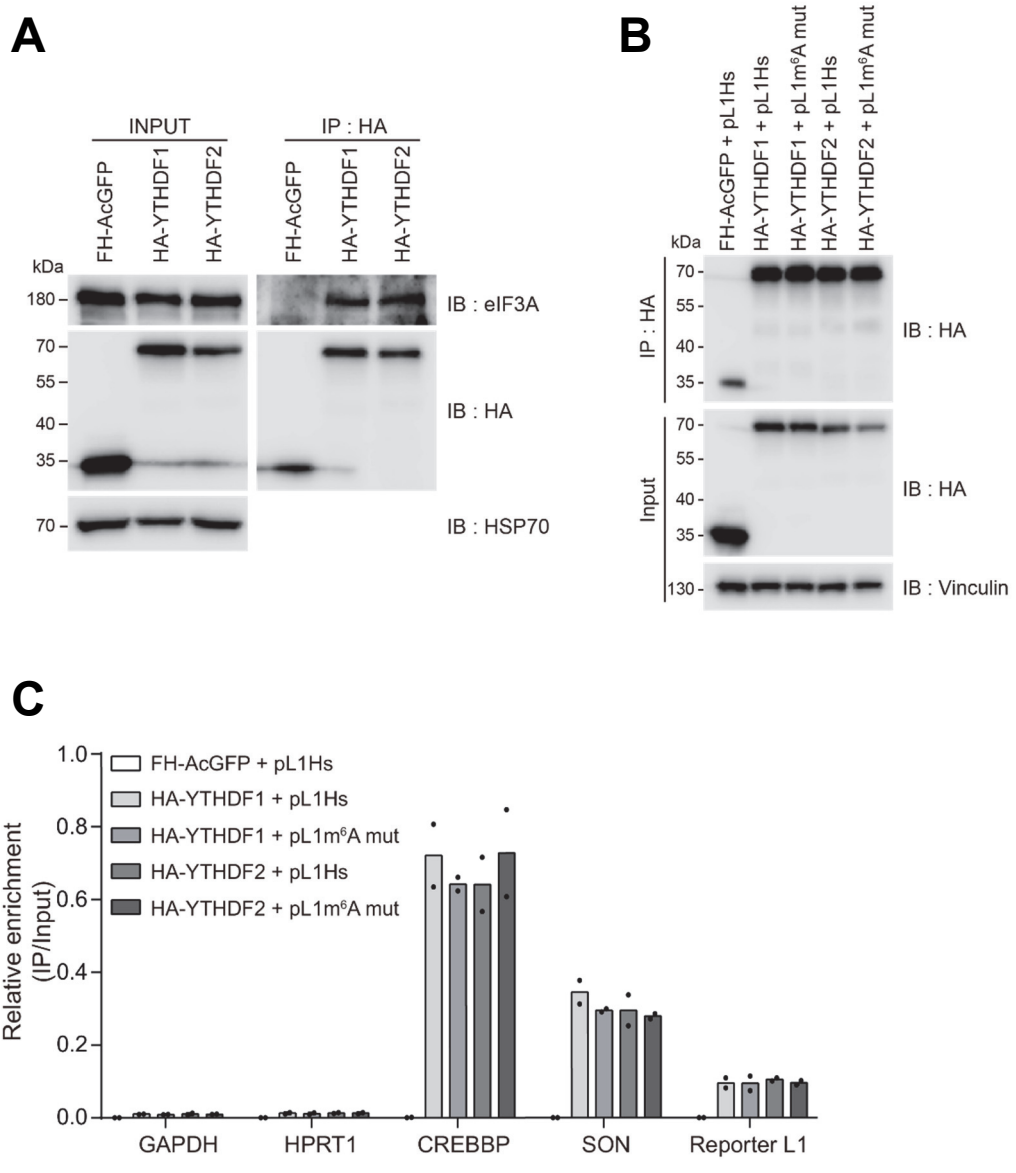

A
INPUT
IP : HA
FH-AcGFP
HA-YTHDF1
HA-YTHDF2
kDa
180
70
55
40
35
IB : eIF3A
IB : HA
IB : HSP70
B
FH-AcGFP + pL1Hs
HA-YTHDF1 + pL1Hs
HA-YTHDF1 + pL1m6A mut
HA-YTHDF2 + pL1Hs
HA-YTHDF2 + pL1m6A mut
IP : HA
Input
130
IB : Vinculin
C
FH-AcGFP + pL1Hs
HA-YTHDF1 + pL1Hs
HA-YTHDF1 + pL1m6A mut
HA-YTHDF2 + pL1Hs
HA-YTHDF2 + pL1m6A mut
Relative enrichment (IP/Input)
1.0
0.8
0.6
0.4
0.2
0.0
GAPDH
HPRT1
CREBBP
SON
Reporter L1

Figure 30. YTHDF1 and 2 interact with L1 RNA independently of 5′ UTR m^{6}A cluster.

(A) Co-immunoprecipitation assay of HA-YTHDF1 and 2 using HA-antibody. (B and C) RNA-immunoprecipitation assay of HA-YTHDF and pL1 co-transfected HeLa cells. Immunoblot assay of HA-immunoprecipitation samples (B). Specific mRNA expression of HA-immunoprecipitation eluates using qPCR (C). Immunoprecipitation using FH-AcGFP served as negative control. The levels of RNA were normalized to the spike-in control and to the levels of input RNA (n = 2 independent samples, mean of two replicates). The immunoblot images of (A) and (B) are representative of two independent experiments.

4.5. 5′ UTR m^{6}A cluster is necessary to produce a functional unit for L1 retrotransposition.

For successful L1 retrotransposition, both ORF1p and ORF2p are required to generate the L1 RNP with the encoding L1 RNA (Wei et al., 2001). Though I observed m^{6}A-mediated regulation in ORF1p synthesis, it is necessary to determine whether m^{6}A modification at the 5′ UTR influences L1 ORF2p expression. Therefore, I tried to detect L1 ORF2p levels using pAD3TE1-transfected cells (Figure 31A). Through western blot analysis, I obtained similar results of 5′ UTR m6A-mediated ORF1p expression in pAD3TE1 transfected cells (Figure 31B). Unlike L1 ORF1p, it has been hypothesized that as few as one molecule of ORF2p is translated per L1 RNA molecule (Alisch et al., 2006). This has made it difficult to detect ORF2p expression even using ectopic expression systems (Ergun et al., 2004; Goodier et al., 2004). Consistently, I confirmed extremely low expression levels of TAP-tagged ORF2p (Figure 31B).

To determine whether m^{6}A modification at the 5′ UTR influences L1 RNP formation, I obtained the cellular RNP fraction as previously reported (Kulpa and Moran, 2006). Briefly, I prepared lysates from pL1-transfected cells and purified L1 RNPs using sucrose cushion ultracentrifugation (Figure 32). I detected comparable levels of L1 RNA in the RNP fractions from pL1Hs- and

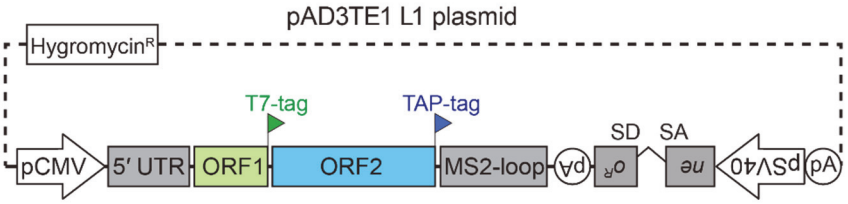
pL1 m^{6}A mut-expressing cells (Figure 33A, B). cDNA synthesis reaction in absence of reverse transcriptase revealed that neither genomic DNA nor plasmid contamination was present in the RNP fraction (Figure 33A). Immunoblotting of the RNP fraction showed that the levels of RNP-associated ORF1p were diminished by L1 5′ UTR m^{6}A mutation (Figure 33C). This indicates that the m^{6}A cluster mutation abolished the sufficient production of ORF1p for L1 RNP formation.

Since ORF2p expression level is too low to observe changes in the m^{6}A mutant (Figure 31B), I introduced the L1 element amplification protocol (LEAP) to gauge the reverse transcriptase activity of ORF2p (Kulpa and Moran, 2006) (Figure 32). Incubation of RNPs with LEAP primer facilitates ORF2p-mediated L1 cDNA synthesis. I amplified LEAP products using PCR with specific primers for reporter L1 and RACE adapter, which yielded products of 300 to 400 base pairs (bp) (Figure 33D). However, m^{6}A-abrogated L1 RNP produced cDNA at significantly lower levels than the wild-type L1 RNP did (Figure 33D). These results reveal that the m^{6}A cluster is necessary for L1 cDNA production, which suggests that the m^{6}A cluster regulates ORF2p expression or its activity.

ORF1p oligomerization is critical for successful L1 retrotransposition (Naufer et al., 2015). I examined whether inefficient ORF1p synthesis results in a failure

of L1 RNP formation. For a quantitative assessment of individual L1 RNP formation, I introduced the pAD3TE1 construct carrying T7-tagged ORF1p and MS2 stem-loop structures in the L1 3′ UTR (Figure 34A). I performed RNA fluorescence in situ hybridization (FISH) with fluorescent Q670-labeled probes complementary to the linker regions between the MS2 loops and immunofluorescence experiments with anti-T7 antibody (Figure 34B). Through z-stack analysis, I obtained the coordinates for the fluorescent signals of L1 RNA and ORF1p and identified the L1 RNPs by sorting out co-localizing particles within an intermolecular distance of 330 nm between L1 RNA and ORF1p. Consistent with the previous study (Doucet *et al.*, 2010), I observed co-localizing signals of L1 RNP as cytoplasmic aggregates (Figure 34B). However, L1 m^6A mut-expressing cells showed a significant reduction in both the number of L1 RNP foci and the signal intensity of co-localizing ORF1p (Figure 34B, C, D). These data indicate that the abrogation of the m^6A cluster reduces the levels of ORF1p in L1 RNP and causes a concomitant decrease in the number of L1 RNP particles.

A


pAD3TE1 L1 plasmid
Hygromycin^R
T7-tag
TAP-tag
SD
SA
pCMV
5′ UTR
ORF1
ORF2
MS2-loop
pA

B

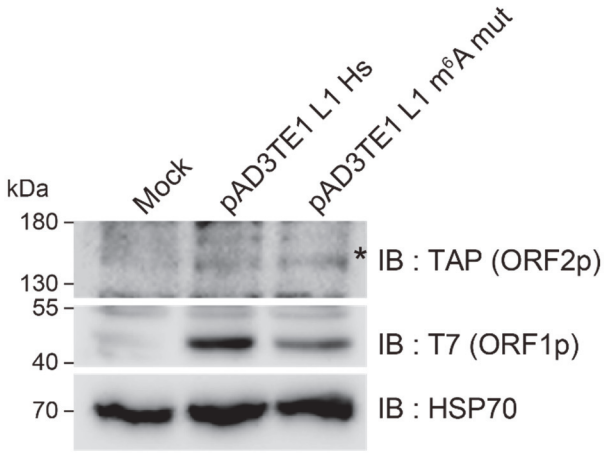

kDa
Mock
pAD3TE1 L1 Hs
pAD3TE1 L1 m^{6}A mut
180
130
55
40
70
*
IB : TAP (ORF2p)
IB : T7 (ORF1p)
IB : HSP70

Figure 31. Detection of L1 ORF2p using pAD3TE1-transfected HeLa cells

(A) A schematic of pAD3TE1 L1 plasmid used in this study. (B) Immunoblot assay of HeLa cells expressing pAD3TE1 L1Hs or pAD3TE1 L1 m^6A mut. HSP70 served as a loading control.

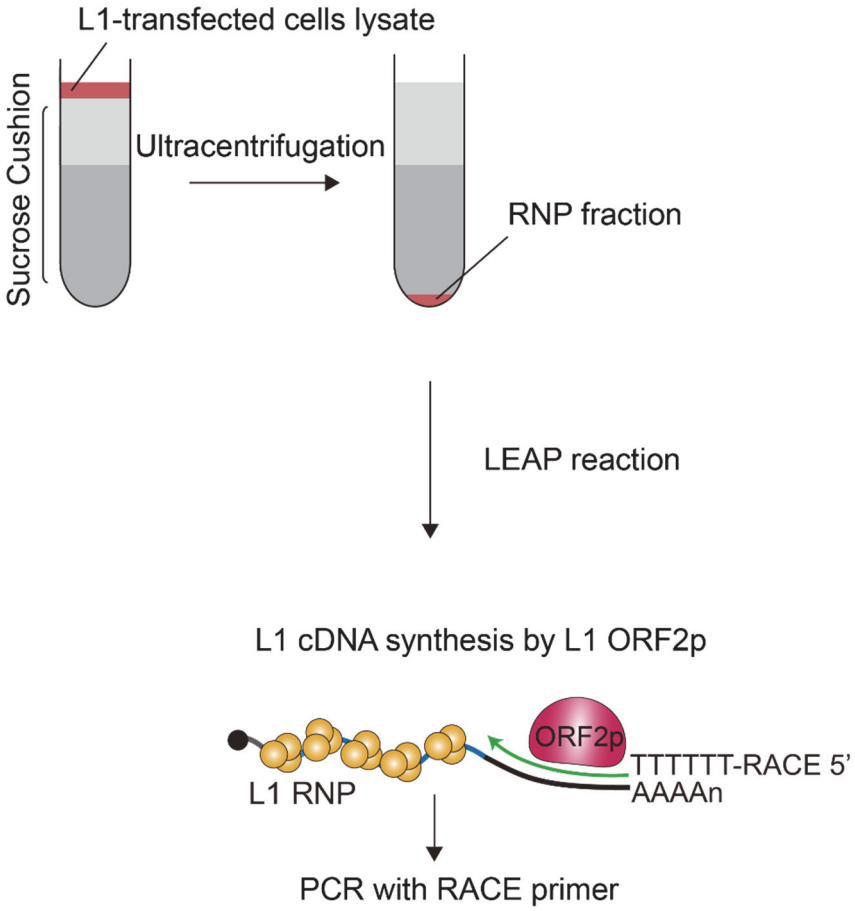

L1-transfected cells lysate
Sucrose Cushion
Ultracentrifugation
RNP fraction
LEAP reaction
L1 cDNA synthesis by L1 ORF2p
ORF2p
TTTTTT-RACE 5'
AAAAn
L1 RNP
PCR with RACE primer

Figure 32. Scheme of L1 RNP purification and LEAP assay

Ultracentrifugation and purification cellular RNP (left part of scheme, cell lysate in red, and sucrose cushion in gray). LEAP reaction (right part of scheme, L1 ORF1p in yellow, L1 ORF2p in magenta, 3′ RACE adapter in green).

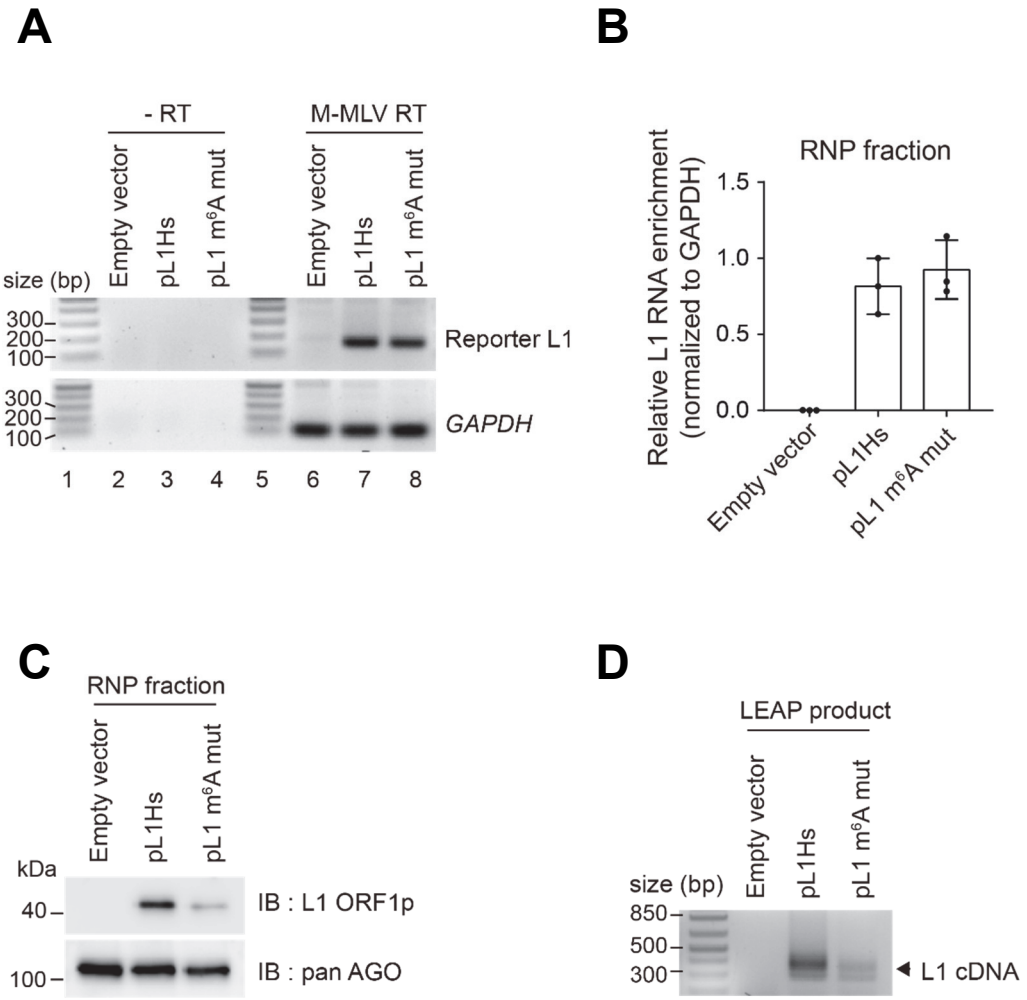

A
- RT
M-MLV RT
Empty vector
pL1Hs
pL1 m6A mut
size (bp)
300
200
100
Reporter L1
GAPDH
1 2 3 4 5 6 7 8
B
RNP fraction
Relative L1 RNA enrichment (normalized to GAPDH)
1.5
1.0
0.5
0.0
C
RNP fraction
kDa
40
100
IB : L1 ORF1p
IB : pan AGO
D
LEAP product
850
500
300
L1 cDNA

Figure 33. L1 5′ UTR m^{6}A cluster is crucial for retrotransposition-competent L1 RNP formation.

(A) Quantification of mRNA levels in the RNP fraction of pL1-expressing HeLa cells. cDNA synthesis in the absence of reverse transcriptase (lane 2-4) and transfection of empty vector (lane 6) served as negative controls. The RT-PCR products of reporter L1 and GAPDH are of 158 and 106 bp, respectively (lane 6-8). Lane 1 and 5 show the DNA ladder. (B) RT-qPCR analysis using the purified RNP fraction of pL1-expressing HeLa cells. The levels of L1 RNA were normalized to the spike-in control and GAPDH mRNA (n = 3 independent experiments, mean ± s.d.). (C) Immunoblot assay of the RNP fraction from pL1-expressing HeLa cells. pan AGO served as a loading control. (D) LEAP assay using RNP fraction from pL1-expressing HeLa cells. The LEAP product is a diffuse band of 300-400 bp. The images of (A), (C), and (D) are representative of three independent experiments.

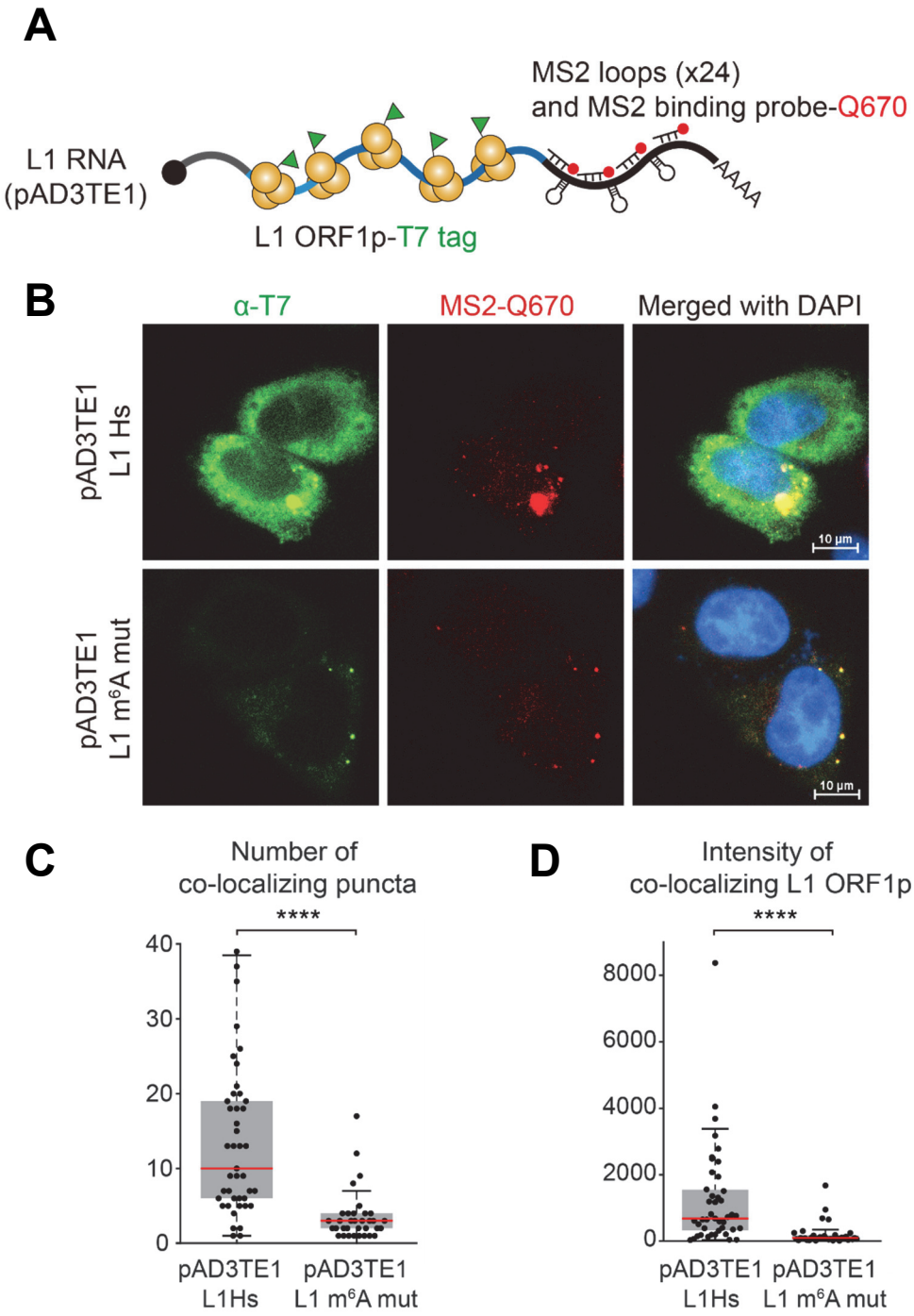

A
MS2 loops (x24)
and MS2 binding probe-Q670
L1 RNA
(pAD3TE1)
AAAA
L1 ORF1p-T7 tag
B
α-T7
MS2-Q670
Merged with DAPI
pAD3TE1
L1 Hs
pAD3TE1
L1 m6A mut
10 μm
10 μm
C
Number of
co-localizing puncta
****
40
30
20
10
0
pAD3TE1
L1Hs
pAD3TE1
L1 m6A mut
D
Intensity of
co-localizing L1 ORF1p
****
8000
6000
4000
2000
0
pAD3TE1
L1Hs
pAD3TE1
L1 m6A mut

Figure 34. m^{6}A modification regulates the formation of cellular L1 RNP bodies.

(A) A schematic of the L1-MS2 construct (pAD3TE1) carrying T7-tagged ORF1p (green) and MS2 stem-loops with Q670-labeled MS2 binding probes (red). (B) Immunofluorescence and RNA FISH images depicting HeLa cells transfected with pAD3TE1 L1Hs (top) or L1 m^{6}A mut (bottom). Images for T7-tagged ORF1p (green), L1-MS2 RNA (red), and the merged images with DAPI (blue) are indicated. The images are representative of two independent experiments (C) The number of L1 RNP foci in pAD3TE1-expressing HeLa cells. Co-localizing puncta within an intermolecular distance of 330 nm were counted as L1-RNP using z-stack analysis. (D) Intensity of L1 ORF1p in co-localizing puncta. Each point represents the intensity of L1 ORF1p per cell. (C and D) Box plots indicate median (red middle line), 25th, 75th percentile (gray box) and 5th and 95th percentile (whiskers). 43 cells for L1Hs and 34 cells for L1 m^{6}A mut, two-sided Kolmogorov-Smirnov test, ****$p < 0.0001$.

4.6. m^6A is a driving force for L1 evolution.

Over the last 40 million years of human evolution, L1 subfamilies have frequently acquired novel 5′ UTRs (Khan *et al.*). Since a new L1 lineage will emerge only through its successful replication, the genetic novelty that promotes L1 mobility must remain preserved in the genomic fossils of L1s (Furano, 2008). Considering that RNA methyltransferase installs m^6A in a sequence-specific manner, I speculated that nucleotide mutations might lead to the acquisition or loss of the m^6A consensus motif during L1 evolution. To unravel the evolutionary history of L1 5′ UTR m^6A cluster regions, I analyzed 443 human-specific full-length L1s (Tang *et al.*, 2018) and compared the three m^6A motif sites, A332, A495, and A600. Given that adenosine residue should be followed by cytosine residue to form the m^6A consensus motif DRA^{m}CH, A332 m^6A positive L1s constitute a considerably small part in the L1PA3 lineage (12.4%). In L1PA2 and younger lineages, the number of A332 m^6A positive L1s increased drastically (92.9%) (Figure 35), and the same was observed in the youngest L1Hs (Figure 36A). On the contrary, A495 and A600 are tightly conserved in all human-specific L1 subfamilies (Figure 36B). I investigated this tendency of the A332 m^6A motif in L1s of chimpanzee and gorilla, which share L1PA2 and L1PA3 lineages with humans. Comparative analysis of chimpanzee-

or gorilla-specific full-length L1s revealed the seismic shift toward the population of A332 m^6A positive L1s, while the chimpanzee- or gorilla-specific L1s continue to harbor the m^6A motifs of A495 and A600 (Figure 35, Figure 37A, B and Figure 38A, B). As in the L1Hs subfamily, the majority of the youngest chimpanzee-specific L1 subfamily (L1Pt) harbor the A332 m^6A motif (Figure 37A). In summary, I found that A332 m^6A motif acquisition by single nucleotide substitution (T333C) first appeared in L1PA3 or older lineages, which indicates that the productive potential of m^6A has allowed positive selection of A332 m^6A-positive L1s during the evolution from the common ancestor (Figure 39).

To evaluate the consequence of A332 m^6A acquisition in ancestral L1 5′ UTR, I generated a chimeric pL1 construct that contained L1PA2 5′ UTR and L1Hs ORF1/2 with the *mblastI* reporter (Figure 40A). Based on the m^6A consensus motif DRA^{m}CH, T333 of pL1PA2$^{5'UTR}$ had no m^6A motif at A332, whereas T333C point mutation enabled A332 m^6A modification (Figure 40A). The retrotransposition assay revealed that T333C mutation enhanced the mobility of pL1PA2$^{5'\ UTR}$, while the mutagenesis control (T333G) did not exert the same effect (Figure 40B, C). As expected, the T333C m^6A-gain mutation enhanced ORF1p synthesis of pL1PA2$^{5'\ UTR}$ (Figure 40D). Although the acquisition of

A332 m^6A motif only led to a 1.4-fold increase in the cultured cell-based L1 retrotransposition assays (Figure 40B), the 12 million years of L1 evolution would have been sufficient to amplify the profound effect of m^6A. These results suggest that m^6A modification in the L1 5' UTR region may have played a crucial role in the L1 evolution of primates.

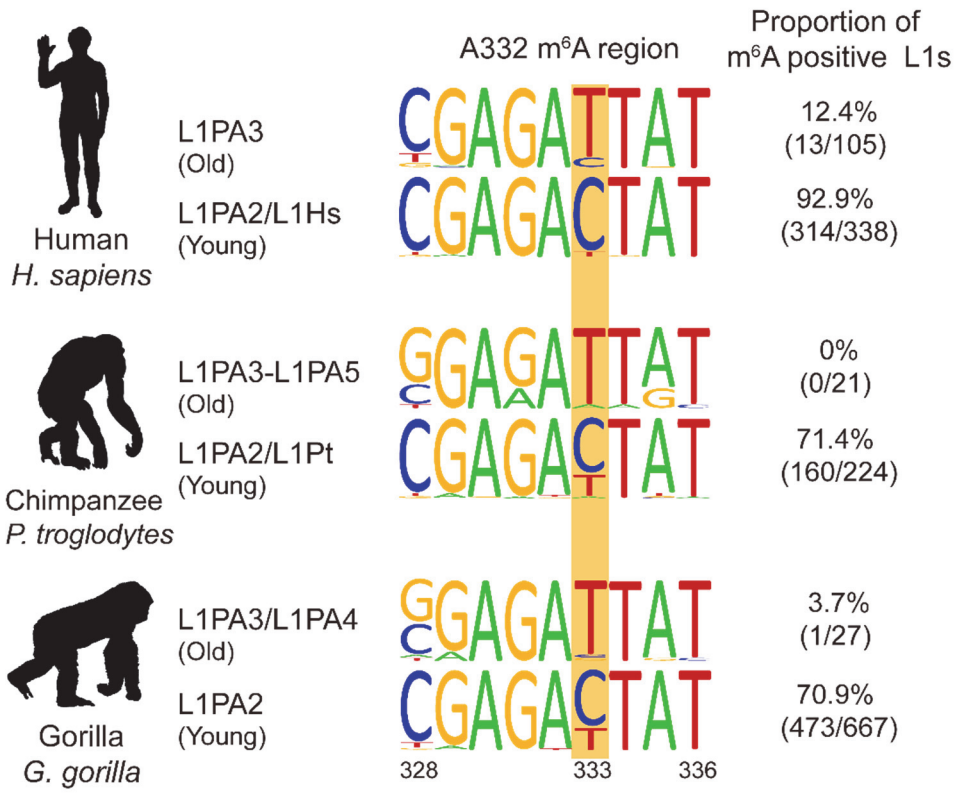
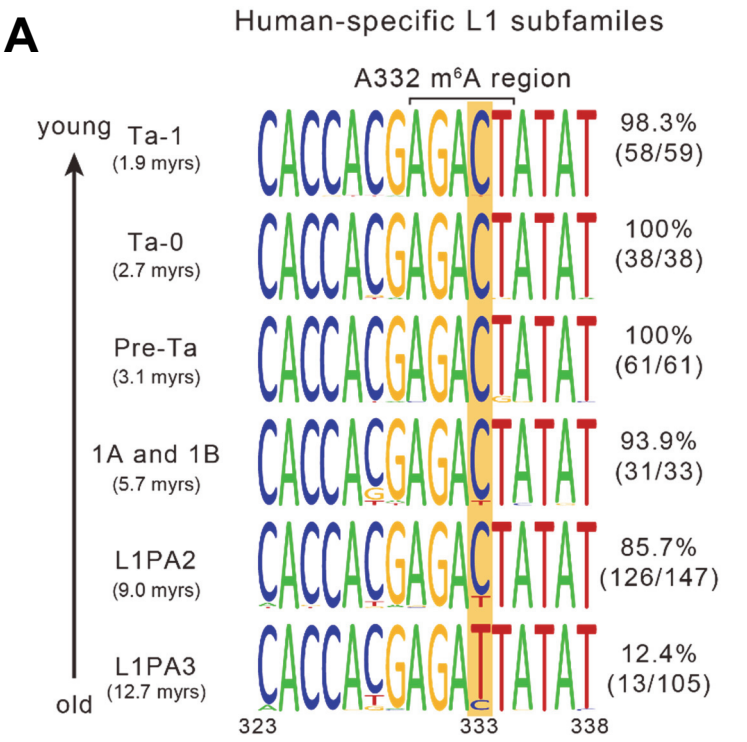

A332 m6A region
Proportion of m6A positive L1s
Human
H. sapiens
L1PA3
(Old)
L1PA2/L1Hs
(Young)
12.4%
(13/105)
92.9%
(314/338)
Chimpanzee
P. troglodytes
L1PA3-L1PA5
(Old)
L1PA2/L1Pt
(Young)
0%
(0/21)
71.4%
(160/224)
Gorilla
G. gorilla
L1PA3/L1PA4
(Old)
L1PA2
(Young)
3.7%
(1/27)
70.9%
(473/667)
328
333
336

Figure 35. Comparative analysis of L1 A332 m^6A sites in species-specific full-length L1s from three different primates

Changes in the A332-m^6A motif region from L1PA3 or older L1s to L1PA2 and a younger L1. The substitution site wherein the residue converts from T to C (333) is highlighted in yellow. The percentage indicates the proportion of m^6A motif-positive L1s with nucleotide C to total L1s. The height of the nucleotide indicates the frequency of the corresponding nucleotide.


A
Human-specific L1 subfamiles
A332 m6A region
young
old
Ta-1
(1.9 myrs)
Ta-0
(2.7 myrs)
Pre-Ta
(3.1 myrs)
1A and 1B
(5.7 myrs)
L1PA2
(9.0 myrs)
L1PA3
(12.7 myrs)
98.3%
(58/59)
100%
(38/38)
100%
(61/61)
93.9%
(31/33)
85.7%
(126/147)
12.4%
(13/105)
323
333
338

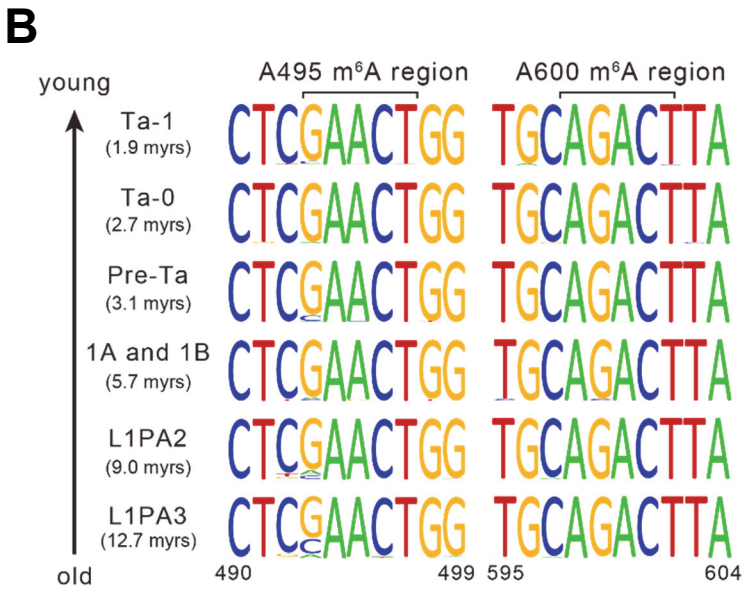

B
young
old
A495 m6A region
A600 m6A region
Ta-1
(1.9 myrs)
Ta-0
(2.7 myrs)
Pre-Ta
(3.1 myrs)
1A and 1B
(5.7 myrs)
L1PA2
(9.0 myrs)
L1PA3
(12.7 myrs)
490
499
595
604

Figure 36. Comparative analysis of the m^{6}A cluster in Human-specific full-length L1 subfamilies

(a) Comparative analysis of the A332 m^{6}A region (A323-A338) in human-specific full-length L1 subfamilies. The yellow box indicates the T to C substitution site at the 333rd residue. The percentage indicates the proportion of m^{6}A positive L1s to total L1s. (b) Comparative analysis of other m^{6}A regions at the A495 and the A600 residues in human-specific full-length L1 subfamilies (left: A495, right: A600). Differences in the substitution patterns in different subfamilies were not observable at the two other m^{6}A sites.

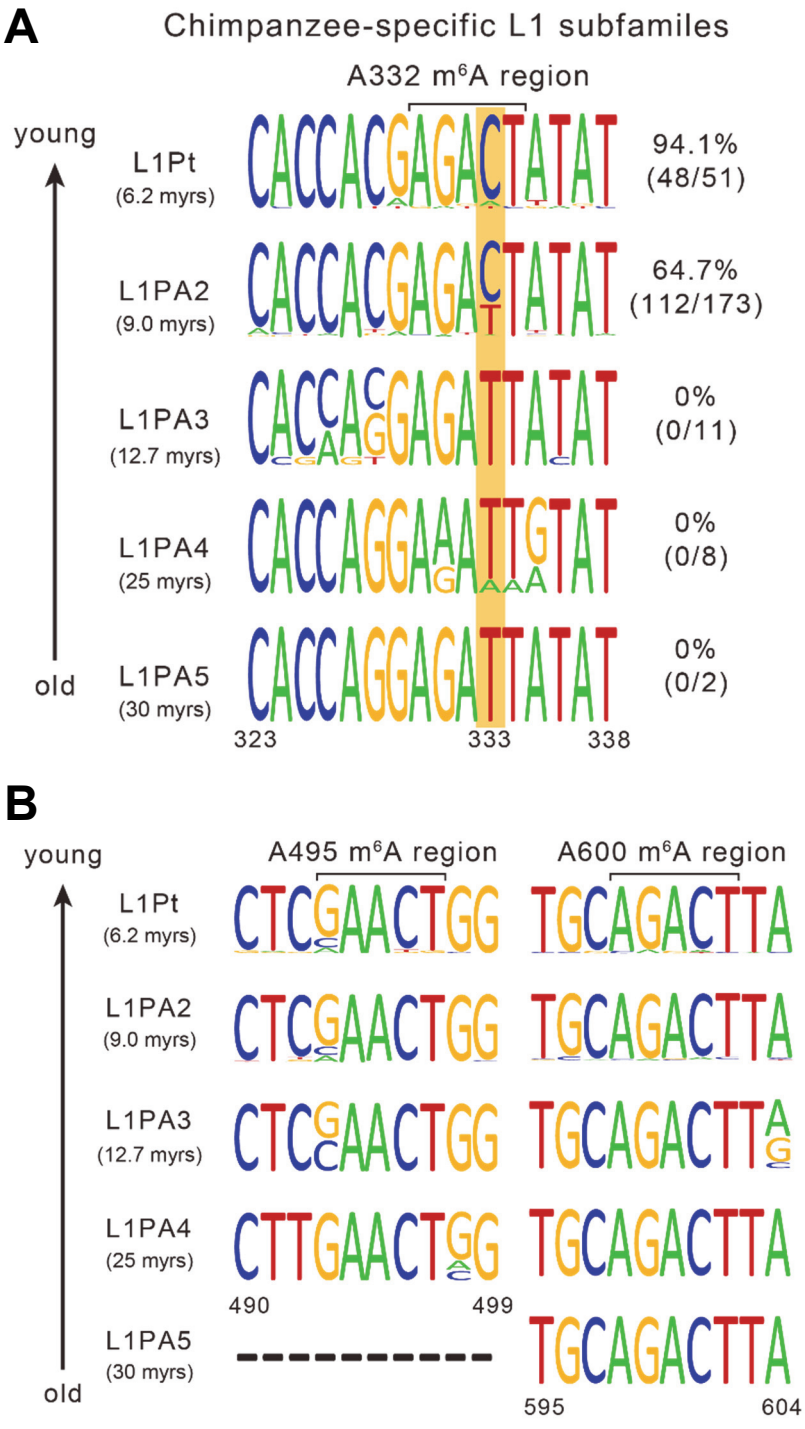
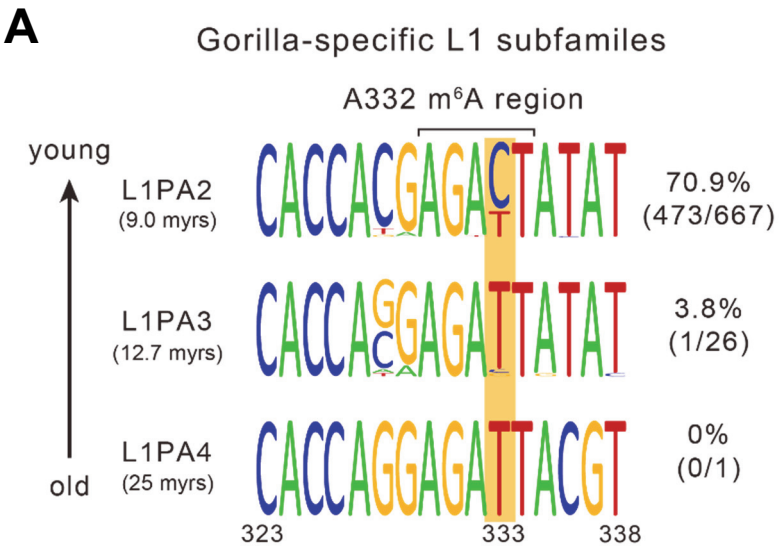

A
Chimpanzee-specific L1 subfamiles
A332 m6A region
young
old
L1Pt
(6.2 myrs)
94.1%
(48/51)
L1PA2
(9.0 myrs)
64.7%
(112/173)
L1PA3
(12.7 myrs)
0%
(0/11)
L1PA4
(25 myrs)
0%
(0/8)
L1PA5
(30 myrs)
0%
(0/2)
323
333
338
B
young
old
A495 m6A region
A600 m6A region
L1Pt
(6.2 myrs)
L1PA2
(9.0 myrs)
L1PA3
(12.7 myrs)
L1PA4
(25 myrs)
L1PA5
(30 myrs)
490
499
595
604

Figure 37. Comparative analysis of the m^{6}A cluster in Chimpanzee-specific full-length L1 subfamilies

(c) Comparative analysis of the A332 m^{6}A region (A323-A338) in chimpanzee-specific full-length L1 subfamilies. The yellow box indicates the T to C substitution site at the 333rd residue. The percentage indicates the proportion of m^{6}A positive L1s to total L1s. (d) Comparative analysis of other m^{6}A regions at the A495 and the A600 residues in chimpanzee-specific full-length L1 subfamilies (left: A495, right: A600). The sequence reads of the A495 region were undefined in L1PA5 subfamilies. Differences in the substitution patterns in different subfamilies were not observable at the two other m^{6}A sites.


A
Gorilla-specific L1 subfamiles
A332 m6A region
young
old
L1PA2
(9.0 myrs)
L1PA3
(12.7 myrs)
L1PA4
(25 myrs)
70.9%
(473/667)
3.8%
(1/26)
0%
(0/1)
323
333
338

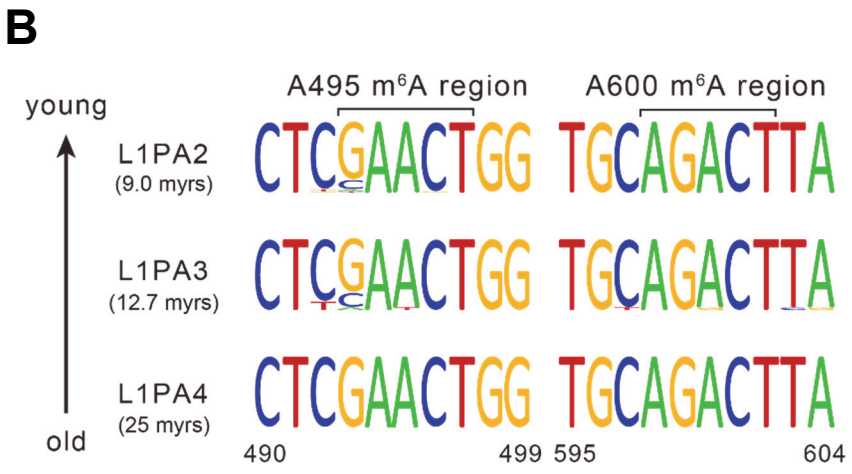

B
A495 m6A region
A600 m6A region
young
old
L1PA2
(9.0 myrs)
L1PA3
(12.7 myrs)
L1PA4
(25 myrs)
490
499
595
604

Figure 38. Comparative analysis of the m^{6}A cluster in Gorilla-specific full-length L1 subfamilies

(e) Comparative analysis of the A332 m^{6}A region (A323-A338) in gorilla-specific full-length L1 subfamilies. The yellow box indicates the T to C substitution site at the 333rd residue. The percentage indicates the proportion of m^{6}A positive L1s to total L1s. (f) Comparative analysis of other m^{6}A regions at the A495 and the A600 residues in gorilla-specific full-length L1 subfamilies (left: A495, right: A600). Differences in the substitution patterns in different subfamilies were not observable at the two other m^{6}A sites. The age of the L1 lineage is specified in parentheses. The height of the nucleotide indicates the frequency of the corresponding nucleotide.

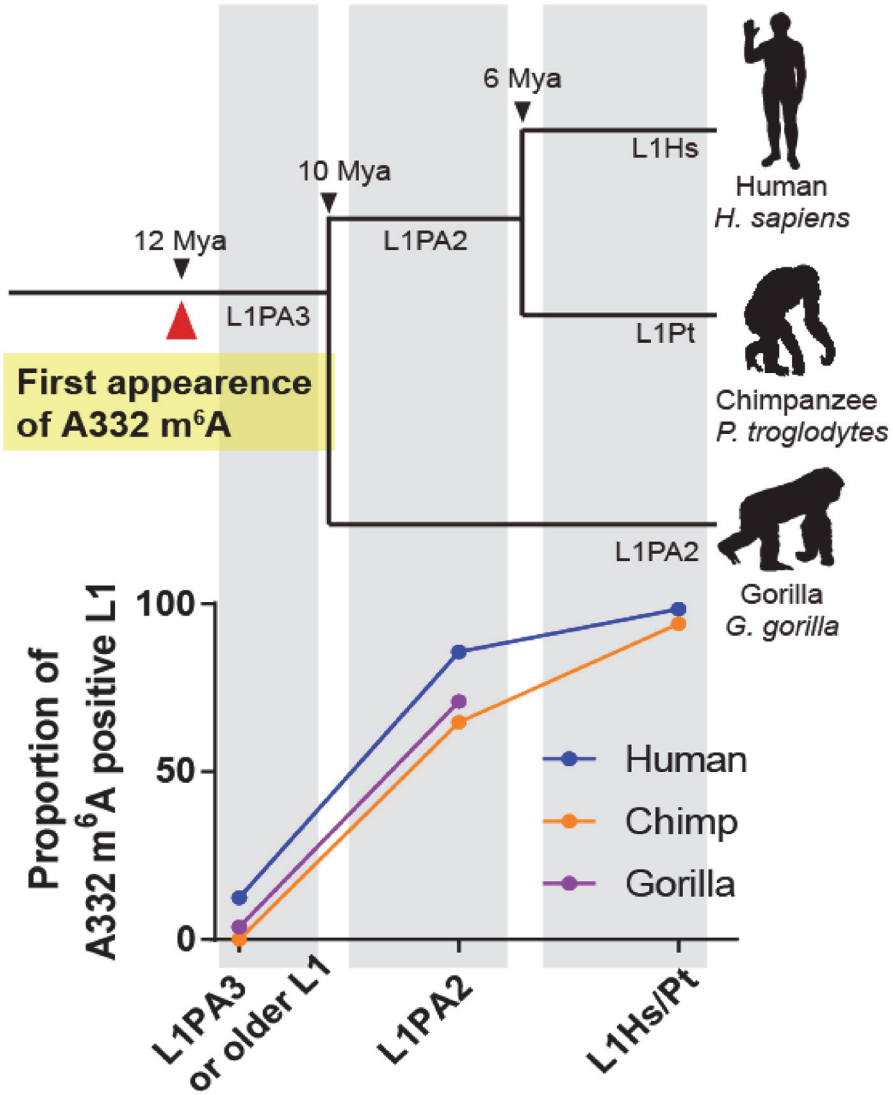

6 Mya
10 Mya
12 Mya
L1Hs
L1PA2
L1PA3
L1Pt
First appearence
of A332 m6A
Human
H. sapiens
Chimpanzee
P. troglodytes
L1PA2
Gorilla
G. gorilla
Proportion of
A332 m6A positive L1
100
50
0
Human
Chimp
Gorilla
L1PA3
or older L1
L1PA2
L1Hs/Pt

Figure 39. Changes in A332 m^{6}A motif proportion during primate evolution

Comparative analysis of L1 A332 m^{6}A sites in species-specific full-length L1s from three different primates. Phylogenetic tree of gorilla, chimpanzee, and human L1s with predicted age and the corresponding L1 subfamily lineages (upper panel). Plot showing changes in the A332-m^{6}A motif region from L1PA3 or older L1s to L1PA2 and a younger L1 of three different primates (lower panel).

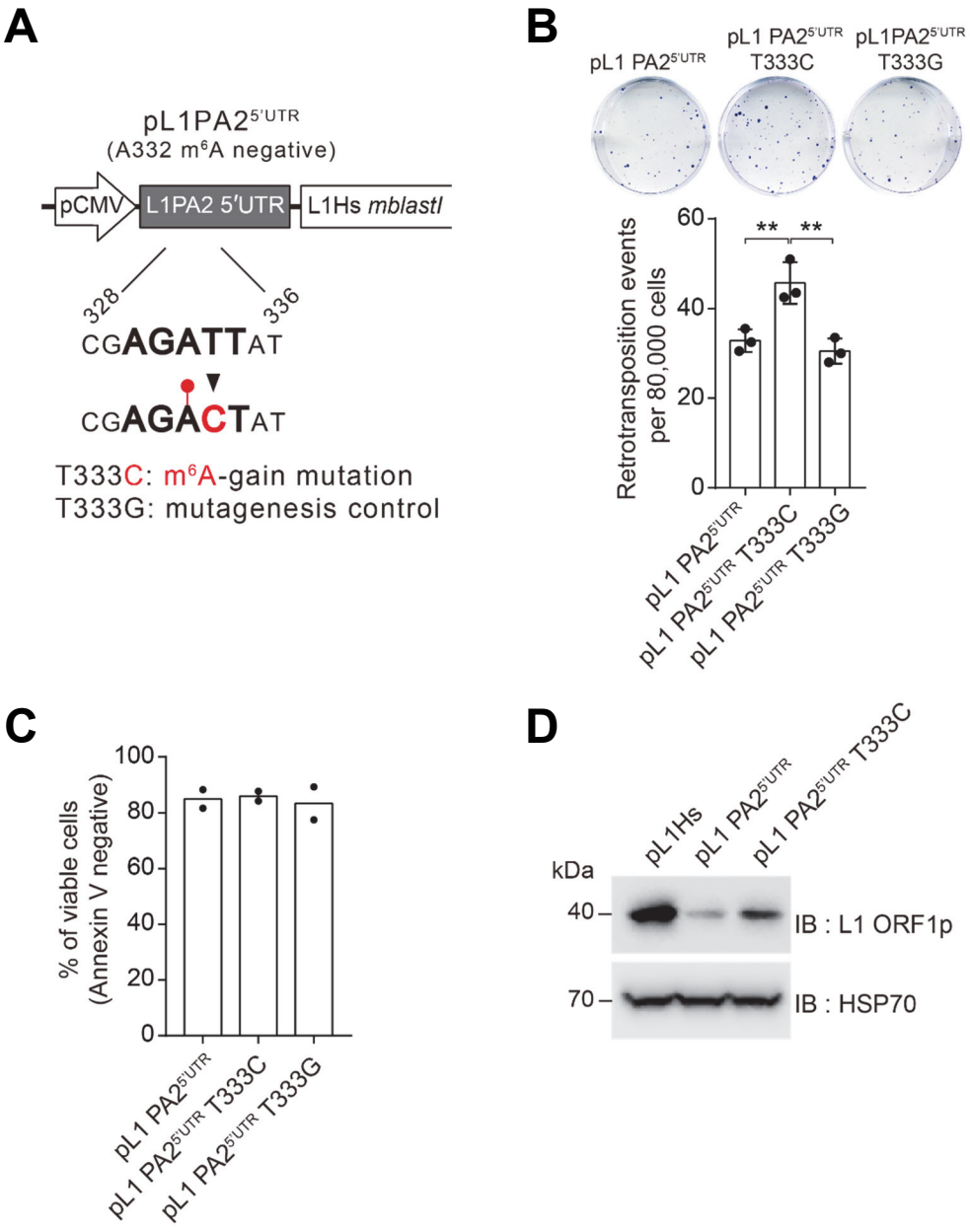

A
pL1PA2 5'UTR
(A332 m6A negative)
pCMV
L1PA2 5'UTR
L1Hs mblastI
328
336
CGAGATTAT
CGAGACTAT
T333C: m6A-gain mutation
T333G: mutagenesis control
B
pL1 PA2 5'UTR
pL1 PA2 5'UTR T333C
pL1PA2 5'UTR T333G
Retrotransposition events per 80,000 cells
60
40
20
0
**
**
pL1 PA2 5'UTR
pL1 PA2 5'UTR T333C
pL1 PA2 5'UTR T333G
C
% of viable cells (Annexin V negative)
100
80
60
40
20
0
pL1 PA2 5'UTR
pL1 PA2 5'UTR T333C
pL1 PA2 5'UTR T333G
D
pL1Hs
pL1 PA2 5'UTR
pL1 PA2 5'UTR T333C
kDa
40
IB : L1 ORF1p
70
IB : HSP70

Figure 40. The impact of A332 m^{6}A acquisition in old L1 activity

(A) A schematic of retrotransposition assay using pL1PA2$^{5' UTR}$ construct that is generated by substituting 5′ UTR of pL1Hs with A332 m^{6}A negative 5′ UTR of L1PA2. A schematic of T333C m^{6}A acquisition mutagenesis in L1 5′ UTR 328-336 region was indicated in red. (B) Retrotransposition assays for assessing the effect of A332 m^{6}A acquisition in pL1PA2$^{5' UTR}$ with T333C mutation. T333G mutation served as negative control (n = 3 independent samples, mean ± s.d., one-way ANOVA and Tukey's multiple comparison test; **p < 0.01). (C) Viability test of pL1 PA2$^{5'UTR}$- or its mutant constructs-expressing HeLa cells through Annexin V assay (n = 2 independent samples, mean of two replicates). (D) Immunoblot assay showing L1 ORF1p expression in the indicated pL1-transfected HeLa cells. HSP70 served as a loading control. The immunoblot images are representative of three independent experiments.

5. DISCUSSION

The role of m^6A modification in pathogenic viral transcripts has been reported in the past decade (Dang et al., 2019). However, the role of m^6A in L1s as genomic parasites have been poorly understood. In this study, I demonstrated that the proper formation of the m^6A cluster in 5' UTR of L1 RNA is essential for L1 retrotransposition. The evolutionary history of the m^6A cluster in primate-specific L1s revealed the most influential m^6A region (A332) that was obtained in the past 12 million years. This suggests the potential role of m^6A as a driving force in L1 evolution (Figure 41).

Two recent studies have revealed that the m^6A modification decreases the stability of L1 RNA with respect to R-loop or chromatin regulation (Abakir *et al.*, 2020; Liu *et al.*, 2020). However, I revealed that the L1 5′ UTR m^6A cluster did not affect RNA stability but promoted translation. Abakir *et al*., and Liu *et al.* observed the role of m^6A in genome-wide L1 repetitive elements, which are mostly inactive by 5′ truncations or inversions (Brouha *et al.*, 2003; Sassaman et al., 1997). Considering that this study focused on the functions of m^6A in the replication cycle of retrotransposition-competent L1s, which have intact 5′ UTR, this difference in the scope of L1 RNA types may contribute to the discrepancy.

Indeed, m^{6}A enzymes regulate L1 expression only when L1 contains its 5′ UTR. The presence of 5′ UTR in L1 transcripts affects retrotransposition efficiency (An *et al.*, 2011). Despite the unique characteristics of L1 5′ UTR that is lengthy, GC rich, and exhibits promoter activity, its regulatory function at the post-transcriptional level has posed a long-standing question. My findings demonstrated that L1 5′ UTR m^{6}A modification is essential for L1 translation, L1 RNP formation, and thus retrotransposition. Therefore, I provide a new perspective on the regulatory function of L1 5′ UTR as a hub for RNA modification.

I demonstrated that m^{6}A promotes not only ORF1p production via enhancing the translational efficiency, but also L1 cDNA synthesis. Since ORF2p can proceed reverse transcription regardless of association with ORF1p (Doucet *et al.*, 2010; Kulpa and Moran, 2006), it remains to clarify whether m^{6}A modification upregulates ORF2p translation or m^{6}A- modified L1 RNA indirectly influences reverse transcriptase activity of ORF2p. The unconventional translational mechanism of ORF2p, which relies on the translation of the upstream ORF (Alisch *et al.*, 2006), suggests that enhanced ORF1p translation rates by m^{6}A cluster successively stimulate ORF2p synthesis. In addition, m^{6}A modifications could alter RNA-protein interactome (Arguello

et al., 2017; Perez-Perri et al., 2018) or RNA secondary structure (Liu et al., 2015), which might affect L1 ORF2p enzymatic activity. Therefore, future studies could reveal the role of m^6A in ORF2p regulation. By adopting a microscopic approach, I confirmed that m^6A is critical for the formation of L1 RNP aggregates. The rate of ORF1p oligomerization is the limiting factor in the production of successful L1 RNPs (Naufer *et al.*, 2015). Therefore, I speculated that 5′ UTR m^6As enable L1 RNA to produce sufficient ORF1p, which further accelerates the oligomerization of ORF1p. Since the process of L1 RNP formation is more complicated than the biochemical interaction between L1 RNA and its protein, the process by which m^6A orchestrates the assembly of retrotransposition-competent L1 RNP remains to be understood.

eIF3 recognizes an m^6A residue in the 5′ UTR and promotes the translation of mRNAs (Meyer *et al.*, 2015). I assumed that the L1 5′ UTR recruits eIF3 to the m^6A cluster for efficiently translating the L1 mRNA. Indeed, eIF3 PAR-CLIP-seq data reveal the interaction between eIF3 and L1 5′ UTR m^6A residue. I also demonstrated that the eIF3-bound portion of L1 decreases in the absence of the 5′ UTR m^6A. A single m^6A residue is sufficient to induce eIF3-mediated translation (Meyer *et al.*, 2015). This could explain the synergetic effects of triple m^6A residues in L1 5′ UTR, which suggests that each m^6A residue can

serve as a docking site for eIF3. Moreover, under cellular stress, the 5′ UTR m^6A facilitates the cap-independent translation of mRNA (Coots *et al.*, 2017; Meyer *et al.*, 2015; Zhou *et al.*, 2015). These studies raise the possibility that m^6A initiates the cap-independent translation of L1 RNA. Although Dmitriev *et al*. revealed that human L1 mRNA is translated in a cap-dependent manner, m^6A modification was not considered in their experiments (Dmitriev et al., 2007). Although I confirmed that heat shock stress promotes L1 RNA translation via 5′ UTR m^6A cluster, in future studies, it is important to determine whether m^6A modification enables the cap-independent translation of L1 and whether m^6A acts as a molecular switch for L1 expression under cellular stress.

L1s have been continuously active since the origin of mammals (Boissinot et al., 2004). One of the previous studies on L1 evolution revealed that several distinct L1 lineages coexisted and were in a simultaneously activated state in the ancestral primate genome. However, since the emergence of the L1PA lineage, the L1 subfamily has evolved and maintained itself as a single lineage in the last 25 million years of the evolution of human and its close relatives (Khan *et al.*). The study proposed that the competition between or coexistence of L1 lineages is determined by the status of the 5′ UTR of L1s and acquisition of

novel 5′ UTR is a fundamental feature in mammalian L1 evolution (Khan *et al.*, 2006). Given that m^6A methyltransferase marks m^6A in a sequence-specific manner, the accumulation of mutations in L1 might cause the loss or acquisition of putative m^6A motifs. To further elucidate the history of m^6A in L1 evolution, I analyzed species-specific full-length L1s from the human, chimpanzee, and gorilla genome. Notably, the A332 m^6A motif first appeared in L1PA3 or older L1 lineages more than 12 million years ago. During the evolution of the three different primates, humans, chimpanzees, and gorillas, the A332 m^6A-positive L1s have propagated their progenies and have become the dominant L1 subfamilies. As m^6A modification promotes L1 mobility, the acquisition of m^6A would have resulted in the positive selection of A332 m^6A-containing L1s. Over the extended periods of L1 evolution, L1s have competed for survival against host restriction (Castro-Diaz et al., 2014a; Jacobs *et al.*, 2014; Khan *et al.*, 2006). KRAB-zinc finger proteins (ZFP), which have evolved with L1s, suppress the old L1 transcription in a sequence-specific manner (Castro-Diaz *et al.*, 2014a; Jacobs *et al.*, 2014). However, L1Hs, which is the youngest L1 lineage in the human genome, escapes KRAB-ZFP restriction and is not recognized by any KRAB-ZFPs (Castro-Diaz *et al.*, 2014a). Instead, the host defense utilizes post-transcriptional suppression mechanisms, such as small RNA interference (e.g., piRNA) or APOBECs, to restrict the replication of L1s

(Marchetto et al., 2013); however, the youngest L1s are still active. My findings provide clues on how the youngest L1s continuously replicate under host surveillance. The emergence and the propagation of the A332 m^6A-positive L1s suggest that 5′ UTR m^6A modification was a countermeasure against the host post-transcriptional restriction.

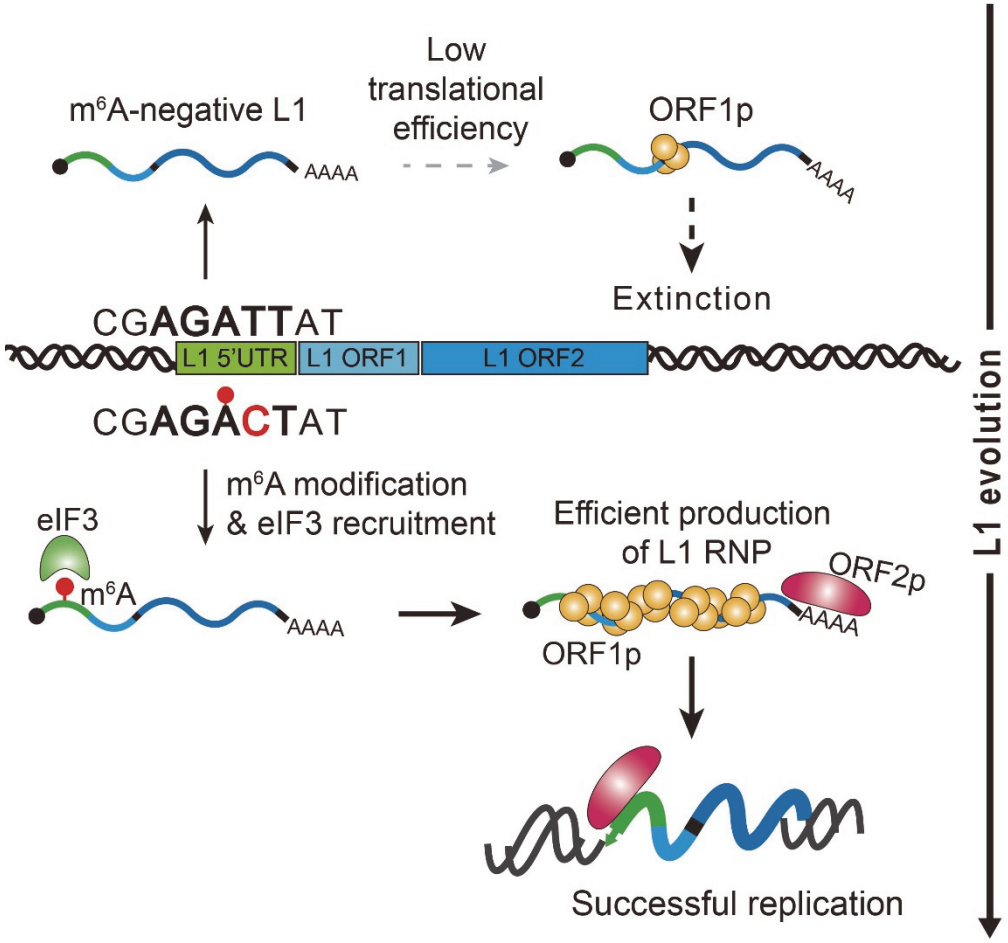

m6A-negative L1
AAAA
Low
translational
efficiency
ORF1p
AAAA
Extinction
CGAGATTAT
L1 5'UTR
L1 ORF1
L1 ORF2
CGAGACTAT
m6A modification
& eIF3 recruitment
eIF3
m6A
AAAA
Efficient production
of L1 RNP
ORF2p
AAAA
ORF1p
Successful replication
L1 evolution

Figure 41. m^{6}A is a driving force in L1 evolution.

A proposed model for the role of m^{6}A in L1 replication and evolution. Full-length young L1s RNA have m^{6}A in 5′ UTR A332 residue (lower part of scheme, m^{6}A in red circle and m^{6}A-gain mutation at T333C in red). 5′ UTR m^{6}A cluster recruits eIF3 complex (green) and promotes translation of L1 ORF1p (yellow). Increased L1 ORF1p synthesis leads efficient production of L1 ribonucleotide particle (RNP) formation with its parental mRNA (line with poly A) and the reverse transcriptase, L1 ORF2p (magenta). The L1 RNP enters the nucleus and then generate the progeny through insertion of its cDNA. Old L1s with no A332 m^{6}A motif have lower efficiency of translation and replication than those of A332 m^{6}A-positive L1s (upper part of scheme). Since the A332 m^{6}A motif first appeared ~12 million years ago, m^{6}A-stimulated L1 replication has allowed m^{6}A-positive L1s survive during evolution, but made old L1s out of competition.

6. REFERENCES

Abakir, A., Giles, T.C., Cristini, A., Foster, J.M., Dai, N., Starczak, M., Rubio-Roldan, A., Li, M., Eleftheriou, M., Crutchley, J., et al. (2020). N6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. Nature Genetics *52*, 48-55. 10.1038/s41588-019-0549-x.

Alarcon, C.R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S.F. (2015). N6-methyladenosine marks primary microRNAs for processing. Nature *519*, 482-485. 10.1038/nature14281.

Alisch, R.S., Garcia-Perez, J.L., Muotri, A.R., Gage, F.H., and Moran, J.V. (2006). Unconventional translation of mammalian LINE-1 retrotransposons. Genes & development *20*, 210-224. 10.1101/gad.1380406.

An, W., Dai, L., Niewiadomska, A.M., Yetil, A., O'Donnell, K.A., Han, J.S., and Boeke, J.D. (2011). Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. Mobile DNA *2*, 2. 10.1186/1759-8753-2-2.

Arguello, A.E., DeLiberto, A.N., and Kleiner, R.E. (2017). RNA Chemical Proteomics Reveals the N6-Methyladenosine (m6A)-Regulated Protein–RNA Interactome. Journal of the American Chemical Society *139*, 17249-17252. 10.1021/jacs.7b09213.

Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M.,

Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell *15*, 707-719. 10.1016/j.stem.2014.09.019.

Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M., and Moran, J.V. (2010). LINE-1 retrotransposition activity in human genomes. Cell *141*, 1159-1170. 10.1016/j.cell.2010.05.021.

Boissinot, S., Roos, C., and Furano, A.V. (2004). Different Rates of LINE-1 (L1) Retrotransposon Amplification and Evolution in New World Monkeys. Journal of Molecular Evolution *58*, 122-130. 10.1007/s00239-003-2539-x.

Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V., and Kazazian, H.H., Jr. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A *100*, 5280-5285. 10.1073/pnas.0831042100.

Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., Duc, J., Jang, S.M., Turelli, P., and Trono, D. (2014a). Evolutionally dynamic L1 regulation in embryonic stem cells. Genes & development *28*, 1397-1409. 10.1101/gad.241661.114.

Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., Duc, J., Jang, S.M., Turelli, P., and Trono, D. (2014b). Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev *28*,

1397-1409. 10.1101/gad.241661.114.

Coots, R.A., Liu, X.-M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X., and Qian, S.-B. (2017). m(6)A Facilitates eIF4F-Independent mRNA Translation. Molecular cell *68*, 504-514.e507. 10.1016/j.molcel.2017.10.002.

Cost, G.J., Feng, Q., Jacquier, A., and Boeke, J.D. (2002). Human L1 element target-primed reverse transcription in vitro. EMBO J *21*, 5899-5910. 10.1093/emboj/cdf592.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res *14*, 1188-1190. 10.1101/gr.849004.

Dang, W., Xie, Y., Cao, P., Xin, S., Wang, J., Li, S., Li, Y., and Lu, J. (2019). N6-Methyladenosine and Viral Infection. Frontiers in Microbiology *10*. 10.3389/fmicb.2019.00417.

Dewannieux, M., Harper, F., Richaud, A., Letzelter, C., Ribet, D., Pierron, G., and Heidmann, T. (2006). Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. Genome Res *16*, 1548-1556. 10.1101/gr.5565706.

Dmitriev, S.E., Andreev, D.E., Terenin, I.M., Olovnikov, I.A., Prassolov, V.S., Merrick, W.C., and Shatsky, I.N. (2007). Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal

ribosome entry site mediated. Mol Cell Biol *27*, 4685-4697. 10.1128/mcb.02138-06.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2012). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.

Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. Nat Protoc *8*, 176-189. 10.1038/nprot.2012.148.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201-206. 10.1038/nature11112.

Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.V., and Gilbert, N. (2010). Characterization of LINE-1 Ribonucleoprotein Particles. PLOS Genetics *6*, e1001150. 10.1371/journal.pgen.1001150.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797. 10.1093/nar/gkh340.

Ergun, S., Buschmann, C., Heukeshoven, J., Dammann, K., Schnieders, F.,

Lauke, H., Chalajour, F., Kilic, N., Stratling, W.H., and Schumann, G.G. (2004). Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. J Biol Chem *279*, 27753-27763. 10.1074/jbc.M312985200.

Feng, Q., Moran, J.V., Kazazian, H.H., Jr., and Boeke, J.D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell *87*, 905-916. 10.1016/s0092-8674(00)81997-2.

Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m(6)A RNA methylation. Nat Rev Genet *15*, 293-306. 10.1038/nrg3724.

Furano, A.V.a.B., S. (2008). Long Interspersed Nuclear Elements (LINEs): Evolution. In In: Encyclopedia of LifeSciences (ELS). 10.1002/9780470015902.a0005304.pub2.

Fustin, J.M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., Isagawa, T., Morioka, M.S., Kakeya, H., Manabe, I., and Okamura, H. (2013). RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell *155*, 793-806. 10.1016/j.cell.2013.10.026.

Garcia-Perez, J.L., Marchetto, M.C.N., 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. Human Molecular Genetics *16*, 1569-1577.

10.1093/hmg/ddm105.

Gokhale, N.S., McIntyre, A.B.R., McFadden, M.J., Roder, A.E., Kennedy, E.M., Gandara, J.A., Hopcraft, S.E., Quicke, K.M., Vazquez, C., Willer, J., et al. (2016). N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. Cell Host Microbe *20*, 654-665. 10.1016/j.chom.2016.09.015.

Goodier, J.L. (2016). Restricting retrotransposons: a review. Mob DNA *7*, 16. 10.1186/s13100-016-0070-z.

Goodier, J.L., and Kazazian, H.H., Jr. (2008). Retrotransposons revisited: the restraint and rehabilitation of parasites. Cell *135*, 23-35. 10.1016/j.cell.2008.09.022.

Goodier, J.L., Ostertag, E.M., Engleka, K.A., Seleme, M.C., and Kazazian, H.H., Jr. (2004). A potential role for the nucleolus in L1 retrotransposition. Hum Mol Genet *13*, 1041-1048. 10.1093/hmg/ddh118.

Grimaldi, G., Skowronski, J., and Singer, M.F. (1984). Defining the beginning and end of KpnI family segments. EMBO J *3*, 1753-1759.

Hohjoh, H., and Singer, M.F. (1996). Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. EMBO J *15*, 630-639.

Holcik, M., and Sonenberg, N. (2005). Translational control in stress and apoptosis. Nature Reviews Molecular Cell Biology *6*, 318-327. 10.1038/nrm1618.

Jacobs, F.M., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A.D., Katzman, S., Paten, B., Salama, S.R., and Haussler, D. (2014). An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. Nature *516*, 242-245. 10.1038/nature13760.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G., and He, C. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol *7*, 885-887. 10.1038/nchembio.687.

Jurka, J. (1997). Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc Natl Acad Sci U S A *94*, 1872-1877. 10.1073/pnas.94.5.1872.

Kapitonov, V.V., Tempel, S., and Jurka, J. (2009). Simple and fast classification of non-LTR retrotransposons based on phylogeny of their RT domain protein sequences. Gene *448*, 207-213. 10.1016/j.gene.2009.07.019.

Khan, H., Smit, A., and Boissinot, S. (2006). Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome Res *16*, 78-87. 10.1101/gr.4001406.

Kim, B., and Kim, V.N. (2019). fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA. Methods *152*, 3-11. 10.1016/j.ymeth.2018.06.004.

Klawitter, S., Fuchs, N.V., Upton, K.R., Munoz-Lopez, M., Shukla, R., Wang, J., Garcia-Canadas, M., Lopez-Ruiz, C., Gerhardt, D.J., Sebe, A., et al. (2016). Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. Nat Commun *7*, 10286. 10.1038/ncomms10286.

Kulpa, D.A., and Moran, J.V. (2005). Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. Hum Mol Genet *14*, 3237-3248. 10.1093/hmg/ddi354.

Kulpa, D.A., and Moran, J.V. (2006). Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat Struct Mol Biol *13*, 655-660. 10.1038/nsmb1107.

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. 10.1038/35057062.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357-359. 10.1038/nmeth.1923.

Lee, A.S.Y., Kranzusch, P.J., and Cate, J.H.D. (2015). eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. Nature *522*, 111-114. 10.1038/nature14267.

Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G.M., Bansal, V., Wang, Y.,

Mason, C.E., and Rana, T.M. (2016). Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. Nat Microbiol *1*, 16011. 10.1038/nmicrobiol.2016.11.

Lionnet, T., Czaplinski, K., Darzacq, X., Shav-Tal, Y., Wells, A.L., Chao, J.A., Park, H.Y., de Turris, V., Lopez-Jones, M., and Singer, R.H. (2011). A transgenic mouse for in vivo detection of endogenous labeled mRNA. Nat Methods *8*, 165-170. 10.1038/nmeth.1551.

Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M.M., Zhao, S., Shen, B., Gao, Y., Han, D., and He, C. (2020). N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. Science *367*, 580-586. 10.1126/science.aay6018.

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol *10*, 93-95. 10.1038/nchembio.1432.

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature *518*, 560-564. 10.1038/nature14234.

Luan, D.D., Korman, M.H., Jakubczak, J.L., and Eickbush, T.H. (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal

target site: a mechanism for non-LTR retrotransposition. Cell *72*, 595-605. 10.1016/0092-8674(93)90078-5.

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. Molecular and Cellular Biology *31*, 300-316. 10.1128/mcb.00561-10.

Mager, D.L., and Stoye, J.P. (2015). Mammalian Endogenous Retroviruses. Microbiol Spectr *3*, Mdna3-0009-2014. 10.1128/microbiolspec.MDNA3-0009-2014.

Malik, H.S., Burke, W.D., and Eickbush, T.H. (1999). The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol *16*, 793-805. 10.1093/oxfordjournals.molbev.a026164.

Marchetto, M.C.N., Narvaiza, I., Denli, A.M., Benner, C., Lazzarini, T.A., Nathanson, J.L., Paquola, A.C.M., Desai, K.N., Herai, R.H., Weitzman, M.D., et al. (2013). Differential L1 regulation in pluripotent stem cells of humans and apes. Nature *503*, 525-529. 10.1038/nature12686.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 *17*, 3. 10.14806/ej.17.1.200.

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. 10.1128/MCB.21.2.467-475.2001.

Mathias, S.L., Scott, A.F., Kazazian, H.H., Jr., Boeke, J.D., and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. Science *254*, 1808-1810. 10.1126/science.1722352.

Meyer, K.D., and Jaffrey, S.R. (2014). The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol *15*, 313-326. 10.1038/nrm3785.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. Cell *163*, 999-1010. 10.1016/j.cell.2015.10.012.

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell *149*, 1635-1646. 10.1016/j.cell.2012.05.003.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917-927. 10.1016/s0092-8674(00)81998-4.

Morrish, T.A., Gilbert, N., Myers, J.S., Vincent, B.J., Stamato, T.D., Taccioli, G.E., Batzer, M.A., and Moran, J.V. (2002). DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nature Genetics *31*, 159-

165. 10.1038/ng898.

Naufer, M.N., Callahan, K.E., Cook, P.R., Perez-Gonzalez, C.E., Williams, M.C., and Furano, A.V. (2015). L1 retrotransposition requires rapid ORF1p oligomerization, a novel coiled coil-dependent property conserved despite extensive remodeling. Nucleic Acids Research *44*, 281-293. 10.1093/nar/gkv1342.

Ostertag, E.M., and Kazazian, H.H., Jr. (2001). Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. Genome Res *11*, 2059-2065. 10.1101/gr.205701.

Payer, L.M., and Burns, K.H. (2019). Transposable elements in human genetic disease. Nature Reviews Genetics *20*, 760-772. 10.1038/s41576-019-0165-8.

Perez-Perri, J.I., Rogell, B., Schwarzl, T., Stein, F., Zhou, Y., Rettel, M., Brosig, A., and Hentze, M.W. (2018). Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. Nature Communications *9*, 4408. 10.1038/s41467-018-06557-8.

Sassaman, D.M., Dombroski, B.A., Moran, J.V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D., and Kazazian, H.H., Jr. (1997). Many human L1 elements are capable of retrotransposition. Nat Genet *16*, 37-43. 10.1038/ng0597-37.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to

ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675. 10.1038/nmeth.2089.

Smit, A., Hubley, R & Green, P. (2013-2015). RepeatMasker Open-4.0.

Song, H.J., Gallie, D.R., and Duncan, R.F. (1995). m7GpppG cap dependence for efficient translation of Drosophila 70-kDa heat-shock-protein (Hsp70) mRNA. Eur J Biochem *232*, 778-788.

Stein, S.C., and Thiart, J. (2016). TrackNTrace: A simple and extendable open-source framework for developing single-molecule localization and tracking algorithms. Sci Rep *6*, 37947. 10.1038/srep37947.

Swergold, G.D. (1990). Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol Cell Biol *10*, 6718-6729. 10.1128/mcb.10.12.6718.

Tang, W., Mun, S., Joshi, A., Han, K., and Liang, P. (2018). Mobile elements contribute to the uniqueness of human genome with 15,000 human-specific insertions and 14 Mbp sequence increase. DNA Res *25*, 521-533. 10.1093/dnares/dsy022.

Wang, X., Zhao, Boxuan S., Roundtree, Ian A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N6-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell *161*, 1388-1399. https://doi.org/10.1016/j.cell.2015.05.014.

Wang, Y., Li, Y., Toth, J.I., Petroski, M.D., Zhang, Z., and Zhao, J.C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat Cell Biol *16*, 191-198. 10.1038/ncb2902.

Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D., and Moran, J.V. (2001). Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol *21*, 1429-1439. 10.1128/MCB.21.4.1429-1439.2001.

Xie, Y., Rosser, J.M., Thompson, T.L., Boeke, J.D., and An, W. (2011). Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. Nucleic Acids Res *39*, e16. 10.1093/nar/gkq1076.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137. 10.1186/gb-2008-9-9-r137.

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vagbo, C.B., Shi, Y., Wang, W.L., Song, S.H., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell *49*, 18-29. 10.1016/j.molcel.2012.10.015.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock

response. Nature *526*, 591-594. 10.1038/nature15377.

Zhou, Y., Zeng, P., Li, Y.H., Zhang, Z., and Cui, Q. (2016). SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. Nucleic Acids Res *44*, e91. 10.1093/nar/gkw104.

7. ABSTRACT IN KOREAN

인간 유전체 내에는 스스로 복제하며 유전체의 새로운 자리로 끼어들어갈 수 있는 성질의 유전자가 있다. 이들은 소위 점핑 (jumping) 유전자, 혹은 트랜스포존이라 불린다. 점핑 유전자 중 레트로트랜스포존 (retrotransposon)은 자신의 유전자를 RNA 중간체를 거치는 방식으로 유전체 내에 자신의 유전자를 복제 및 삽입한다. 이는 유전체의 새로운 돌연변이를 만들어 내며 암이나, 유전 질환 등의 질병을 유발할 수도 있지만, 종의 진화에서도 중요한 역할을 한다.

Long interspersed nuclear element 1 (LINE-1; L1)은 가장 활발한 점핑 유전자이며, 오랜 기간 생존 및 복제로 인해 인간 유전체의 약 17%를 차지한다. L1 은 일종의 기생 유전자로 이들의 발현은 세포 수준에서 인식되어 항-바이러스 작용에 의해 억제된다. 그러나 이와 같은 억제에도 불구하고, 인간 진화 과정 속에서 어떻게 L1 이 살아남아 현재도 활발히 복제 및 점핑을 하는지는 아직 밝혀지지 않았다.

본 연구에서는 유인원 유전체의 L1 서열 변화를 계통학적으로 분석하여, L1 이 어떻게 인간 유전체 내에 성공적으로 기생하게 되었는지를 규명하였다. 먼저, 인간, 침팬지, 고릴라의 공통 조상에서 생성된 L1 돌연변이를 발견하고, 이 돌연변이가 유인원 진화 과정에서 L1 의 생존에 결정적 역할을 하였음을 밝혀냈다.

상기된 L1 의 돌연변이는 L1 유전자의 전령 RNA 에 대해 메틸화 변형을 유도하는 모티프 (motif)를 지니고 있음을 확인하고, RNA 메틸화 효소 METTL3 와 메틸화 제거효소인 ALKBH5 가 L1 증식을 조절하는 것을 발견하였다. 나아가 RNA 메틸화는 단백질 번역 개시 인자인 eIF3 가 L1 RNA 에 대한 접근을 촉진시켜 L1 단백질 생성을 촉진하고, L1 RNA-단백질 복합체 형성을 유도하는 것을 밝혀냈다.

이를 통해 RNA 변형 기작이 유전적인 형질이 되어 L1 점핑 유전자의 생존 및 진화에 기여한 기작을 새롭게 제시하였다.