L1 retrotransposons exploit RNA m6A modification as an evolutionary driving force

Sung-Yeon Hwnag
Seoul National University
Hyunchul Jung
KAIST
Seyoung Mun
Dankook University
Sungwon Lee
Seoul National University
S. Chan Baek
Center for RNA Research, Institute for Basic Science
Hyungseok Moon
Seoul National University
Baekgyu Kim
Center for RNA Research, Institute for Basic Science
Yongkuk Choi
Center for RNA Research, Institute for Basic Science
Young-Hyun Go
Sogang University
Wanxiangfu Tang
Brock University
Jongsu Choi
Dr. von Hauner Children's Hospital https://orcid.org/0000-0001-5007-7469
Jung Kyoon Choi
Korea Advanced Institute of Science and Technology
Hyuk-Jin Cha
Seoul National University https://orcid.org/0000-0001-9277-2662
Hye Yoon Park
Seoul National University https://orcid.org/0000-0003-4704-8178
Ping Liang
Brock University
V. Narry Kim
Institute for Basic Science https://orcid.org/0000-0002-3213-4965
Kyudong Han

Dankook University

Kwangseog Ahn (ksahn@snu.ac.kr)

Seoul National University

Article

Keywords: L1 retrotransposons, mechanisms underlying L1 propagation, RNA m6A modification, METTL3, ALKBH5

Posted Date: July 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-40207/v1

Version of Record: A version of this preprint was published at Nature Communications on February 9th, 2021. See the published version at https://doi.org/10.1038/s41467-021-21197-1.

L1 retrotransposons exploit RNA m^{6}A modification as an evolutionary driving force

Sung-Yeon Hwang[1,2], Hyunchul Jung[3], Seyoung Mun[4], Sungwon Lee[1,2], S. Chan Baek[1,2], Hyungseok C. Moon[5], Baekgyu Kim[1,2], Yongkuk Choi[1,2], Young-Hyun Go[6], Wanxiangfu Tang[8], Jongsu Choi[7], Jung Kyoon Choi[3], Hyuk-Jin Cha[6], Hye Yoon Park[5], Ping Liang[8], V. Narry Kim[1,2], Kyudong Han[4*] and Kwangseog Ahn[1,2*]

[1]Center for RNA Research, Institute for Basic Science, Seoul 08826, Republic of Korea
[2]School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea
[3]Department of Bio and Brain Engineering, KAIST, Daejeon 34141, Republic of Korea
[4]Department of Nanobiomedical Science and BK21 PLUS NBM Global Research Center for Regenerative Medicine, Dankook University, Cheonan, 31116, Republic of Korea.
[5]Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea.
[6]Department of Pharmacy, Seoul National University, Seoul 08826, Republic of Korea
[7]Dr. von Hauner Children's Hospital, Ludwig-Maximilians-University of Munich, Lindwurmstrasse 4, D-80337 Munich, Germany.
[8]Department of Biological Sciences, Brock University, St. Catharines, Ontario, Canada
[*]Correspondence: ksahn@snu.ac.kr (K.A.), jim97@dankook.ac.kr (K.H.)

ABSTRACT

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, we propose a novel survival strategy of L1, which exploits RNA m^{6}A modification. We discover that m^{6}A 'writer' METTL3 facilitates L1 retrotransposition, whereas m^{6}A 'eraser' ALKBH5 suppresses it. The essential m^{6}A 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, through the comparative analysis of human- and primate-specific L1 lineages, we find that the most functional m^{6}A motif-containing L1s have been positively selected and became a distinctive feature of evolutionarily young L1s. Thus, our findings demonstrate that L1 retrotransposons hijack RNA m^{6}A modification system for its successful replication.

INTRODUCTION

Long interspersed element-1 (L1) is currently an active autonomous retrotransposon, and constitutes ~17% of the human genome[1]. The average human genome contains 80-100 copies of retrotransposition-competent L1s[2,3]. A retrotransposition-competent L1 is 6 kb in length and consists of a 5′ untranslated region (UTR) containing an internal promoter[4], two open reading frames (ORF1 and ORF2), and a short 3′ UTR. ORF1 encodes a nucleic acid chaperon protein (ORF1p)[5], while ORF2 encodes a protein with endonuclease and reverse transcriptase activity (ORF2p)[6,7]. ORF1p and ORF2p associate preferentially with their parental mRNA to form an L1 ribonucleoprotein (RNP) particle[8]. The L1 RNP enters the nucleus and then generate the progeny through de novo insertion of its cDNA[9,10]. The mobility of L1s contributed to a source of genetic variation, but also pose a threat to genome integrity[11]. Although several host factors have evolved to suppress L1 retrotransposition, the youngest L1 subfamilies are still active and replicated continuously[12-14]. To date, however, the mechanism of how L1s have propagated under host surveillance remains unknown.

N6-methyladenosine (m^6A) is the most prevalent internal modification in eukaryotic mRNAs, which determines RNA function and fate[15]. 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 DRA^{m}CH (where D = G/A/U, R = G/A, and H = U/C/A)[16,17]. Conversely, m^6A is removed by the demethylases α-ketoglutarate-dependent dioxygenase AlkB homolog 5 (ALKBH5) or fat mass and obesity-associated protein (FTO)[18,19]. Emerging studies have revealed that m^6A modification in viral transcripts affect the gene expression and replication of viruses such as HIV-1[20]. Despite the critical role of m^6A in pathogenic viral transcripts, it remains unclear whether m^6A participates in the regulation of the endogenous parasites, L1 retrotransposons.

Here, we show that L1 retrotransposon exploits m^6A modification to facilitate its mobility. We figured out that m^6A machinery plays a role in L1 regulation and identified the functional m^6A cluster located on 5′ UTR of retrotransposition-competent full-length L1. Our results show that L1 5′ UTR m^6A cluster recruits eukaryotic initiation factor 3 (eIF3) for efficient translation and promotes the formation of L1 RNP, which are essential for L1 mobility. Lastly, we traced a recent episode of human- and primate-specific L1 evolution and revealed that the most functional m^6A site (A332 residue in L1 5′ UTR) first appeared ~12 million years ago. During the primate evolution, A332 m^6A-positive L1s have been selected and became a distinctive feature of evolutionarily

young L1s, which suggests that the acquisition of m^{6}A motif has acted as an evolutionary driving force for L1 retrotransposons.

RESULTS

METTL3 and ALKBH5 regulate L1 retrotransposition

To determine whether RNA m^{6}A modification affects L1 retrotransposition, we evaluated the effects of the RNA m^{6}A machinery on L1 retrotransposition using a cell-based engineered L1-reporter assay[21]. For the assay, we 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[22] (Fig. 1A). When L1 is successfully integrated into the host chromosome, the cells acquire resistance to blasticidin (Fig. 1A).

We 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 (Fig. 1B). Conversely, the silencing of ALKBH5 increased L1 mobility, while the silencing of FTO did not affect L1 retrotransposition (Fig. 1B). The depletion of the m^{6}A machinery did not vitiate cell viability (Supplementary Fig. 1A). In a reciprocal experiment, we performed an L1 retrotransposition assay with the ectopic expression of RNA m^{6}A 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 (Fig. 1C). We hypothesized that ALKBH5 may function as an L1 restriction factor by removing essential m^{6}A for L1 mobility. To examine whether the enzymatic function of ALKBH5 is critical for L1 mobility suppression, we 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 (Fig. 1D). The viability of transfected cells remained unaffected (Supplementary Fig. 1B).

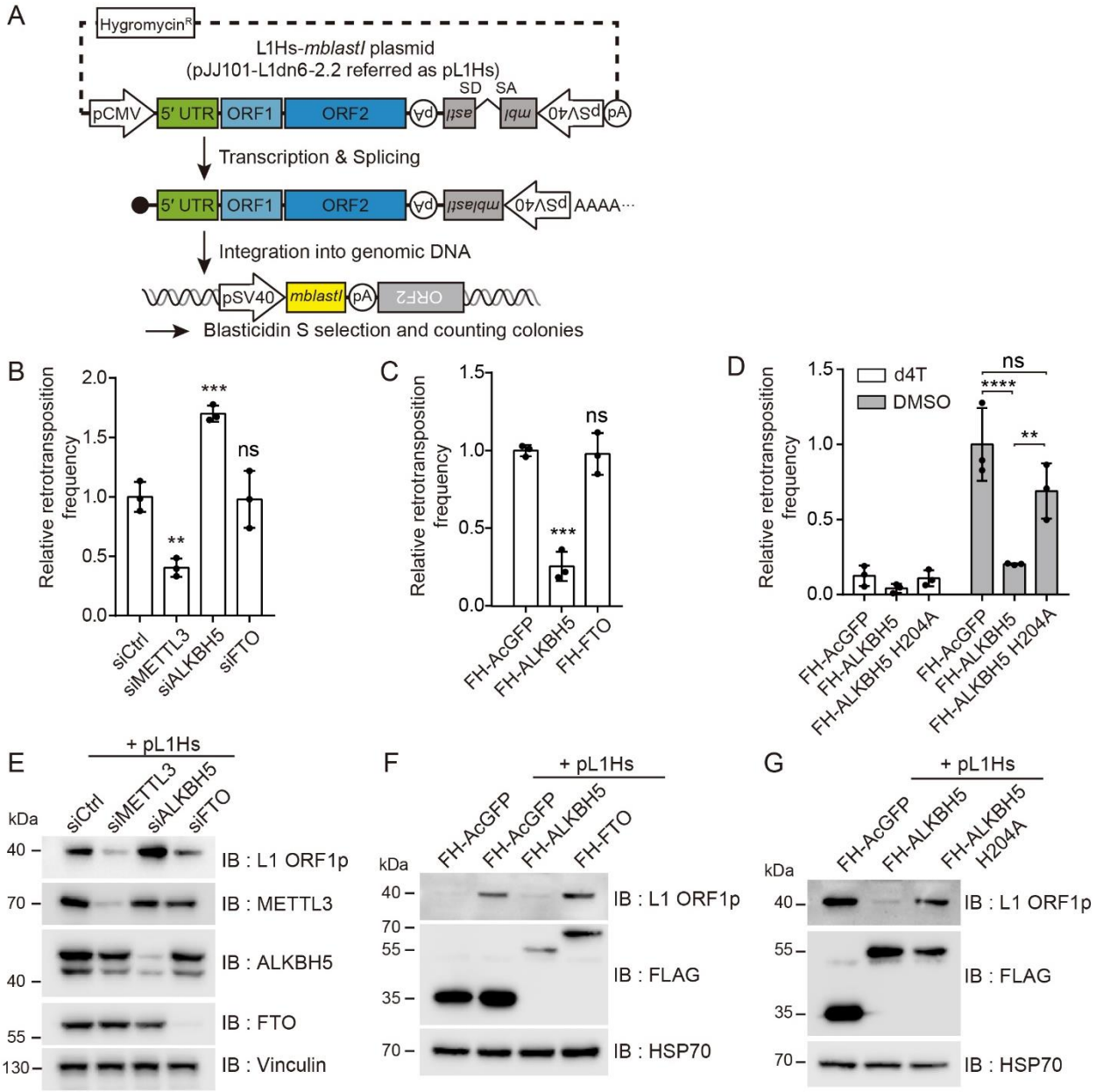

Figure 1. RNA methylation machinery controls L1 retrotransposition. **(A)** A schematic of the L1 construct and an overview of the L1 retrotransposition assay using engineered human L1 construct. **(B)** Retrotransposition assay in HeLa cells treated with siRNA that targets METTL3, ALKBH5, or FTO. A non-targeting siRNA (siCtrl) was used as a control. The results of independent experiments were normalized relative to those of the control. **(C)** Retrotransposition assays performed by co-transfecting the pL1Hs expression cassette with the indicated m^{6}A enzyme-expressing vectors into HeLa cells. Retrotransposition frequency was normalized to that of the control (FH-AcGFP). **(D)** 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, error = 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). **(E)** 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. **(F and G)** Immunoblot assay using pL1Hs-expressing HeLa cells. AcGFP, ALKBH5, FTO, or ALKBH5^{H204A} 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 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^{6}A modification, we 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 us to assess L1 mobility by measuring luminescence as previously reported[23] (Supplementary Fig. 1C). Overexpression of ALKBH5 impaired L1 retrotransposition, regardless of the presence of the CMV promoter (Supplementary Fig. 1D). In line with this result, depletion of METTL3 or ALKBH5 regulates L1 mobility, whereas FTO knockdown did not affect (Supplementary Fig. 1E). These results indicate that ALKBH5-specific m^{6}A substrates are necessary for L1 expansion. To summarize, our data support the functional role of the m^{6}A machinery in regulating L1 retrotransposition.

RNA m^{6}A metabolism regulates gene expression at post-transcriptional levels. Therefore, we speculated that the m^{6}A machinery would influence the protein expression of L1. Immunoblot analysis of HeLa cells devoid of m^{6}A enzymes revealed that m^{6}A enzymes regulate the expression of ORF1p (Fig. 1E). Overexpression of ALKBH5 inhibited ORF1p expression, while the ectopic overexpression of FTO and ALKBH5^{H204A} did not affect the ORF1p expression (Fig. 1F, G). Neither the depletion of RNA m^{6}A machinery nor the overexpression of ALKBH5 altered the levels of expression of the control EGFP (Supplementary Fig. 1F, G), which indicates that m^{6}A enzymes do not affect transfection efficiency. These results suggest that m^{6}A-mediated L1 regulation affects both retrotransposition and L1 protein expression.

L1 RNA is modified by m^{6}A

Although the possibility of L1 m^{6}A modification was demonstrated in recent studies[24,25], 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, we performed methyl-RNA immunoprecipitation (MeRIP) using human embryonic stem cells (H9 hESCs) that express endogenous L1 at sufficient levels[26]. Through qRT-PCR analysis of the MeRIP eluates, we 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 (Fig. 2A). To minimize bias resulting from primers in L1 RNA detection, we 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 (Fig. 2A). Similar to the endogenously expressed L1 RNA

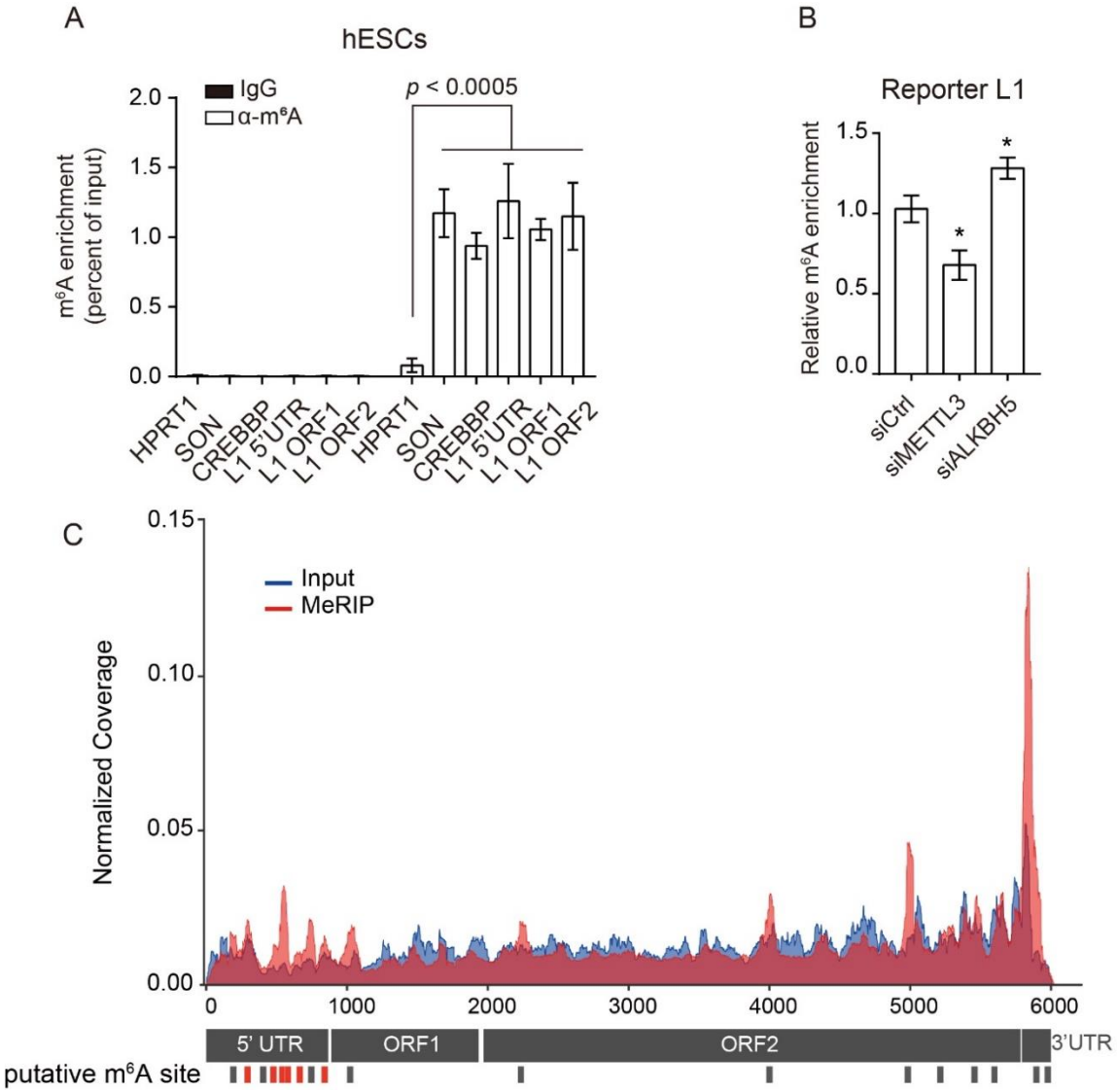

Figure 2. L1 RNA is modified by m^{6}A. **(A)** MeRIP-qPCR analysis of mRNA from H9 hESCs. Eluates from IgG immunoprecipitation served as negative control. Eluted RNA was quantified to determine the percentage of input. (n = 3 independent samples, error = s.e.m., one-way ANOVA and Dunnett's multiple 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, error = s.e.m., unpaired Student's t test, *p < 0.05). **(C)** 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.

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 (Supplementary Fig. 2A). We 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 (Fig. 2B). 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, we analyzed the m^{6}A transcriptome of hESCs reported previously[27] and mapped reads to the consensus sequence of L1Hs, the youngest L1[28]. We identified 18 peaks across the L1Hs sequence using two biological replicates (Fig. 2C). Given that the reads from L1s may yield false-positive results, we segregated peaks containing putative m^{6}A motifs based on the m^{6}A prediction score algorithm (SRAMP)[29]. Six out of eighteen peaks with a high score motif were identified through SRAMP prediction (Supplementary Table 1A). Notably, all six m^{6}A-putative peaks were located in the 5′ UTR.

Next, we mapped the sites of m^{6}A modifications in reporter L1-transfected HeLa cells using MeRIP-seq. Consistent with findings from previous studies[16,17], our 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 (Supplementary Fig. 2B). By mapping the reads on reporter L1, we obtained five candidate peaks (Supplementary Fig. 2C), and further sorted according to the approach based on the m^{6}A prediction as described above. All five peaks were classified as m^{6}A-putative regions with high scores (Supplementary Table 1B). 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^{6}A modification sites commonly detected in endogenous and exogenous L1 RNA are A332 and A839, both located in the 5′ UTR (Supplementary Table 1). This is a notable phenomenon since m^{6}A 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^{6}A modification.

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

Given that the L1 5′ UTR has a potential m^{6}A cluster, we next examined whether the L1 5′ UTR is necessary for m^{6}A-dependent L1 regulation. We 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 (Supplementary Fig. 3A). Furthermore, using the codon-optimized synthetic L1 construct that encoded ORFs with the same amino acids yet different nucleotide sequences, we examined whether alterations in ORF1 and ORF2 nucleotide sequences

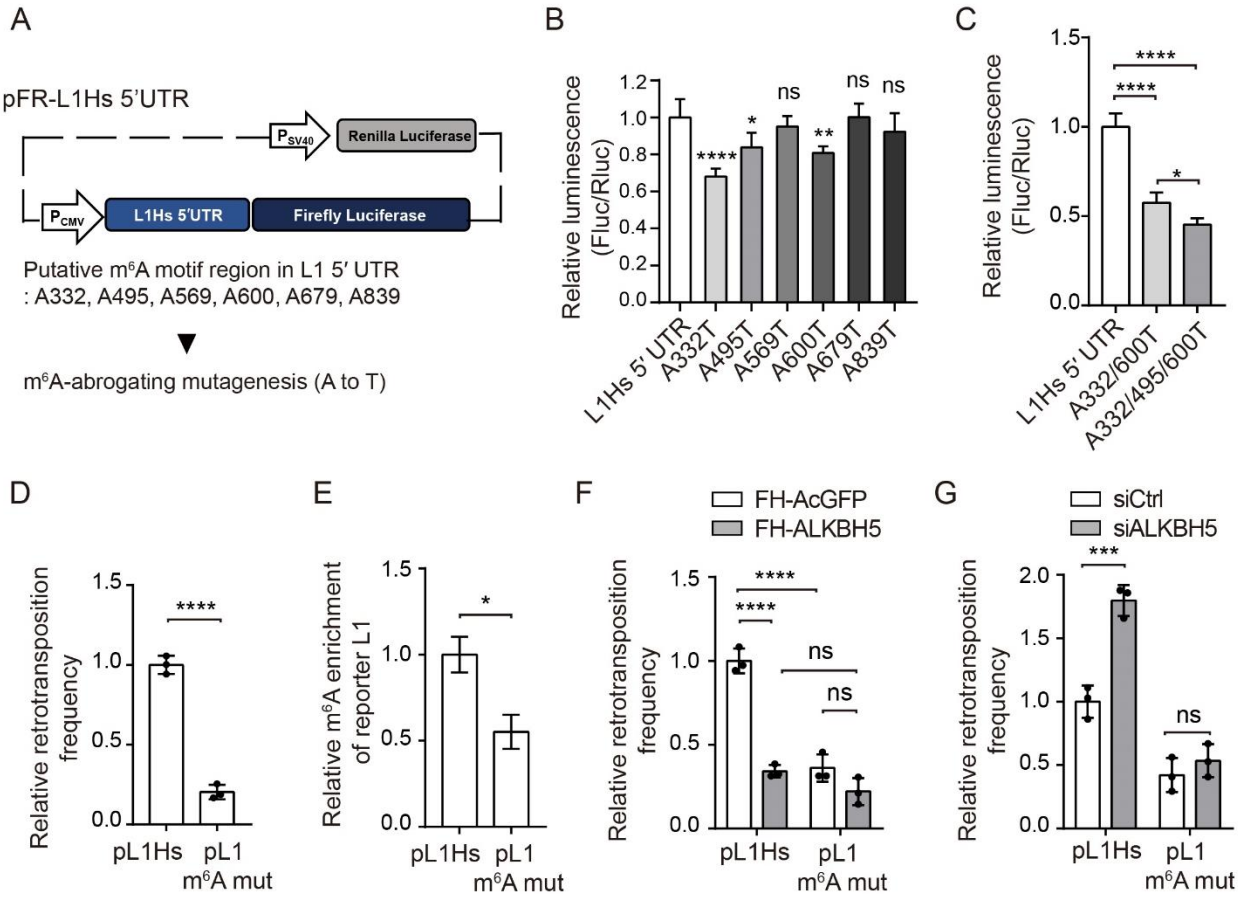

Figure 3. L1 5′ UTR m^{6}A cluster promotes L1 activity. **(A)** Schematic of the dual-luciferase plasmid carrying L1 5′ UTR upstream of the firefly luciferase gene (pFR-L1Hs 5′ UTR). Firefly luciferase luminescence reflected the effect of 5′ UTR and of its mutations. **(B and C)** Dual-luciferase assay 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. (error = s.d., four (B) or five (C) independent samples) **(D)** L1 assays using the triple m^{6}A mutated L1 construct (pL1 m^{6}A mut) in HeLa cells. Retrotransposition frequency was normalized to that of pL1Hs. (n = 3 independent samples, error = s.d.) **(E)** 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, error = s.e.m.) **(F and G)** Retrotransposition assay using pL1Hs-or pL1 m^{6}A mut-expressing HeLa cells ALKBH5 overexpression (F) or silencing (G). L1 mobility was normalized to that in pL1Hs-expressing cells transfected with AcGFP (F) or non-targeting siRNA (G). (n = 3 independent samples, error = s.d.) In (B-G), Statistical significance was calculated by one-way ANOVA with Dunnett's (B), Tukey's multiple comparisons test (C, F, and G), and unpaired Student's t test (D and E) (****p < 0.0001, ***p < 0.001, and **p < 0.01, *p < 0.05, ns: not significant.)

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 (Supplementary Fig. 3B). 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, we selected six adenosine candidates of m^{6}A modification (332, 495, 569, 600, 679, and 839) through MeRIP-seq analysis in either hESCs or L1 reporter-expressing HeLa cells (Supplementary Table 1). We generated a set of firefly luciferase reporter plasmids encoding L1 5′ UTR or its m^{6}A-silencing A to T mutants (Fig. 3A). To quantify the effect of L1 5′ UTR m^{6}A mutation without the bias from transfection efficiency, we 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 (Fig. 3B). However, the weak effect of these single mutants led us to hypothesize that multiple m^{6}A 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 (Fig. 3C).

We next performed the L1 retrotransposition assay using the 5′ UTR m^{6}A mutants of the pL1Hs construct. Mutations at each m^{6}A 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^{6}A mut) markedly inhibited L1 mobility (Fig. 3D and Supplementary Fig. 3C). We validated the effect of the m^{6}A cluster using the L1-luciferase reporter construct pYX014. Indeed, the triple m^{6}A mutant of the L1-luciferase construct (pYX014 L1 m^{6}A mut) induced approximately 50% decline in L1 mobility compared to that induced by the wild-type L1 (Supplementary Fig. 3D).

To assess the effect of the triple mutation in the m^{6}A modification level of L1, we 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* (Fig. 3E and Supplementary Fig. 3E). 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 our finding that ALKBH5 inhibits L1 mobility, we 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 (Fig. 3F). However, ALKBH5 overexpression caused only marginal effects in pL1 m^{6}A mut-expressing cells (Fig. 3F). Moreover, silencing the triple m^{6}A modification led to the suppression of L1 mobility in AcGFP-expressing cells, but not

153 in ALKBH5-expressing cells (Fig. 3F). In a reciprocal experiment, we measured the L1 retrotransposition
154 frequency of pL1Hs and pL1 m^{6}A mut in ALKBH5-lacking cells. Notably, ALKBH5 knockdown led to the
155 enhancement of L1 mobility in pL1Hs-expressing cells, whereas no measurable changes were observed in pL1
156 m^{6}A mut-expressing cells (Fig. 3G). Consistent with this result, ALKBH5 was not able to suppress L1 ORF1p
157 expression in the absence of the m^{6}A cluster (Supplementary Fig. 3G, H). In summary, we demonstrated that
158 ALKBH5 suppresses L1 expression in the 5′ UTR m^{6}A cluster-dependent manner, which suggests that the L1
159 5′ UTR m^{6}As serve as the substrates for ALKBH5 demethylation.

160 **m^{6}A modification promotes the translational efficiency of L1 RNA**

161 Given that m^{6}A regulates L1 ORF1p expression, we investigated the stages in the L1 life cycle that are
162 regulated by m^{6}A modification. First, we quantified L1 RNA expression in the presence or absence of the 5′
163 UTR m^{6}A cluster using two different plasmids, pL1Hs and pYX014. Irrespective of the vectors used, L1 m^{6}A
164 mutation did not influence the levels of L1 RNA expression (Supplementary Fig. 4A, B). We next assessed the
165 stability of reporter L1 mRNAs with or without the 5′ UTR m^{6}A mutation using the transcription inhibitor,
166 actinomycin D. L1 RNA was more stable in both pL1Hs- and pL1 m^{6}A mut-expressing HeLa cells in compared
167 to positive control, c*MYC* mRNA (Supplementary Fig. 4C). We did not observe any significant difference in L1
168 RNA stability by m^{6}A mutation (Supplementary Fig. 4C). We next examined the distribution of reporter L1
169 mRNAs in the nuclear and cytoplasmic fractions. In comparison to that of *GAPDH* (abundant in the cytoplasm)
170 and *MALAT1* (abundant in the nucleus), over 80% of the L1 mRNA was present in the cytoplasmic fraction and
171 the m^{6}A-deficient mutation did not affect the cellular localization of L1 RNA (Supplementary Fig. 4D).

172 Several recent studies have linked 5′ UTR m^{6}A modification to translational efficiency in the context of
173 cellular stress[30-32]. Besides, a previous study raised the possibility that the presence of the L1 5′ UTR determines
174 the quality of L1 RNA[33]. Therefore, we reasoned that the L1 5′ UTR m^{6}A cluster could modulate the translation
175 of L1 RNA. To test this hypothesis, we performed an immunoblot assay in HeLa cells that expressed a single
176 to triple m^{6}A mutant of the pL1 construct. The expression levels of ORF1p gradually decreased as the number
177 of mutations increased (Fig. 4A). In addition, through polysome profiling, we captured polysome-bound RNA
178 to assess the translational efficiency of L1 RNA. The deletion of the m^{6}A cluster significantly reduced the

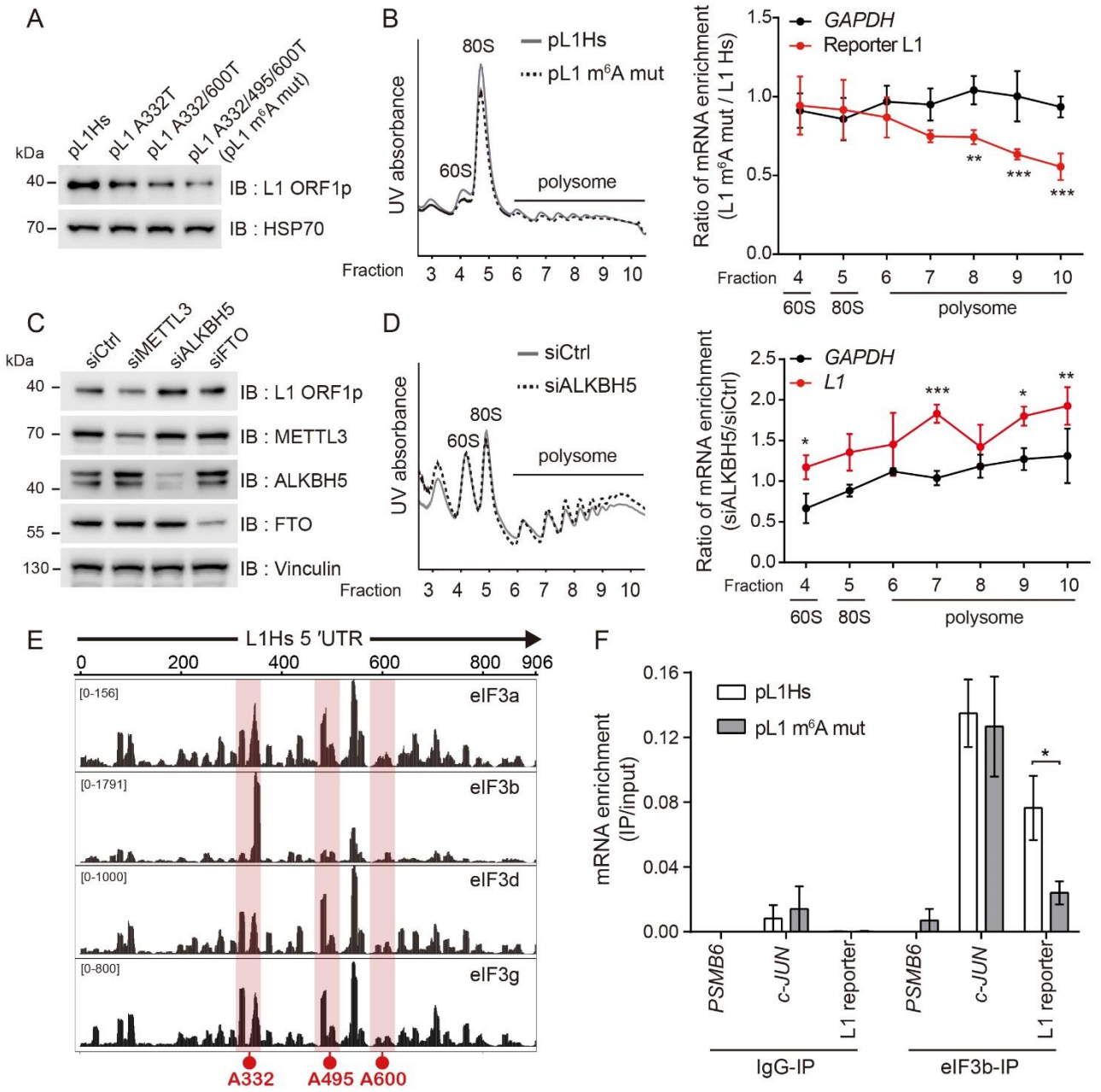

Figure 4. L1 5′ UTR m^{6}A cluster enhances the translational efficiency through the recruitment of eIF3. **(A)** Immunoblot analysis for assessing the effect of m6A mutation in L1 ORF1p levels. HSP70 served as a loading control. **(B)** Polysome profiling of pL1Hs- or pL1 m^{6}A mut-expressing HeLa 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, error = 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) **(C)** Immunoblot assay for determining endogenous L1 ORF1p levels in PA-1 cells treated with indicated siRNAs. Vinculin served as a loading control. **(D)** Polysome profiling of PA-1 cells lacking ALKBH5 compared to siCtrl (left panels). The levels of endogenous L1 RNA was measured as in (B) using L1 5′ UTR-specific primers (right panel). (n = 3 independent samples, error = s.d., statistical significance was determined as in (B)). **(E)** Identification of eIF3 binding sites in L1Hs 5′ UTR using the previously reported eIF3 PAR-CLIP data set (GSE65004)[1]. The red boxes indicate the m^{6}A sites-containing region. **(F)** 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, error = s.e.m., unpaired Student's t test; *p < 0.05).

enrichment of polysome-bound L1 RNA compared to that of pL1Hs constructs (Fig. 4B and Supplementary Fig. 5A). To validate these results, we investigated whether m^{6}A regulates the translational efficiency of endogenous L1 mRNAs in PA-1 human embryonic carcinoma cells. Consistent with the effects of m^{6}A machinery depletion in pL1Hs-expressing HeLa cells (Fig. 1E), ALKBH5 knockdown augmented the production of endogenous ORF1p while METTL3 knockdown reduced ORF1p synthesis (Fig. 4C). 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 (Supplementary Fig. 5B). 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 (Fig. 4D and Supplementary Fig. 5C), which indicates that ALKBH5 regulates L1 retrotransposition by suppressing the efficiency of L1 RNA translation.

eIF3 is an m^{6}A-binding protein and promotes the selective translation of mRNAs that bear m^{6}A in 5′ UTR[30]. These characteristics of eIF3 lead us to hypothesize that the L1 5′ UTR m^{6}A cluster serves as a docking site for eIF3 to promote translation. To define the functional relationship between eIF3 and the L1 m^{6}A cluster, we analyzed previously reported data from photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation sequencing (PAR-CLIP seq) of eIF3 subunits a, b, d, and g[34]. By mapping the reads from PAR-CLIP of the eIF3 subunits along the endogenous L1Hs, we revealed that eIF3 exhibits preferential binding to the L1 5′ UTR (Supplementary Fig. 6A). Furthermore, the PAR-CLIP clusters were significantly enriched in the A332 m^{6}A region in all four eIF3 subunits, while the A495 m^{6}A region contained PAR-CLIP clusters of three eIF3 subunits: eIF3a, d, and g (Fig. 4E). We were unable to detect the comparable eIF3-binding sites in the A600 m^{6}A region (Fig. 4E). To verify the interaction between eIF3 and the L1 m^{6}A cluster, we transfected pL1Hs or pL1 m^{6}A mut into HeLa cells and performed UV crosslinking immunoprecipitation using eIF3b antibody (Supplementary Fig. 6B). Through qRT-PCR analysis of the immunoprecipitated eluates, we observed the enrichment of L1 RNA comparable to *c-JUN,* a known eIF3-bound mRNA, in pL1Hs-expressing cells (Fig. 4F). *PSMB6* and eluates from IgG immunoprecipitation served as negative controls. Remarkably, the silencing of the m^{6}A cluster reduced the quantity of eIF3-bound L1 RNA by approximately 70%, which indicates that the L1 5′ UTR m^{6}A cluster bears the eIF3 docking site (Fig. 4F). Collectively, these data suggest that the 5′ UTR m^{6}A cluster recruits eIF3 for the efficient translation of L1 RNA.

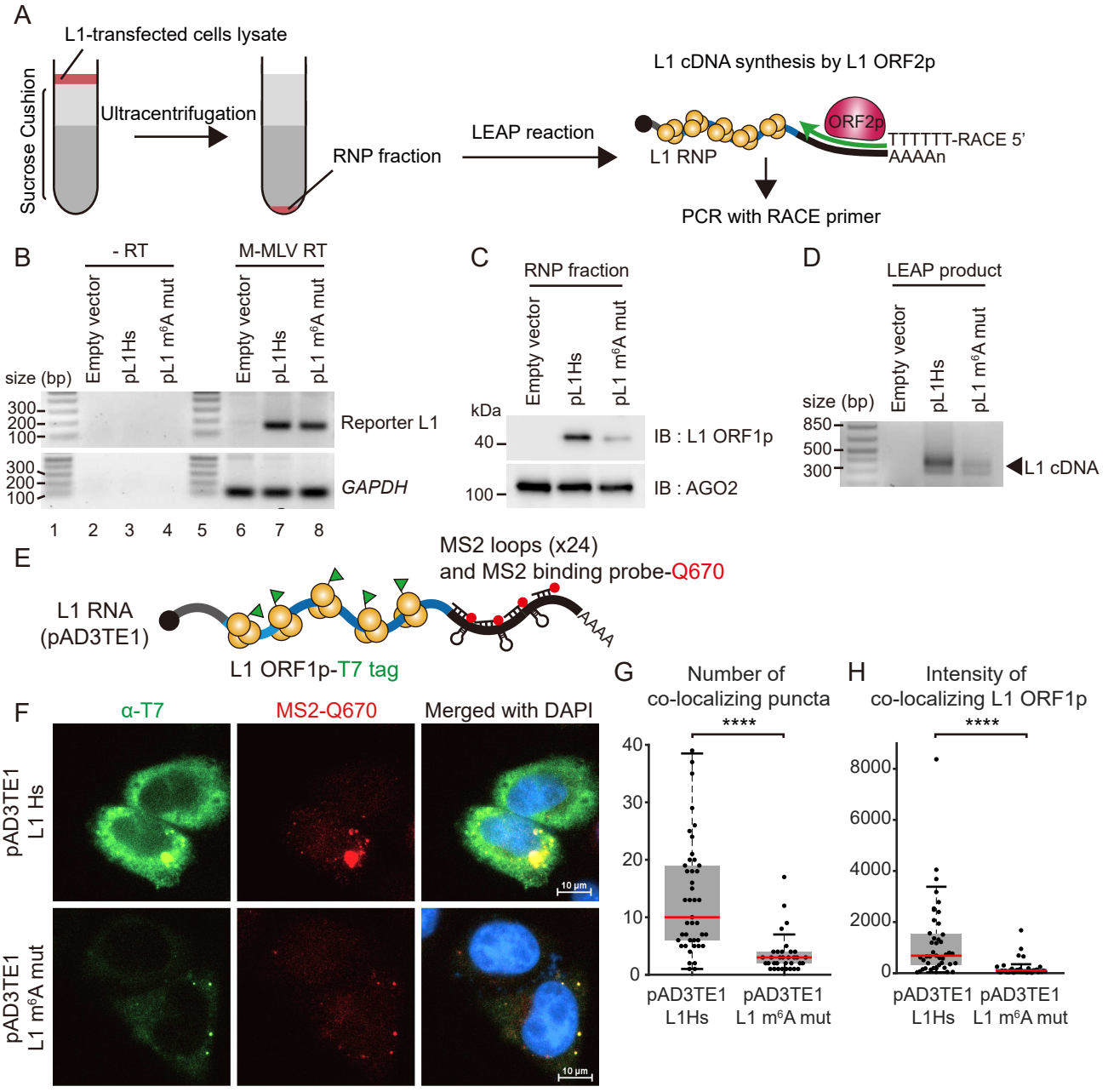

Figure 5. m^{6}A modification is crucial for generating retrotransposition-competent L1 RNPs. **(A)** Scheme of LEAP assay with description. **(B)** Quantification of mRNA levels in the RNP fraction of pL1-expressing HeLa cells. cDNA synthesis in the absence of reverse transcriptase (lane 2-4) and transfection 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. **(C)** Immunoblot assay of the RNP fraction from pL1-expressing HeLa cells. AGO2 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. **(E)** A schematic of the L1-MS2 construct (pAD3TE1) carrying T7-tagged ORF1p (green) and MS2 stem-loops with Q670-labeled MS2 binding probes (red). **(F)** 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. **(G)** 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. **(H)** Intensity of L1 ORF1p in co-localizing puncta. Each point represents the intensity of L1 ORF1p per cell. Data are presented with Tukey's boxes and whiskers. (43 cells for L1Hs and 34 cells for L1 m^{6}A mut; Kolmogorov-Smirnov test; ****$p < 0.0001$).

206 **5′ UTR m^{6}A cluster is necessary to produce a functional unit for L1 retrotransposition**

207 For successful L1 retrotransposition, both ORF1p and ORF2p are required to generate the L1 RNP with the
208 encoding L1 RNA[35]. Though we observed m^{6}A-mediated regulation in ORF1p synthesis, it is necessary to
209 determine whether m^{6}A modification at the 5′ UTR influences L1 RNP formation. To investigate L1 RNP
210 regulation by m^{6}A, we obtained the cellular RNP fraction as previously reported[36]. Briefly, we prepared lysates
211 from pL1-transfected cells and purified L1 RNPs using sucrose cushion ultracentrifugation (Fig. 5A). We
212 detected comparable levels of L1 RNA in the RNP fractions from pL1Hs- and pL1 m^{6}A mut-expressing cells
213 (Fig. 5B). cDNA synthesis reaction in absence of reverse transcriptase revealed that neither genomic DNA nor
214 plasmid contamination was present in the RNP fraction (Fig. 5B). Immunoblotting of the RNP fraction showed
215 that the levels of RNP-associated ORF1p were diminished by L1 5′ UTR m^{6}A mutation (Fig. 5C). This indicates
216 that the m^{6}A cluster mutation abolished the sufficient production of ORF1p for L1 RNP formation.

217 Since ORF2p expression level is too low to observe changes in the m^{6}A mutant[37,38], we introduced the L1
218 element amplification protocol (LEAP) to gauge the reverse transcriptase activity of ORF2p[36] (Fig. 5A).
219 Incubation of RNPs with LEAP primer facilitates ORF2p-mediated L1 cDNA synthesis. We amplified LEAP
220 products using PCR with specific primers for reporter L1 and RACE adapter, which yielded products of 300 to
221 400 base pairs (bp) (Fig. 5D). However, m^{6}A-abrogated L1 RNP produced cDNA at significantly lower levels
222 than the wild-type L1 RNP did (Fig. 5D). These results reveal that the m^{6}A cluster is necessary for L1 cDNA
223 production, which suggests that the m^{6}A cluster regulates ORF2p expression or its activity.

224 ORF1p oligomerization is critical for successful L1 retrotransposition[40]. We examined whether inefficient
225 ORF1p synthesis results in a failure of L1 RNP formation. For a quantitative assessment of individual L1 RNP
226 formation, we introduced the pAD3TE1 construct carrying T7-tagged ORF1p and MS2 stem-loop structures in
227 the L1 3′ UTR (Fig. 5E). We performed RNA fluorescence in situ hybridization (FISH) with fluorescent Q670-
228 labeled probes complementary to the linker regions between the MS2 loops and immunofluorescence
229 experiments with anti-T7 antibody (Fig. 5F). Through z-stack analysis, we obtained the coordinates for the
230 fluorescent signals of L1 RNA and ORF1p and identified the L1 RNPs by sorting out co-localizing particles
231 within an intermolecular distance of 330 nm between L1 RNA and ORF1p. Consistent with the previous study[41],
232 we observed co-localizing signals of L1 RNP as cytoplasmic aggregates (Fig. 5F). However, L1 m^{6}A mut-
233 expressing cells showed a significant reduction in both the number of L1 RNP foci and the signal intensity of

co-localizing ORF1p (Fig. 5F, G, H). These data indicate that the abrogation of the m^{6}A cluster reduces the levels of ORF1p in L1 RNP and causes a concomitant decrease in the number of L1 RNP particles.

m^{6}A is a driving force for L1 evolution

Over the last 40 million years of human evolution, L1 subfamilies have frequently acquired novel 5′ UTRs[28]. 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[42]. Considering that RNA methyltransferase installs m^{6}A in a sequence-specific manner, we speculated that nucleotide mutations might lead to the acquisition or loss of the m^{6}A consensus motif during L1 evolution. To unravel the evolutionary history of L1 5′ UTR m^{6}A cluster regions, we analyzed 443 human-specific full-length L1s[43] and compared the three m^{6}A motif sites, A332, A495, and A600. Given that adenosine residue should be followed by cytosine residue to form the m^{6}A consensus motif DRA^{m}CH, A332 m^{6}A positive L1s constitute a considerably small part in the L1PA3 lineage (12.4%). In L1PA2 and younger lineages, the number of A332 m^{6}A positive L1s increased drastically (92.9%) (Fig. 6A), and the same was observed in the youngest L1Hs (Supplementary Fig. 7A). On the contrary, A495 and A600 are tightly conserved in all human-specific L1 subfamilies (Supplementary Fig. 7B). We investigated this tendency of the A332 m^{6}A 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^{6}A positive L1s, while the chimpanzee- or gorilla-specific L1s continue to harbor the m^{6}A motifs of A495 and A600 (Fig. 6A, and Supplementary Fig. 7C, D, E, F). As in the L1Hs subfamily, the majority of the youngest chimpanzee-specific L1 subfamily (L1Pt) harbor the A332 m^{6}A motif (Supplementary Fig. 7C). In summary, we found that A332 m^{6}A motif acquisition by single nucleotide substitution (T333C) first appeared in L1PA3 or older lineages, which indicates that the productive potential of m^{6}A has allowed positive selection of A332 m^{6}A-positive L1s during the evolution from the common ancestor.

To evaluate the consequence of A332 m^{6}A acquisition in ancestral L1 5′ UTR, we generated a chimeric pL1 construct that contained L1PA2 5′ UTR and L1Hs ORF1/2 with the *mblastI* reporter (Fig. 6B). Based on the m^{6}A consensus motif DRA^{m}CH, T333 of pL1PA2$^{5'UTR}$ had no m^{6}A motif at A332, whereas T333C point

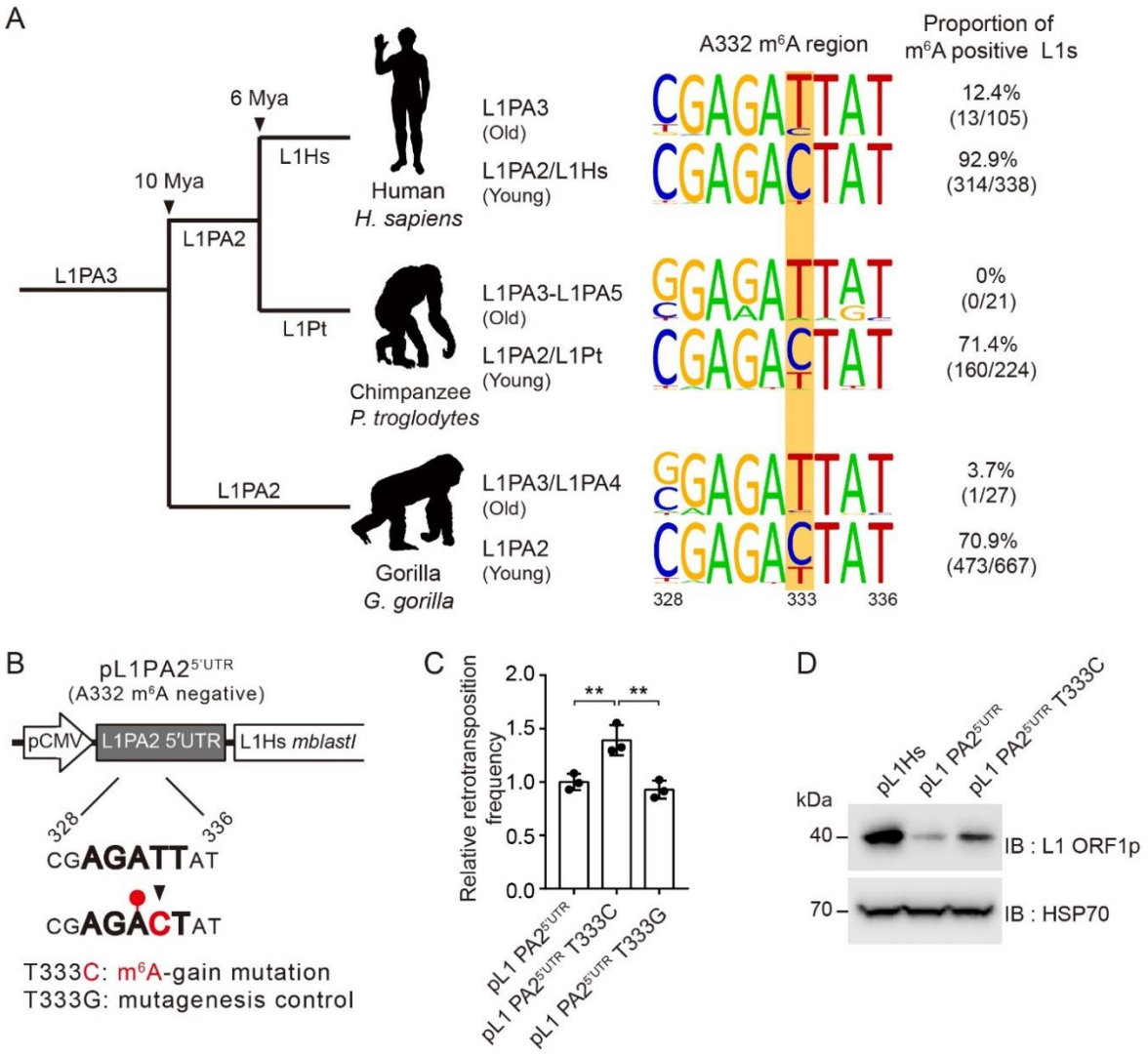

Figure 6. m^{6}A is a driving force in L1 evolution. **(A)** Comparative analysis of L1 A332 m^{6}A sites in species-specific full-length L1s from three primates. Phylogenetic tree of gorilla, chimpanzee, and human L1s with predicted age and the corresponding L1 subfamily lineages (left). Changes in the A332-m^{6}A motif region from L1PA3 or older L1s to L1PA2 and a younger L1 (right). The substitution site wherein the residue converts from T to C (333) is highlighted in yellow. The percentage indicates the proportion of m^{6}A motif-positive L1s with nucleotide C to total L1s. **(B)** 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. **(C)** 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. The values are normalized to that of pL1PA2$^{5'\,UTR}$. (n = 3 independent samples, error = s.d., one-way ANOVA and Tukey's multiple comparison test; **p < 0.01) **(D)** Immunoblot assay showing L1 ORF1p expression in the indicated pL1-transfected HeLa cells. HSP70 served as a loading control.

mutation enabled A332 m^{6}A modification (Fig. 6B). 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 (Fig. 6C). As expected, the T333C m^{6}A-gain mutation enhanced ORF1p synthesis of pL1PA2$^{5'\,UTR}$ (Fig. 6D). Although the acquisition of A332 m^{6}A motif only led to a 1.4-fold increase in the cultured cell-based L1 retrotransposition assays (Fig. 6C), the 12 million years of L1 evolution would have been sufficient to amplify

the profound effect of m^{6}A. These results suggest that m^{6}A modification in the L1 5' UTR region may have played a crucial role in the L1 evolution of primates.

DISCUSSION

The role of m^{6}A modification in pathogenic viral transcripts has been reported in the past decade[44]. However, the role of m^{6}A in L1s as genomic parasites have been poorly understood. Although recent studies have revealed the m^{6}A modification of L1 repetitive elements with respect to R-loop or chromatin regulation[24,25], the functions of m^{6}A in the life cycle of retrotransposition-competent L1s have not been established. In our study, we demonstrated that the proper formation of the m^{6}A cluster in 5' UTR of L1 RNA is essential for L1 retrotransposition. The evolutionary history of the m^{6}A cluster in primate-specific L1s revealed the most influential m^{6}A region (A332) obtained in the past 12 million years, which suggests the potential role of m^{6}A as a driving force in L1 evolution (Fig. 7).

Our results revealed that the m^{6}A 'writer' METTL3 and the m^{6}A 'eraser' ALKBH5 regulates L1 mobility in an opposite manner, and showed that m^{6}A modifications occurs predominantly in L1 5′ UTR. 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[33]. 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. Our findings demonstrated that L1 5′ UTR m^{6}A modification is essential for L1 translation, L1 RNP formation, and thus retrotransposition. Therefore, we provide a new perspective on the regulatory function of L1 5′ UTR as a hub for RNA modification.

We 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[36,41], 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[39], 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[45,46] or RNA secondary structure[47], which might affect L1 ORF2p enzymatic activity. Therefore, future studies could reveal the role of m^{6}A in ORF2p regulation. By adopting a microscopic

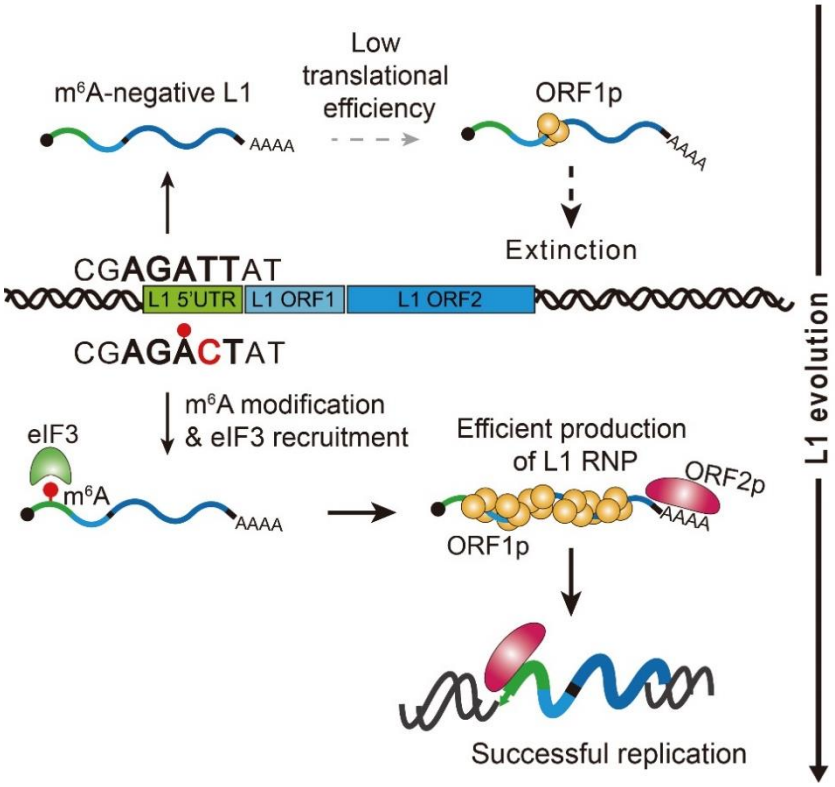

Figure 7. A proposed model for the role of m^{6}A in L1 replication and evolution

292 approach, we confirmed that m^{6}A is critical for the formation of L1 RNP aggregates. The rate of ORF1p
293 oligomerization is the limiting factor in the production of successful L1 RNPs[40]. Therefore, we speculated that
294 5′ UTR m^{6}As enable L1 RNA to produce sufficient ORF1p, which further accelerates the oligomerization of
295 ORF1p. Since the process of L1 RNP formation is more complicated than the biochemical interaction between
296 L1 RNA and its protein, the process by which m^{6}A orchestrates the assembly of retrotransposition-competent
297 L1 RNP remains to be understood.

298 eIF3 recognizes an m^{6}A residue in the 5′ UTR and promotes the translation of mRNAs[30]. We assumed that the
299 L1 5′ UTR recruits eIF3 to the m^{6}A cluster for efficiently translating the L1 mRNA. Indeed, eIF3 PAR-CLIP-
300 seq data reveal the interaction between eIF3 and L1 5′ UTR m^{6}A residue. We also demonstrated that the eIF3-
301 bound portion of L1 decreases in the absence of the 5′ UTR m^{6}A. A single m^{6}A residue is sufficient to induce
302 eIF3-mediated translation[30]. This could explain the synergetic effects of triple m^{6}A residues in L1 5′ UTR, which
303 suggests that each m^{6}A residue can serve as a docking site for eIF3. Moreover, under cellular stress, the 5′ UTR
304 m^{6}A facilitates the cap-independent translation of mRNA[30-32]. These studies raise the possibility that m^{6}A
305 initiates the cap-independent translation of L1 RNA. Although Dmitriev *et al*. revealed that human L1 mRNA

306 is translated in a cap-dependent manner, m^{6}A modification was not considered in their experiments[48]. Therefore,
307 in future studies, it is important to determine whether m^{6}A modification enables the cap-independent translation
308 of L1 and whether m^{6}A acts as a molecular switch for L1 expression under cellular stress.
309 L1s have been continuously active since the origin of mammals[49]. One of the previous studies on L1 evolution
310 revealed that several distinct L1 lineages coexisted and were in a simultaneously activated state in the ancestral
311 primate genome. However, since the emergence of the L1PA lineage, the L1 subfamily has evolved and
312 maintained itself as a single lineage in the last 25 million years of the evolution of human and its close relatives[28].
313 The study proposed that the competition between or coexistence of L1 lineages is determined by the status of
314 the 5′ UTR of L1s and acquisition of novel 5′ UTR is a fundamental feature in mammalian L1 evolution[28]. Given
315 that m^{6}A methyltransferase marks m^{6}A in a sequence-specific manner, the accumulation of mutations in L1
316 might cause the loss or acquisition of putative m^{6}A motifs. To further elucidate the history of m^{6}A in L1
317 evolution, we analyzed species-specific full-length L1s from the human, chimpanzee, and gorilla genome.
318 Notably, the A332 m^{6}A motif first appeared in L1PA3 or older L1 lineages more than 12 million years ago.
319 During the evolution of the three different primates, humans, chimpanzees, and gorillas, the A332 m^{6}A-positive
320 L1s have propagated their progenies and have become the dominant L1 subfamilies. As m^{6}A modification
321 promotes L1 mobility, the acquisition of m^{6}A would have resulted in the positive selection of A332 m^{6}A-
322 containing L1s. Over the extended periods of L1 evolution, L1s have competed for survival against host
323 restriction[13,28,50]. KRAB-zinc finger proteins (ZFP), which have evolved with L1s, suppress the old L1
324 transcription in a sequence-specific manner[13,50]. However, L1Hs, which is the youngest L1 lineage in the human
325 genome, escapes KRAB-ZFP restriction and is not recognized by any KRAB-ZFPs[50]. Instead, the host defense
326 utilizes post-transcriptional suppression mechanisms, such as small RNA interference (e.g., piRNA) or
327 APOBECs, to restrict the replication of L1s[51]; however, the youngest L1s are still active. Our findings provide
328 clues on how the youngest L1s continuously replicate under host surveillance. The emergence and the
329 propagation of the A332 m^{6}A-positive L1s suggest that 5′ UTR m^{6}A modification was a countermeasure against
330 the host post-transcriptional restriction. Therefore, we propose a novel survival strategy in the recent history of
331 L1 evolution.

Methods

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.

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[22]. 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$^{5′\ UTR}$, we synthesized the L1PA2 5′ UTR region based on reported consensus sequences using gene synthesis (Cosmogenetech). We then replaced the L1Hs 5′ UTR of pL1Hs with L1PA2 5′ UTR, as described above.

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[41]. pAD3TE1 was gift from Aurélien J. Doucet. The generation of L1 5′ UTR m^6A mutant constructs of pAD3TE1 was performed according to the method for pL1Hs mutant construct generation.

L1-firefly luciferase-tagged plasmids pYX014 and pYX015 were gifts from W. An[23]. 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^{6}A 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.

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.

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 DharmaFECCT 1 transfection reagent (Dharmacon) according to the manufacturer's instructions.

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 with the respective primary antibodies at 4°C. Subsequently, the membranes were washed thrice with TBST and incubated with HRP-conjugated secondary antibodies 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).

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 Supplementary Table 2.

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^{6}A 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 above.

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^{6}A 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 above.

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 $\times 10^6$ 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.

L1 *mblastI* retrotransposition assay. Next, 8×10^4 HeLa cells were plated 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.

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.

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.

Crosslinking immunoprecipitation and qPCR (CLIP-qPCR). eIF3-RNA CLIP-qPCR was performed as described previously[30] 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 1400 rpm 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 1400 rpm shaking incubator. RNA was extracted using TRIzol LS supplemented with 20 ng of spike-in RNA.

Methyl-RNA immunoprecipitation (MeRIP)-seq. MeRIP was performed as described earlier[52] 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[53]. 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 Supplementary Table 3.

For MeRIP-seq analysis, the adapters were trimmed using Cutadapt[54] (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)[55] and peak calling was performed using MACS2[56]. For analyzing the m^{6}A modifications in endogenous L1, the sequence reads from human embryonic stem cells were retrieved[27] (accession code: GSE52600) and were aligned against L1Hs consensus sequence using STAR. The codes are available from https://github.com/hastj7373/merip-seq.

RNA FISH and Immunofluorescence. The L1 MS2-stem-loop constructs pAD3TE1 L1Hs and pAD3TE1 L1 m^{6}A 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[57]; 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 in anti-T7 primary antibody (Abcam) diluted in the blocking solution 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 (Jackson immunoresearch) diluted in blocking solution 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.

Co-localization analysis of RNA FISH and IFA microscope image. Binary masks of cells were generated using the ROI manager in ImageJ[58]. Protein and mRNA particles from z-stack images were detected using the TrackNTrace software[59]. 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.

LEAP assay. The LEAP assay was performed as described previously[36]. 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.

eIF3 PAR-CLIP analysis. We utilized previously published PAR-CLIP data for eIF3a, b, d, and g[34]. 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, we 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;[54]). The remaining reads were mapped to the L1Hs consensus sequence, wherein upto three mismatches were allowed using bowtie2[60] (--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.

Comparison analysis of species-specific m^{6}A site. To identify the species-specific full-length L1s in human, chimpanzee, and gorilla genome, we used BLAT-based and liftOver-based methods with a computational approach. The detailed protocol was previously described[43]. Here, we 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). We collected and retrieved the species-specific full-length L1s. We then classified the L1 subfamilies (L1Hs, L1PA2~L1PA5) using RepeatMasker utility[61]. 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[62]. The conserved sequence motifs at the three sites (A332, A495, and A600) were visualized using the program Weblogo[63]. Species-specific L1 loci are listed in Supplementary Table 4-6.

AUTHOR CONTRIBUTIONS

S.-Y.H., S.L., J.C., and K.A. contributed to the conceptualization and designed the experiments. S.-Y.H. performed the biochemical and cell biological experiments. H.J. and J.K.C. conducted and provided support for the MeRIP-seq analyses. S.M., W.T., P.L., and K.H. worked on the identification of gorilla-specific full-length L1s and performed comparative analyses among primate-specific L1 subfamilies. S.C.B. performed eIF3 PAR-CLIP seq analyses. S.L. performed RNA FISH and IFA. S.L., H.C.M., and H.Y.P. led the microscopic image analyses. B.K. and V.N.K. provided support during the generation of the MeRIP-seq cDNA libraries. Y.C. and V.N.K. contributed to the polysome profiling experiments. Y.-H.G. and H.-J.C. provided the cell lysates of hESCs. S.-Y.H., H.J., S.M., S.L., S.C.B., H.C.M., H.Y.P., K.H., and K.A. wrote the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Institute for Basic Science of the Ministry of Science Grant IBS-R008-D1 and Samsung Science and Technology Foundation SSTF-BA1402-19 (to K. Ahn), funding from the Fellowship for Fundamental Academic Fields of Seoul National University (to S.-Y. Hwang), and the BK21 plus fellowship. We are grateful to J. L. Garica-Perez, W. An, A. Roy-Engel, and A.J. Doucet for providing the L1 plasmids. Computational comparative genomic work was facilitated by the Compute Canada high-performance computing facilities.

CONFLICTS OF INTEREST

None declared.

References

1 Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921, doi:10.1038/35057062 (2001).

2 Beck, C. R. *et al.* LINE-1 retrotransposition activity in human genomes. *Cell* **141**, 1159-1170, doi:10.1016/j.cell.2010.05.021 (2010).

3 Brouha, B. *et al.* Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* **100**, 5280-5285, doi:10.1073/pnas.0831042100 (2003).

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

5 Martin, S. L. & Bushman, F. D. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol Cell Biol* **21**, 467-475, doi:10.1128/MCB.21.2.467-475.2001 (2001).

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

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

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

9 Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. 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, doi:10.1016/0092-8674(93)90078-5 (1993).

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

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

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

13 Jacobs, F. M. *et al.* An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* **516**, 242-245, doi:10.1038/nature13760 (2014).

14 Castro-Diaz, N. *et al.* Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes & development* **28**, 1397-1409, doi:10.1101/gad.241661.114 (2014).

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

16 Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635-1646, doi:10.1016/j.cell.2012.05.003 (2012).

17 Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206, doi:10.1038/nature11112 (2012).

18 Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* **49**, 18-29, doi:10.1016/j.molcel.2012.10.015 (2013).

19 Jia, G. *et al.* N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol* **7**, 885-887, doi:10.1038/nchembio.687 (2011).

20 Lichinchi, G. *et al.* Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells.

Nat Microbiol **1**, 16011, doi:10.1038/nmicrobiol.2016.11 (2016).

21 Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917-927, doi:10.1016/s0092-8674(00)81998-4 (1996).

22 Klawitter, S. *et al.* Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. *Nat Commun* **7**, 10286, doi:10.1038/ncomms10286 (2016).

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

24 Abakir, A. *et al.* N6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nature Genetics* **52**, 48-55, doi:10.1038/s41588-019-0549-x (2020).

25 Liu, J. *et al.* N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580-586, doi:10.1126/science.aay6018 (2020).

26 Macia, A. *et al.* Epigenetic Control of Retrotransposon Expression in Human Embryonic Stem Cells. *Molecular and Cellular Biology* **31**, 300-316, doi:10.1128/mcb.00561-10 (2011).

27 Batista, P. J. *et al.* m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707-719, doi:10.1016/j.stem.2014.09.019 (2014).

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

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

30 Meyer, K. D. *et al.* 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**, 999-1010, doi:10.1016/j.cell.2015.10.012 (2015).

31 Zhou, J. *et al.* Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* **526**, 591-594, doi:10.1038/nature15377 (2015).

32 Coots, R. A. *et al.* m(6)A Facilitates eIF4F-Independent mRNA Translation. *Molecular cell* **68**, 504-514.e507, doi:10.1016/j.molcel.2017.10.002 (2017).

33 An, W. *et al.* Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. *Mobile DNA* **2**, 2, doi:10.1186/1759-8753-2-2 (2011).

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

35 Wei, W. *et al.* Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* **21**, 1429-1439, doi:10.1128/MCB.21.4.1429-1439.2001 (2001).

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

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

38 Ergun, S. *et al.* 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, doi:10.1074/jbc.M312985200 (2004).

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

40 Naufer, M. N. *et al.* L1 retrotransposition requires rapid ORF1p oligomerization, a novel coiled coil-dependent

property conserved despite extensive remodeling. *Nucleic Acids Research* **44**, 281-293, doi:10.1093/nar/gkv1342 (2015).

41 Doucet, A. J. *et al.* Characterization of LINE-1 Ribonucleoprotein Particles. *PLOS Genetics* **6**, e1001150, doi:10.1371/journal.pgen.1001150 (2010).

42 Furano, A. V. a. B., S. in *In: Encyclopedia of LifeSciences (ELS)* (2008).

43 Tang, W., Mun, S., Joshi, A., Han, K. & Liang, P. 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, doi:10.1093/dnares/dsy022 (2018).

44 Dang, W. *et al.* N6-Methyladenosine and Viral Infection. *Frontiers in Microbiology* **10**, doi:10.3389/fmicb.2019.00417 (2019).

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

46 Perez-Perri, J. I. *et al.* Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nature Communications* **9**, 4408, doi:10.1038/s41467-018-06557-8 (2018).

47 Liu, N. *et al.* N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560-564, doi:10.1038/nature14234 (2015).

48 Dmitriev, S. E. *et al.* 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, doi:10.1128/mcb.02138-06 (2007).

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

50 Castro-Diaz, N. *et al.* Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev* **28**, 1397-1409, doi:10.1101/gad.241661.114 (2014).

51 Marchetto, M. C. N. *et al.* Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* **503**, 525-529, doi:10.1038/nature12686 (2013).

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

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

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

55 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2012).

56 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

57 Lionnet, T. *et al.* A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nat Methods* **8**, 165-170, doi:10.1038/nmeth.1551 (2011).

58 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675, doi:10.1038/nmeth.2089 (2012).

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

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

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

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

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

Figures

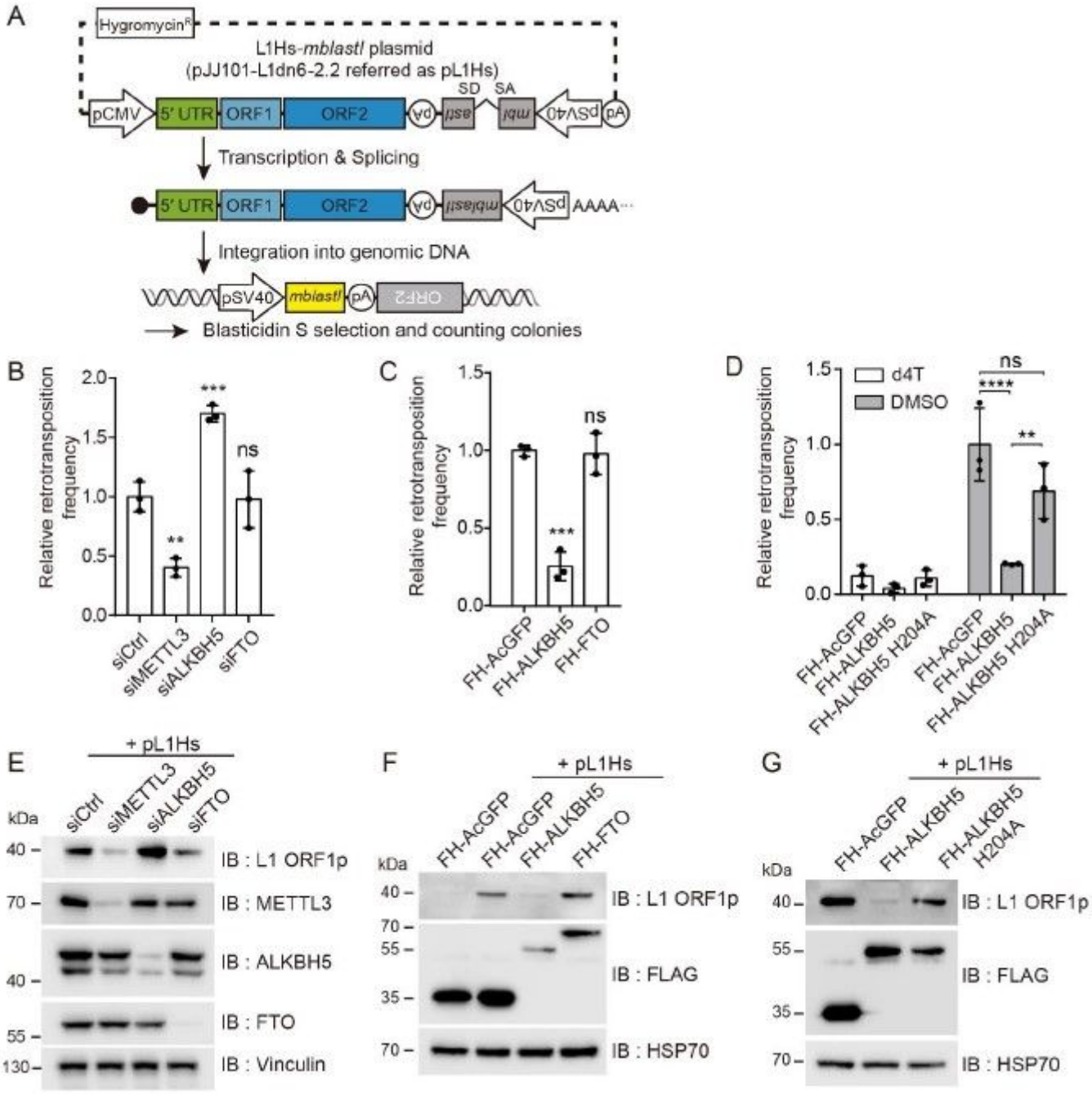

Figure 1

RNA methylation machinery controls L1 retrotransposition. (A) A schematic of the L1 construct and an overview of the L1 retrotransposition assay using engineered human L1 construct. (B) Retrotransposition assay in HeLa cells treated with siRNA that targets METTL3, ALKBH5, or FTO. A non-targeting siRNA (siCtrl) was used as a control. The results of independent experiments were normalized relative to those of the control. (C) Retrotransposition assays performed by cotransfecting the pL1Hs expression cassette with the indicated m6A enzyme-expressing vectors into HeLa cells. Retrotransposition frequency was normalized to that of the control (FH-AcGFP). (D) 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, error = 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). (E) Immunoblot assay of lysates from pL1Hs-transfected HeLa cells treated with indicated siRNAs that target m6A enzymes. Vinculin served as a loading control. (F and G) 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.

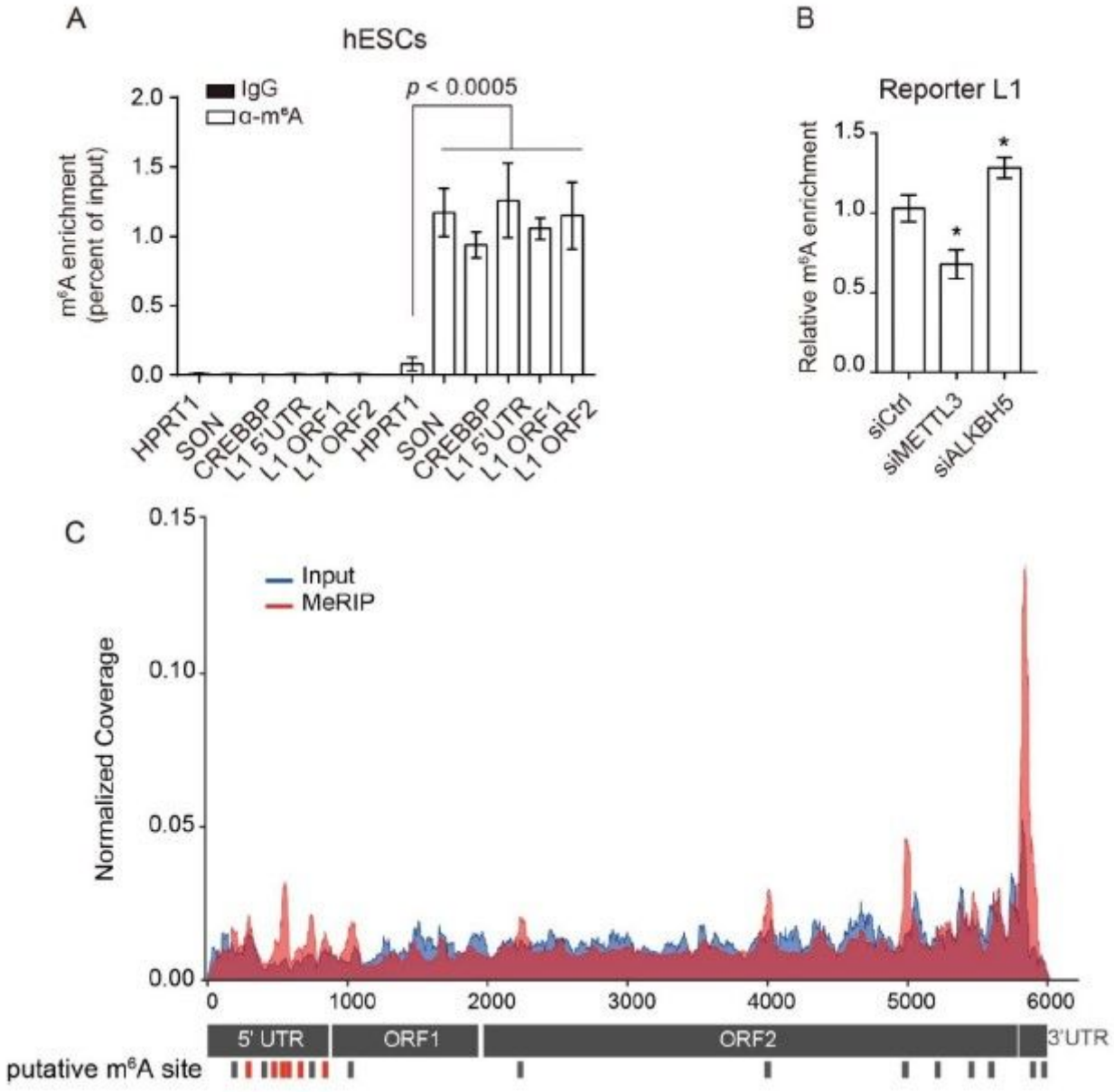

Figure 2

L1 RNA is modified by m6A. (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, error = s.e.m., one-way ANOVA and Dunnett's multiple comparisons test). (B) MeRIP-qPC analysis of pL1Hs-transfected HeLa cells with m6A 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, error = s.e.m., unpaired Student's t test, *p < 0.05). (C) Map of m6A 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 m6A peaks identified by manual inspection in two replicates. m6A peaks in red correspond to peaks containing high score m6A-prediction sites.

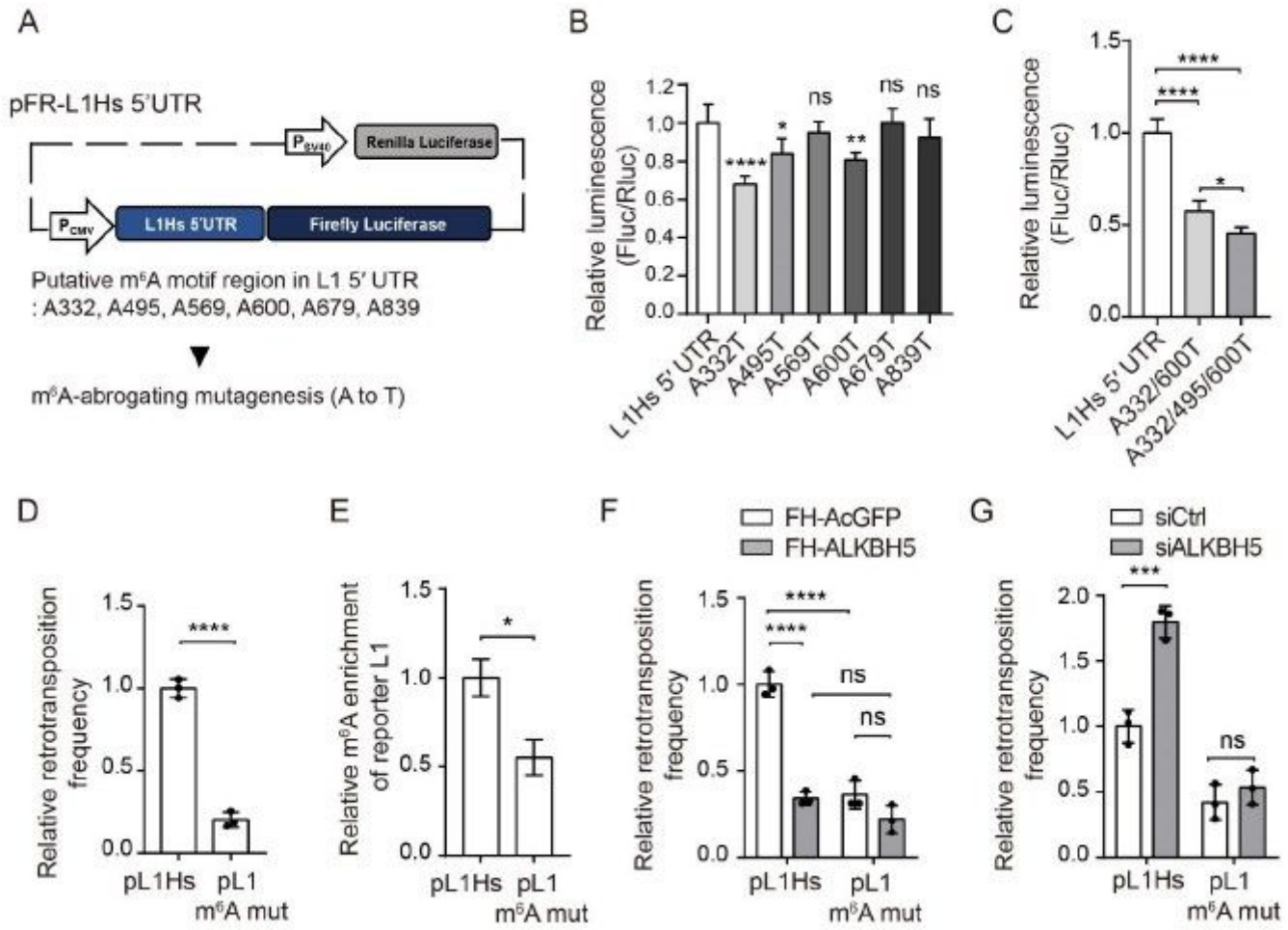

Figure 3

L1 5' UTR m6A cluster promotes L1 activity. (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 Tm6A-abrogating mutant. The ratio of the luminescence of firefly and Renilla luciferase (Fluc/Rluc) was normalized to pFRL1Hs 5' UTR-expressing cells. (error = s.d., four (B) or five (C) independent samples) (D) L1 assays using the triple m6A mutated L1 construct (pL1 m6A mut) in

HeLa cells. Retrotransposition frequency was normalized to that of pL1Hs. (n = 3 independent samples, error = s.d.) (E) MeRIP-qPCR analysis for evaluating the effect of the triple m6A mutation construct (pL1 m6A mut). m6A antibody-bound L1 RNA was normalized to that of pL1Hs-transfected cells. (n = 3 independent samples, error = s.e.m.) (F and G) Retrotransposition assay using pL1Hs-or pL1 m6A mut-expressing HeLa cells ALKBH5 overexpression (F) or silencing (G). L1 mobility was normalized to that in pL1Hs-expressing cells transfected with AcGFP (F) or non-targeting siRNA (G). (n = 3 independent samples, error = s.d.) In (B-G), Statistical significance was calculated by one-way ANOVA with Dunnett's (B), Tukey's multiple comparisons test (C, F, and G), and unpaired Student's t test (D and E) (****p < 0.0001, ***p < 0.001, and **p < 0.01, *p < 0.05, ns: not significant.)

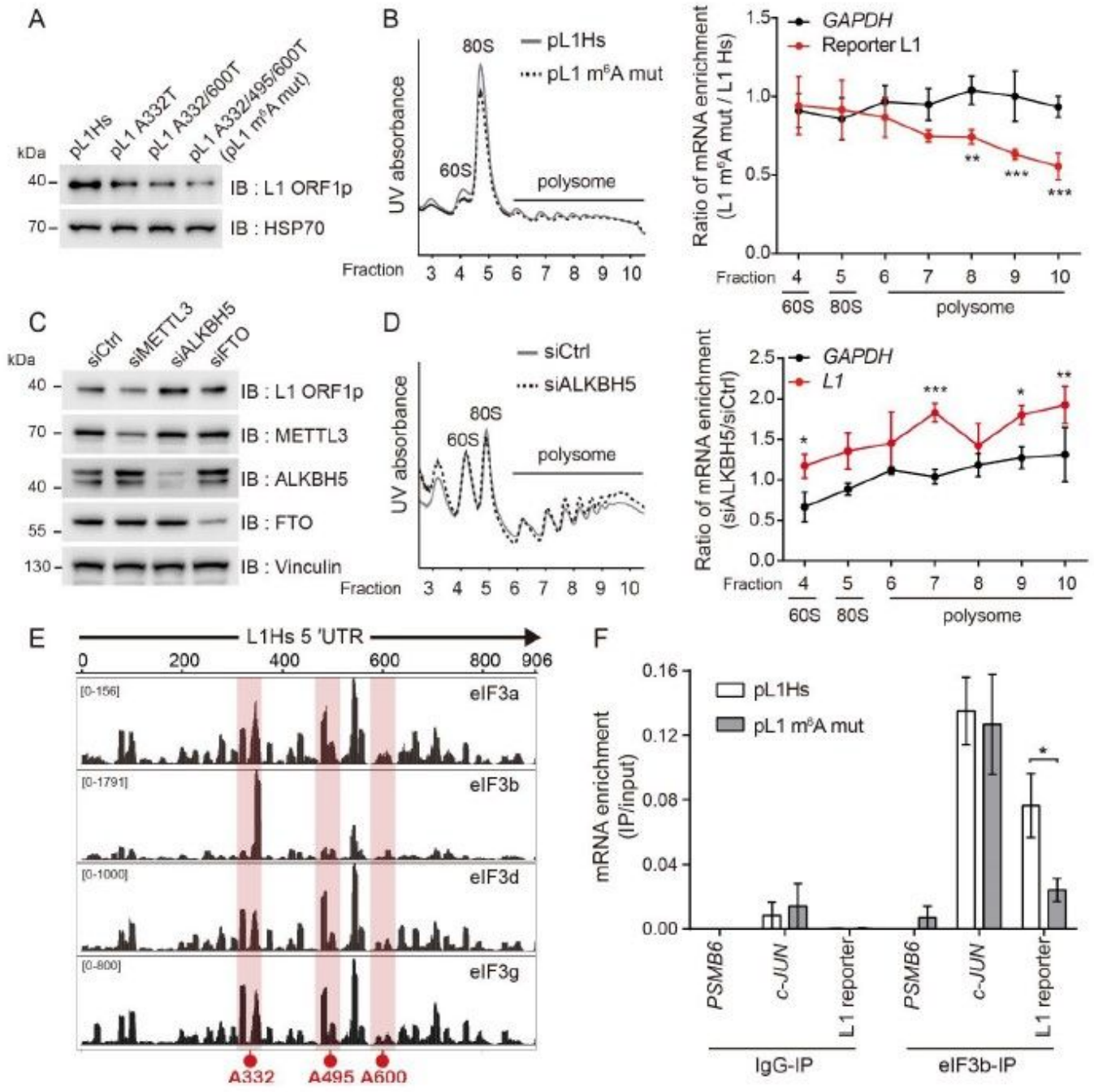

Figure 4

L1 5' UTR m6A cluster enhances the translational efficiency through the recruitment of eIF3. (A) Immunoblot analysis for assessing the effect of m6A mutation in L1 ORF1p levels. HSP70 served as a loading control. (B) Polysome profiling of pL1Hs- or pL1 m6A mut-expressing HeLa cells (left panel). Ratio of the polysome-bound mRNA levels in pL1 m6A 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, error = 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) (C) Immunoblot assay for determining endogenous L1 ORF1p levels in PA-1 cells treated with indicated siRNAs. Vinculin served as a loading control. (D) Polysome profiling of PA-1 cells lacking ALKBH5 compared to siCtrl (left panels). The levels of endogenous L1 RNA was measured as in (B) using L1 5' UTRspecific primers (right panel). (n = 3 independent samples, error = s.d., statistical significance was determined as in (B)). (E) Identification of eIF3 binding sites in L1Hs 5' UTR using the previously reported eIF3 PAR-CLIP data set (GSE65004)1. The red boxes indicate the m6A sites-containing region. (F) eIF3 UV-CLIP-qPCR using pL1Hs- or pL1 m6A mutexpressing HeLa cells. IgG-IP and PSMB6 served as negative controls. (n = 4 independent samples, error = s.e.m., unpaired Student's t test; *p < 0.05).

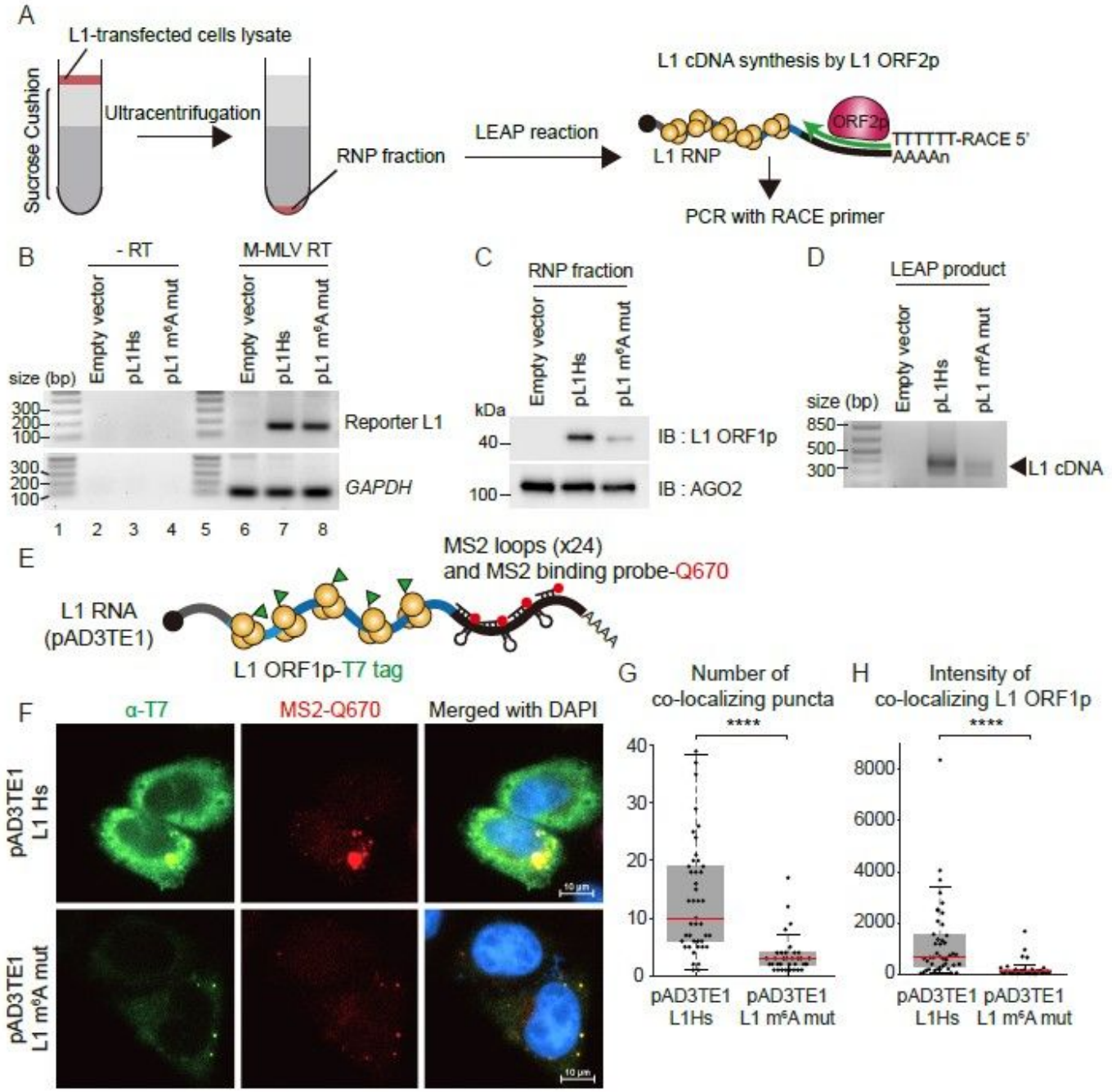

Figure 5

m6A modification is crucial for generating retrotransposition-competent L1 RNPs. (A) Scheme of LEAP assay with description. (B) 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. (C) Immunoblot assay of the RNP fraction from pL1-expressing HeLa cells. AGO2 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. (E) A schematic of the L1-MS2 construct (pAD3TE1) carrying T7-tagged ORF1p (green) and MS2 stem-loops with Q670-

labeled MS2 binding probes (red). (F) Immunofluorescence and RNA FISH images depicting HeLa cells transfected with pAD3TE1 L1Hs (top) or L1 m6A mut (bottom). Images for T7-tagged ORF1p (green), L1-MS2 RNA (red), and the merged images with DAPI (blue) are indicated. (G) 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. (H) Intensity of L1 ORF1p in co-localizing puncta. Each point represents the intensity of L1 ORF1p per cell. Data are presented with Tukey's boxes and whiskers. (43 cells for L1Hs and 34 cells for L1 m6A mut; Kolmogorov-Smirnov test; ****p < 0.0001).

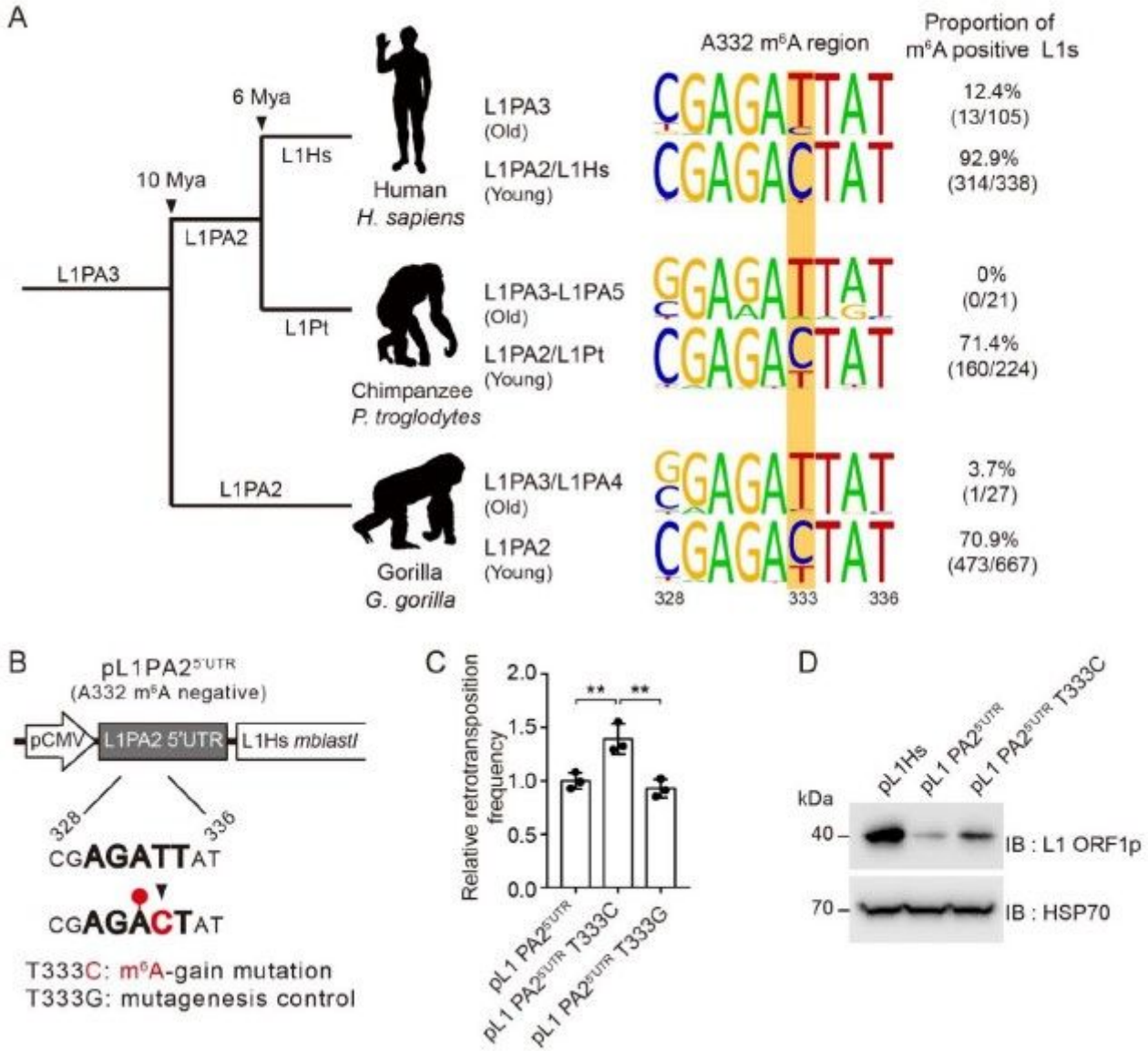

Figure 6

m6A is a driving force in L1 evolution. (A) Comparative analysis of L1 A332 m6A sites in species-specific full-length L1s from three primates. Phylogenetic tree of gorilla, chimpanzee, and human L1s with predicted age and the corresponding L1 subfamily lineages (left). Changes in the A332-m6A motif region from L1PA3 or older L1s to L1PA2 and a younger L1 (right). The substitution site wherein the residue

converts from T to C (333) is highlighted in yellow. The percentage indicates the proportion of m6A motif-positive L1s with nucleotide C to total L1s. (B) A schematic of retrotransposition assay using pL1PA25' UTR construct that is generated by substituting 5' UTR of pL1Hs with A332 m6A negative 5' UTR of L1PA2. (C) Retrotransposition assays for assessing the effect of A332 m6A acquisition in pL1PA25' UTR with T333C mutation. T333G mutation served as negative control. The values are normalized to that of pL1PA25' UTR. (n =3 independent samples, error = s.d., one-way ANOVA and Tukey's multiple comparison test; **p < 0.01) (D) Immunoblot assay showing L1 ORF1p expression in the indicated pL1-transfected HeLa cells. HSP70 served as a loading control.

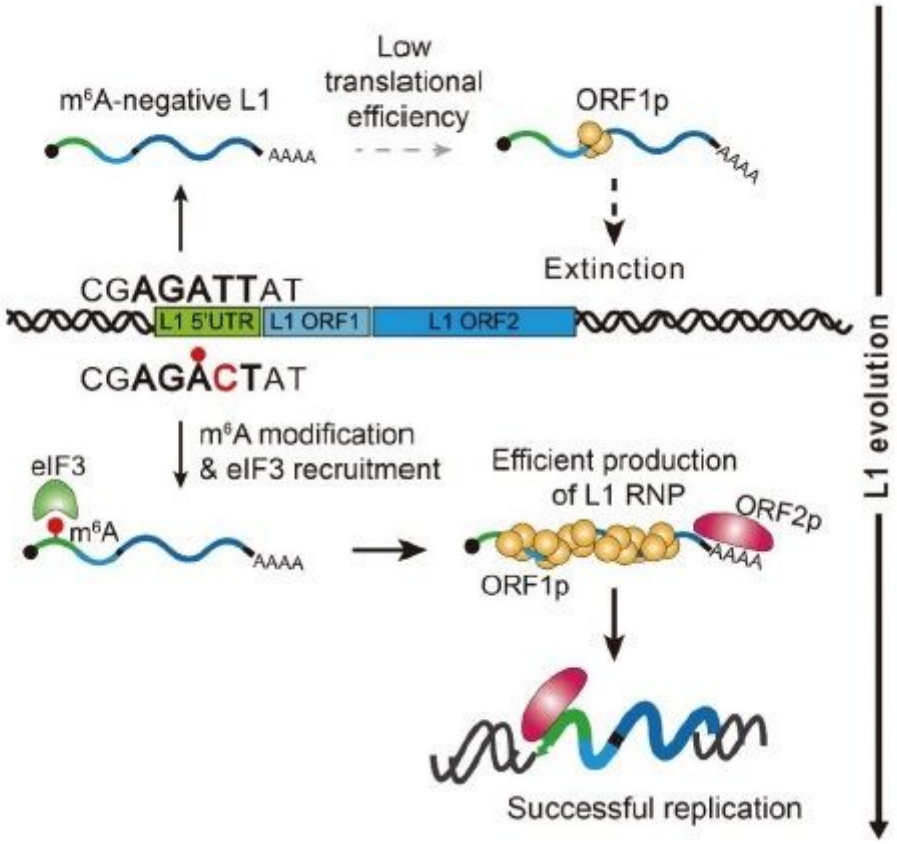

Figure 7

A proposed model for the role of m6A in L1 replication and evolution.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- KSAHNNatureCommunicationsSupplementalInformation.docx