




ARTICLE

# Endogenous reverse transcriptase and RNase H-mediated antiviral mechanism in embryonic stem cells

Junyu Wu<sup>1</sup>, Chunyan Wu<sup>1</sup>, Fan Xing<sup>1</sup>, Liu Cao<sup>1</sup>, Weijie Zeng<sup>1</sup>, Liping Guo<sup>1</sup>, Ping Li<sup>1</sup>, Yongheng Zhong<sup>1</sup>, Hualian Jiang<sup>1</sup>, Manhui Luo<sup>1</sup>, Guang Shi<sup>2</sup>, Lang Bu<sup>1</sup>, Yanxi Ji<sup>1</sup>, Panpan Hou<sup>1</sup>, Hong Peng<sup>1</sup>, Junjiu Huang<sup>2</sup>, Chunmei Li<sup>1</sup> and Deyin Guo<sup>1</sup> 

Nucleic acid-based systems play important roles in antiviral defense, including CRISPR/Cas that adopts RNA-guided DNA cleavage to prevent DNA phage infection and RNA interference (RNAi) that employs RNA-guided RNA cleavage to defend against RNA virus infection. Here, we report a novel type of nucleic acid-based antiviral system that exists in mouse embryonic stem cells (mESCs), which suppresses RNA virus infection by DNA-mediated RNA cleavage. We found that the viral RNA of encephalomyocarditis virus can be reverse transcribed into complementary DNA (vcDNA) by the reverse transcriptase (RTase) encoded by endogenous retrovirus-like elements in mESCs. The vcDNA is negative-sense single-stranded and forms DNA/RNA hybrid with viral RNA. The viral RNA in the heteroduplex is subsequently destroyed by cellular RNase H1, leading to robust suppression of viral growth. Furthermore, either inhibition of the RTase activity or depletion of endogenous RNase H1 results in the promotion of virus proliferation. Altogether, our results provide intriguing insights into the antiviral mechanism of mESCs and the antiviral function of endogenized retroviruses and cellular RNase H. Such a natural nucleic acid-based antiviral mechanism in mESCs is referred to as ERASE (endogenous RTase/RNase H-mediated antiviral system), which is an addition to the previously known nucleic acid-based antiviral mechanisms including CRISPR/Cas in bacteria and RNAi in plants and invertebrates.

*Cell Research* (2021) 31:998–1010; <https://doi.org/10.1038/s41422-021-00524-7>

## INTRODUCTION

Diverse and sophisticated cellular antiviral mechanisms have evolved across kingdoms. For examples, the system with clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) genes employs RNA-guided DNA cleavage to prevent DNA phage infection in bacteria and archaeobacteria, and the RNA interference (RNAi) adopts RNA-guided RNA cleavage to defend against RNA virus infection mainly in plants and invertebrates.<sup>1,2</sup> In mammalian cells, the interferon (IFN)-based innate immunity provides the first line of protection when facing virus infection.<sup>3,4</sup> However, the early embryos and the pluripotent stem cells, including embryonic stem cells (ESCs), have been reported to be defective in IFN production during viral infection and they also showed attenuated response to IFN treatment.<sup>5–9</sup> Nevertheless, stem cells still turn out to be resistant to viral infection.<sup>10–13</sup> Thus, how stem cells respond to virus infection is an interesting and challenging question. Recently, RNAi pathway and intrinsic expression of a subset of interferon-stimulated genes (ISGs) were shown to be involved in protection of ESCs from virus infection.<sup>13,14</sup> However, by knocking out Dicer, which is responsible for siRNA production, the RNAi-deficient stem cells become even more resistant to virus infection,<sup>15</sup> suggesting that there may exist other intricate antiviral mechanisms in ESCs.

In the genome of mammals, there exist a large number of retroviral sequences, known as endogenous retroviruses (ERVs), resulting from retroviral infection, viral genome reverse transcription and integration events.<sup>16</sup> As retro-transposition of ERVs

largely threatens the integrity of host genome, the vast majority of ERVs are mutated and defective in expression of viral proteins during evolution.<sup>16,17</sup> These ERV residuals are proposed to function as transcriptional regulatory elements or to generate long noncoding RNAs.<sup>18–20</sup> However, there are still some ERVs that remain intact and encode active viral proteins including reverse transcriptase (RTase).<sup>21–23</sup> The transcription of ERV elements is strictly repressed by DNA methylation in differentiated somatic cells, while they are regulated by the reversible histone modifications in ESCs and early embryos, where some ERVs show higher level of activation.<sup>24–27</sup> However, the functions of the ERV-coded proteins in mammalian early embryos or ESCs remain enigmatic and are of great interest to biologists.

In this study, we explored the function and mechanism of endogenous RTase in antiviral innate immunity of mESCs. We first demonstrated that inhibiting RTase activity could promote virus proliferation in mESCs and the antiviral function of RTase is independent of RNAi and intrinsic expression of ISGs. After virus infection, the RTase catalyzes the production of viral complementary DNA (vcDNA) that is complementary to the viral genome, resulting in the formation of DNA/RNA hybrid. RNase H1 is then recruited by the heteroduplex and digests the viral RNA, thus leading to inhibition of viral proliferation in mESCs. Taken together, our results provided compelling evidence showing that mESCs can restrict virus infection with a novel nucleic acid-based antiviral mechanism depending on the activities of endogenous RTase and RNase H1.

<sup>1</sup>Centre for Infection and Immunity Study (CIIS), School of Medicine (Shenzhen), Sun Yat-sen University, Guangzhou, Guangdong, China and <sup>2</sup>MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China

Correspondence: Deyin Guo ([guodeyin@mail.sysu.edu.cn](mailto:guodeyin@mail.sysu.edu.cn))

These authors contributed equally: Junyu Wu, Chunyan Wu, Fan Xing, Liu Cao

Received: 23 December 2020 Accepted: 26 May 2021

Published online: 22 June 2021

## RESULTS

Endogenous RTase activity inhibits virus infection in mESCs

We first showed that mES cell lines, E14TG2a and D3, were more resistant to the infection of encephalomyocarditis virus (EMCV), a single-stranded RNA virus replicating in cytoplasm, than mouse embryonic fibroblasts (MEFs) (Fig. 1a). Virus infection induced minimal level of IFN and its downstream ISGs in mESCs in comparison with that in MEFs (Supplementary information, Fig. S1a–c). As expected, markedly higher level of RTase activity was observed in mESCs than somatic cells such as MEFs and BHK21 cells (Fig. 1b). Furthermore, mESCs nearly lost the RTase activity after differentiation (Fig. 1c; Supplementary information, Fig. S2a).

We tried to investigate whether the RTase activity of ERVs contributes to antiviral immunity in mESCs by using azidothymidine (AZT), a widely used RTase inhibitor that can inhibit the endogenous RTase activity in mammalian cells (Supplementary information, Fig. S1d–g).<sup>28</sup> AZT treatment significantly increased the RNA and protein levels of EMCV (Fig. 1d, e; Supplementary information, Fig. S2b, c). AZT also increased the proportion of infected cells and virus titers in the medium (Fig. 1f, g; Supplementary information, Fig. S2d). Accordingly, the cytopathic effects of EMCV were significantly enhanced by AZT treatment (Supplementary information, Fig. S2e).

To investigate whether the antiviral function of endogenous RTase is common to other viruses, we tested the effects of AZT on mouse hepatitis virus (MHV), belonging to the genus *betacoronavirus* of the family *Coronaviridae*, which also includes severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and Middle East respiratory syndrome coronavirus (MERS-CoV). The results showed that AZT treatment significantly promoted MHV proliferation in mESCs (Fig. 1k, l; Supplementary information, Fig. S2i, j).

We further evaluated whether the endogenous RTase is involved in the suppression of DNA virus replication. The E14TG2a and D3 cells were infected with herpes simplex virus 1 (HSV-1), a DNA virus of the family *Herpesviridae*, after AZT treatment. The results demonstrated that inhibiting the RTase activity by AZT had minimal effects on the replication of viral genomic DNA and virus titers of HSV-1 (Supplementary information, Fig. S3a–d). Thus, the endogenous RTase-based mechanism may not be responsible for restricting DNA virus replication in mESCs.

The function of RTase activity was further evaluated in the primary mESCs generated from 3.5-day-old preimplantation embryos of C57BL/6 female mice, named as ESC-18 and ESC-22 (Supplementary information, Fig. S4a). These two mESCs are highly pluripotent in that they express pluripotent markers and have the potential to differentiate into all three embryonic germ layers (Supplementary information, Fig. S4b, c). We infected the mESCs with EMCV and found that AZT treatment significantly increased both viral RNA and viral titers in both ESC-18 and ESC-22 (Supplementary information, Fig. S4d, e). However, in the differentiated mESCs and somatic cells, AZT showed minimal effects on virus infection (Supplementary information, Fig. S5). Altogether, these results indicate that inhibition of endogenous RTase activity can promote virus infection in mESCs.

The expression of ERVs in mESCs and early embryos is regulated by a few of epigenetic factors, including KDM1A (also named as LSD1) which is a histone H3 lysine 4 demethylase.<sup>29–31</sup> Its catalytic inhibitor, GSK-LSD1, has been reported to upregulate the expression of a few ERVs<sup>32</sup> and could increase the endogenous RTase activity in mESCs (Supplementary information, Fig. S1h–k). As expected, GSK-LSD1 treatment significantly lowered the level of EMCV RNA, viral protein, and virus titers in mESCs (Fig. 1h–j; Supplementary information, Fig. S2f–h). The infection of MHV was also restricted by GSK-LSD1 treatment (Fig. 1k, l; Supplementary information, Fig. S2i, j).

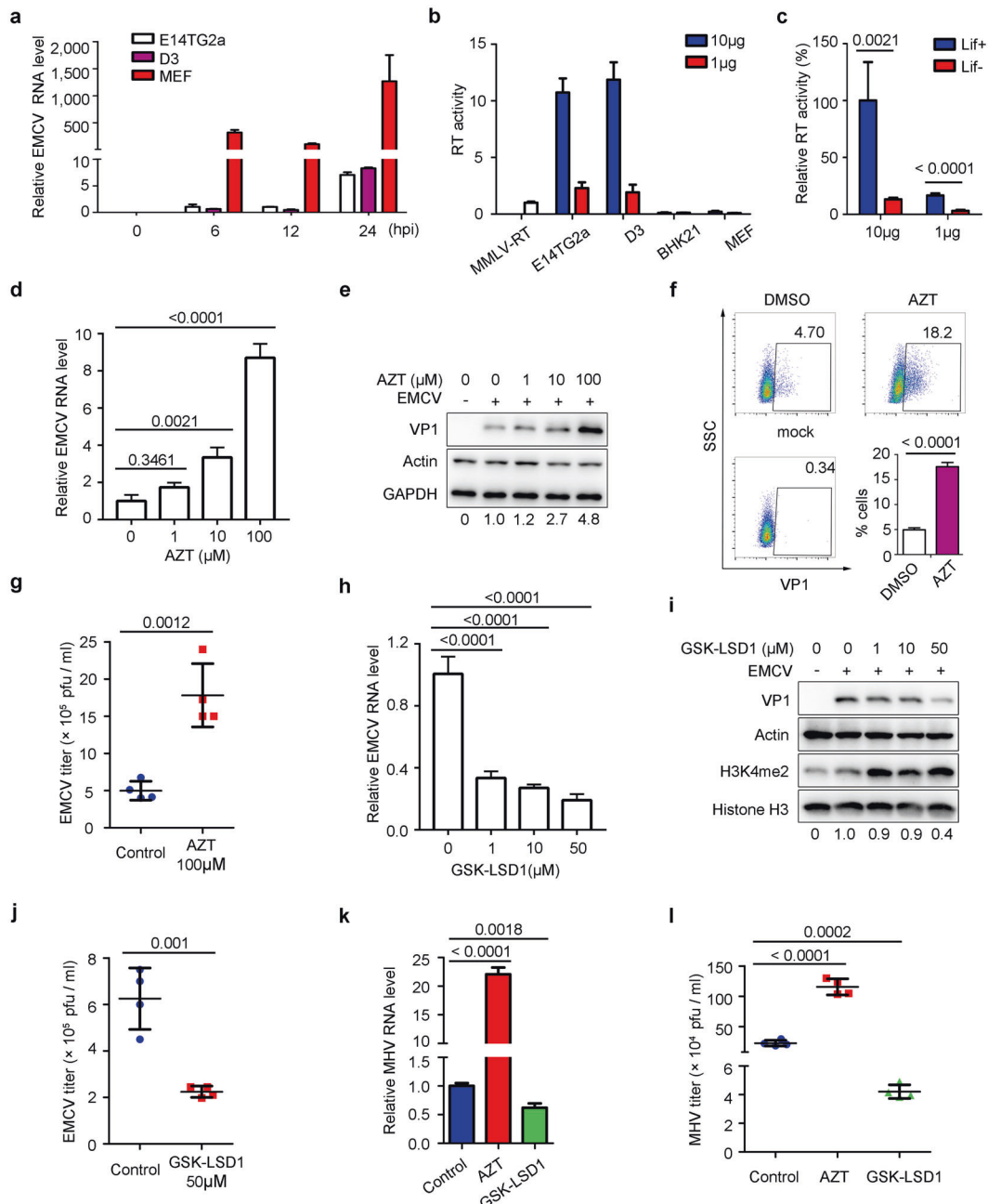
Further, we analyzed the expression levels of a representative set of ERVs, among which MusD showed a much higher level of expression in mESCs than in MEFs (Supplementary information, Fig. S2k). Transfection of the plasmid containing intact MusD could inhibit viral infection in mESCs, whereas IAP, another intact ERV, and LINE-1, the major non-long terminal repeat (LTR) retrotransposon, had minimal effects (Supplementary information, Fig. S2l–q). Furthermore, the effects of MusD on virus infection were ablated by AZT treatment. These results suggested that the RTase activity of MusD contains a potential antiviral effect. To further validate the role of MusD, IFN-deficient somatic cell was generated by CRISPR/Cas9-mediated MAVS knockout in 293T cells (Supplementary information, Fig. S6a–d), which also does not intrinsically express STING.<sup>33</sup> Consistent with that in mESCs, transfection of intact MusD inhibited virus infection in IFN-deficient 293T cells as well (Supplementary information, Fig. S6e–j).

Collectively, these results suggested that RTase activity of ERVs is involved in the suppression of virus infection in mESCs.

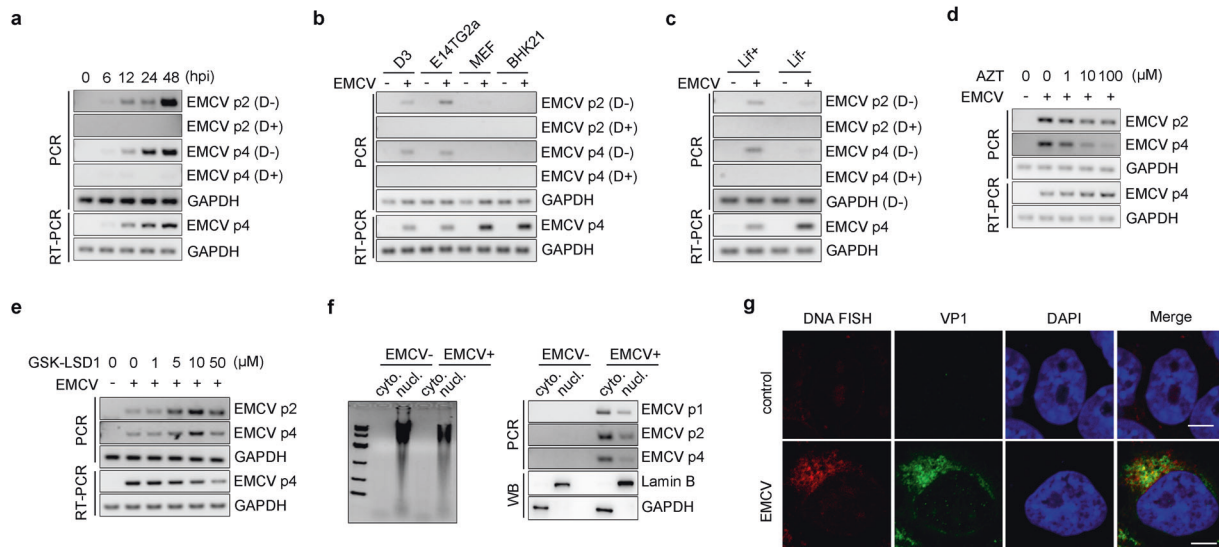
Endogenous RTase activity generates viral DNA in the cytoplasm of infected mESCs

Next, we made efforts to explore whether the antiviral function of RTase relies on one of the two known pathways, the intrinsic expression of some ISGs and Dicer-mediated RNAi, which have been reported to help ESCs restrict virus infection.<sup>13,14</sup> As shown in Supplementary information, Fig. S7a, b, although AZT treatment promoted viral infection, it did not reduce the level of the intrinsically expressed ISGs. Consistently, the expression levels of ISGs were not significantly changed after GSK-LSD1 treatment (Supplementary information, Fig. S7c, d). IFITM1/2/3 are the major intrinsic ISGs involved in the suppression of virus infection in ESCs.<sup>13</sup> Knockdown of IFITM1/2/3 has minimal effects on the viral production in mESCs with AZT treatment (Supplementary information, Fig. S7e, f). Thus, the antiviral function of RTase was independent of the expression of these ISGs. Furthermore, we knocked out Dicer, a key player in RNAi pathway, with CRISPR/Cas9 technology in mESCs (Supplementary information, Fig. S8a). As shown in Supplementary information, Fig. S8b, loss of the miR-16 expression indicated the successful blockage of Dicer activity in the knockout cell lines. However, the effects of AZT on virus infection remained in Dicer-knockout mESCs (Supplementary information, Fig. S8c, d). Consistently, GSK-LSD1 treatment could also restrict virus infection in Dicer-knockout mESCs (Supplementary information, Fig. S8e, f). These results suggested that the role of RTase in virus restriction is also independent of RNAi pathway. Collectively, the RTase activity is involved in a previously unknown antiviral mechanism in mESCs.

A few studies have reported the DNA forms of non-retroviral RNA viruses in mammalian cells,<sup>34–36</sup> but their function has not been clarified. We then tested whether the RTase activity in mESCs led to the synthesis of viral DNA (vDNA) after EMCV infection. As shown in Fig. 2a and Supplementary information, Fig. S9a, vDNA could be readily detected in mESCs as early as 12 h after infection. Furthermore, mESCs produced more vDNA than somatic cells like MEFs and BHK21 cells after viral infection (Fig. 2b), and the level of vDNA was dramatically reduced after the differentiation of mESCs (Fig. 2c; Supplementary information, Fig. S9b). Thus, we supposed that the production of vDNA may be responsible for the antiviral function of RTase in mESCs. In accordance, AZT treatment efficiently inhibited the production of vDNA (Fig. 2d; Supplementary information, Fig. S9c), whereas GSK-LSD1 treatment increased the vDNA level (Fig. 2e; Supplementary information, Fig. S9d). Cell fractionation was performed to determine the subcellular localization of the vDNA. As shown in Fig. 2f and Supplementary information, Fig. S9e, f, different from the genomic DNA, the vDNA was mainly localized in the cytoplasmic fraction with marginal signal in the nuclear fraction. Furthermore, DNA fluorescence in situ hybridization (DNA FISH) confirmed the cytoplasmic



**Fig. 1** Role of endogenous RTase in antiviral responses in mESCs. **a** mESCs were resistant to virus infection. mESCs (E14TG2a and D3) and MEFs were infected with EMCV (MOI = 1) for the indicated time periods. The RNA level of EMCV was determined by quantitative real-time PCR (qRT-PCR). **b** mESCs contain higher endogenous RTase activity than somatic cells. The endogenous RTase activity of D3, E14TG2a, BHK21 and MEF cell extracts was measured as described in Materials and Methods using MS2 RNA as templates. The levels of MS2 cDNA were determined by qRT-PCR with commercial reverse transcriptase, M-MLV, as positive control. The RTase activity was calculated relative to 1 U of M-MLV. **c** The endogenous RTase activity decreased following the differentiation of mESCs. E14TG2a cells were cultured in the medium with or without Lif for 7 days. Cells were lysed and the RTase activity was measured. The relative RTase activity was presented by setting the Lif + (10 µg) group as 100%. **d–g** Inhibition of endogenous RTase activity by AZT promoted virus infection in mESCs. E14TG2a cells were infected with EMCV (MOI = 1) after treatment with AZT at the indicated concentrations for 6 h. The RNA level of EMCV was determined by qRT-PCR (**d**) and the protein level of EMCV VP1 was analyzed by immunoblotting (**e**). Actin and GAPDH were used as loading control. Intensity of VP1 bands was quantitated by ImageJ and normalized to intensity of actin bands and the result is shown at the bottom. The viral infection rates were detected by flow cytometric analysis with VP1 antibody (**f**). The viral titers in the medium were measured by plaque assay (**g**). **h–j** GSK-LSD1 inhibited virus infection in mESCs. E14TG2a cells were pre-treated with GSK-LSD1 at the indicated concentrations for 24 h and were infected with EMCV (MOI = 1) for another 24 h. The RNA level of EMCV was determined by qRT-PCR (**h**). The protein level of EMCV VP1 and the level of histone H3K4me2 were analyzed by immunoblotting (**i**). Viral titers in the medium were measured by plaque assay (**j**). **k, l** The role of endogenous RTase in MHV infection. The E14TG2a cells were pre-treated with 100 µM AZT or 50 µM GSK-LSD1 and infected with MHV (MOI = 1) for 24 h. The RNA level of MHV was determined by qRT-PCR (**k**) and viral titers were measured by plaque assay (**l**). Data are representative of three independent experiments (**e, i**). The graphs represent means ± standard deviation (SD) from three (**a, d, f, h**) or four (**b, c, g, j, k, l**) independent replicates measured in triplicate. Statistics were calculated by the two-tailed unpaired Student's *t*-test (**c, f, g, j, k, l**) or one-way ANOVA with Tukey's post hoc tests (**d, h**).



**Fig. 2** **Generation of viral DNA in the cytoplasm of infected mESCs.** **a** Kinetics of the vDNA synthesis in mESCs. E14TG2a cells were infected with EMCV, followed by extraction of total DNA and RNA at 0–48 hpi (hours after infection). Total DNA was treated with Turbo DNase (D+) or not (D–). The virus RNA and vDNA levels were analyzed by RT-PCR (lower panel) and PCR (upper panel), respectively. GAPDH was used as loading control. The positions of the primers were marked in Fig. 3a. **b** D3, E14TG2a, MEF and BHK21 cells were infected with EMCV for 24 h. The virus RNA and vDNA levels were analyzed by RT-PCR (lower panel) and PCR (upper panel), respectively. **c** The level of vDNA decreased following the differentiation of mESCs. mESCs cultured in the medium with or without LIF for 7 days were infected by EMCV for 24 h. The virus RNA and vDNA levels were analyzed, respectively. **d** AZT inhibits vDNA synthesis in mESCs. E14TG2a cells were infected with EMCV (MOI = 1) after treatment with AZT at the indicated concentrations for 6 h. The virus RNA and vDNA levels were analyzed, respectively. **e** GSK-LSD1 promotes vDNA production in mESCs. E14TG2a cells were pre-treated with GSK-LSD1 at the indicated concentrations for 24 h and were infected with EMCV for another 24 h. The virus RNA and vDNA levels were analyzed, respectively. **f** The vDNA is mainly localized in the cytoplasm. The cytoplasmic and nuclear fractions were isolated from E14TG2a cells with or without virus infection. The DNA of each fraction was extracted and analyzed by agarose gel electrophoresis (left panel). The vDNA was detected by PCR with the indicated primers (right upper panel). Lamin B and GAPDH were used as nuclear and cytoplasmic markers by western blotting, respectively (right lower panel). **g** E14TG2a cells infected with EMCV (MOI = 1, 24 hpi) were examined for the localization of vDNA (red) and VP1 (green) by FISH. Nuclei were stained with DAPI (blue). Scale bars, 10  $\mu$ m. Data are representative of three independent experiments (a–g).

localization of the vDNA (Fig. 2g). Collectively, these results demonstrated that after virus infection, the RTase activity in mESCs could generate vDNA from viral RNA in the cytoplasm and such vDNA may play a role in the restriction of viral replication.

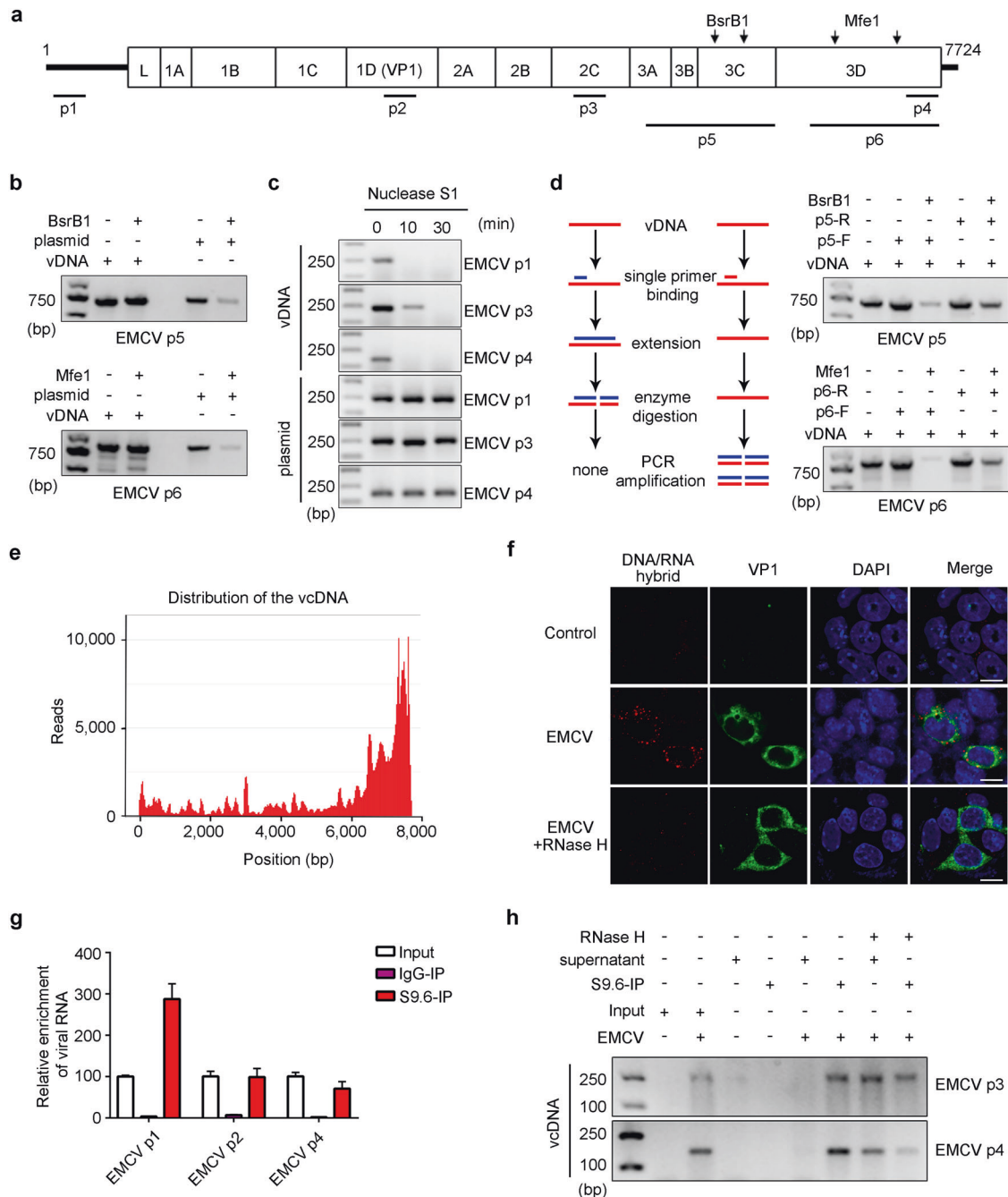
The vDNA is single-stranded and antisense to viral RNA in nature and forms DNA/RNA heteroduplex with viral RNA

We next tried to characterize the nature of the vDNA with the treatment of DNA restrictive endonucleases, BsrB1 and Mfe1 (Fig. 3a), which can only cut double-stranded DNA (dsDNA). In case of dsDNA, the treatment would result in no or much less PCR products, while in case of single-stranded DNA (ssDNA), the treatment should not reduce the yield of PCR products. As shown in Fig. 3b, PCR products of the plasmid pCMV-rNJ08 containing the EMCV genome were remarkably reduced after the treatment of restrictive endonucleases, whereas those of vDNA were not affected by the treatment. Furthermore, the vDNA was sensitive to the ssDNA-specific S1 nuclease (Fig. 3c). These results demonstrated that the vDNA was ssDNA in nature.

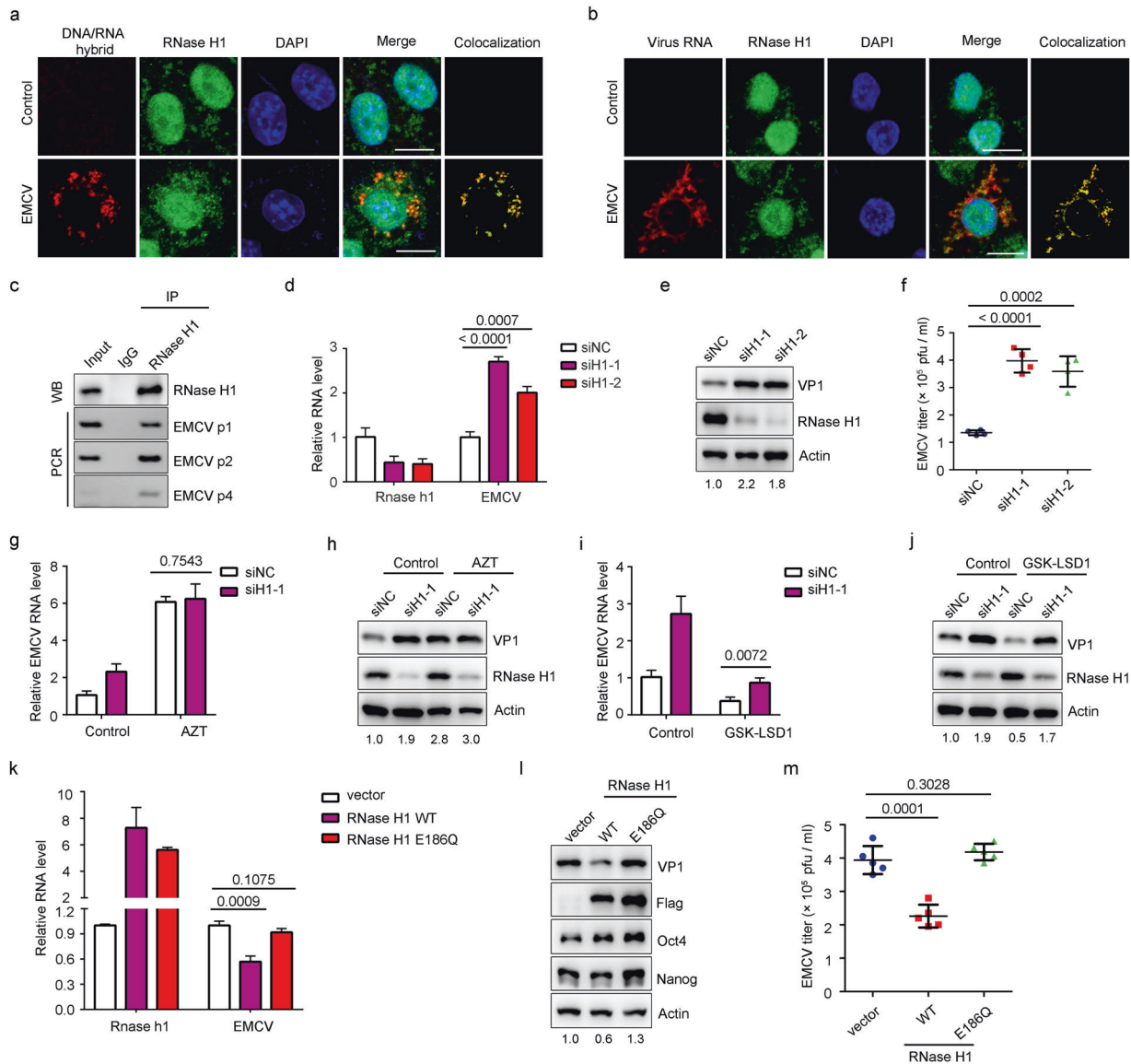
To further determine the vDNA's strand polarity relative to the viral genomic RNA, unidirectional primer extension was performed with the positive-strand primer or the negative-strand primer before the restrictive enzyme digestion.<sup>37</sup> PCR was subsequently performed. As shown in Fig. 3d, the amounts of PCR products were significantly reduced in positive-strand primer (p5-F and p6-F) extension, implying the formation of dsDNA. These results suggested that the vDNA is mainly composed of negative-sense ssDNA that is complementary to viral genomic RNA, and thus we name such viral complementary DNA as vcDNA. To investigate the distribution of the vcDNA along the viral genome, the vcDNA was enriched with specific probes, and then analyzed by unbiased

deep sequencing analysis. As shown in Fig. 3e, vcDNA sequences were not uniformly distributed along the viral genome, and there were more reads that match the 3'-part of the genome.

As the vcDNA was negative-sense single-stranded, we hypothesized that it may exist in the form of DNA/RNA hybrid with viral RNA in infected cells. There are two main types of methods usually used to detect DNA/RNA hybrid in cells: one is based on the monoclonal S9.6 antibody, which specifically recognizes the DNA/RNA hybrid structure,<sup>38</sup> and the other relies on the catalytically inactive RNase H, which recognizes DNA/RNA hybrid but cannot digest its RNA strand.<sup>39</sup> We first verified the specificity of S9.6 antibody in virus-infected somatic cells. Different from the cytoplasmic viral dsRNA signals visualized by the SCICONS-J2 antibody,<sup>40</sup> the S9.6 signals were mainly detected in the nucleus of BHK21 cells because of the R-loop structure on the chromatin<sup>41</sup> (Supplementary information, Fig. S10a), thus confirming the specificity of S9.6 antibody. This nuclear localization was also consistent with the insufficient production of vcDNA in the cytoplasm of BHK21 cells (Fig. 2b). To validate our hypothesis of the formation of DNA/RNA hybrid by vcDNA and viral RNA, virus-infected mESCs were stained with S9.6 antibody. As expected, obvious S9.6 signal was detected in the cytoplasm of virus-infected mESCs (Fig. 3f). Such signal could also be recognized by the catalytically inactive RNase H mutant (E186Q) (Supplementary information, Fig. S10b) and was sensitive to RNase H digestion (Fig. 3f). Furthermore, the DNA-RNA immunoprecipitation (DRIP) experiments were performed. As shown in Fig. 3g, h, both viral RNA and vcDNA could be immunoprecipitated by S9.6 antibody. Altogether, these results demonstrated that the vcDNA mainly existed in DNA/RNA heteroduplex together with viral RNA in infected mESCs.



**Fig. 3 Formation of DNA/RNA heteroduplex by interaction of single-stranded antisense vDNA (vcDNA) with viral RNA.** **a** Schematics of EMCV viral genome. The bottom indicates the positions of the PCR primer pairs used for the detection of vDNA. The arrows above the genome show the sites of restriction endonucleases, *BsrBI* and *MfeI*. **b** The vDNA is resistant to DNA restrictive endonucleases. The vDNA was extracted from E14TG2a cells infected with EMCV and incubated with restrictive endonuclease *BsrBI* or *MfeI* for 1 h. The integrity of DNA was evaluated by PCR with the indicated primers. The plasmid pCMV-rNJ08, which contains full-length EMCV genomic sequence, was used as positive control of dsDNA. **c** The vDNA is single-stranded. The vDNA and plasmid pCMV-rNJ08 were incubated with nuclease S1 at 1 U for the indicated time periods. The integrity of DNA was evaluated by PCR with the indicated primers. **d** The vDNA is complementary with the virus genomic sequence (vcDNA). Unidirectional primer extension was performed as described in Materials and Methods with the indicated primers. The products were then digested by the dsDNA restrictive endonucleases and PCR was performed to evaluate the integrity. **e** The distribution of vcDNA. The vcDNA was enriched with RNA probes targeting the complementary sequence of the virus genome and was sequenced as described in Materials and Methods,  $n = 2$ . **f** DNA/RNA hybrids were accumulated in virus-infected mESCs. E14TG2a cells were infected with EMCV (MOI = 1, 24 hpi). Cells were mock-treated or pre-treated with RNase H for 1 h at room temperature. The DNA/RNA hybrids (S9.6 antibody, red) and VP1 (green) were visualized by immunofluorescent staining. Nuclei were stained with DAPI (blue). Scale bars, 10  $\mu$ m. **g, h** The vcDNA forms DNA/RNA hybrids with viral RNA. The cytoplasmic nucleic acids were extracted from E14TG2a cells infected with EMCV (MOI = 1) for 48 h. The DRIP experiments were performed as described in Materials and Methods with S9.6 antibody. Normal mouse IgG was used as negative control. The viral RNA was analyzed by qRT-PCR (**g**) and the vcDNA level in the supernatant or immunoprecipitates was detected by PCR (**h**). Data are representative of three independent experiments (**b–d, f, h**). In **g**, the graph represents means  $\pm$  SD from three independent replicates measured in triplicate.



**Fig. 4 Engagement of RNase H1 cleavage in antiviral activity of vDNA in mESCs.** **a** DNA/RNA hybrids induced by virus infection recruit endogenous RNase H1. The infected and uninfected E14TG2a cells were fixed and stained with RNA/DNA hybrid-specific S9.6 antibody (red) and RNase H1 antibody (green). Nuclei were stained with DAPI (blue). Scale bars, 10  $\mu$ m. The colocalization area was shown in the right panel. **b** Viral RNA is colocalized with RNase H1. Viral RNA was visualized by RNAscope (red), followed by immunostaining with RNase H1 (green) and DAPI (blue). Scale bars, 10  $\mu$ m. The colocalization area was shown in the right panel. **c** RNase H1 binds to vDNA. EMCV-infected E14TG2a cell lysates were immunoprecipitated with anti-RNase H1 antibody. The pull-down efficiency was validated by western blotting (upper panel). The enrichment of vDNA was analyzed by PCR with the indicated primers (lower panel). Normal rabbit IgG was used as negative control. **d–f** Depletion of RNase H1 promoted virus infection. E14TG2a cells were infected with EMCV (MOI = 1) after transfection with siRNAs for 36 h. The RNA level of EMCV was determined by qRT-PCR (**d**), the protein level of VP1 was analyzed by immunoblotting (**e**), and viral titers were measured by plaque assay (**f**). Intensity of VP1 bands was quantitated by ImageJ and normalized to intensity of actin bands. The result is shown at the bottom. **g, h** RNase H1 functions downstream of endogenous RTase to inhibit viral infection. The siRNA-transfected E14TG2a cells were treated with 100  $\mu$ M AZT for 6 h and then infected with EMCV for another 24 h. Both the levels of viral RNA (**g**) and the VP1 protein (**h**) were analyzed. **i, j** The antiviral function of GSK-LSD1 is partially dependent on RNase H1. The siRNA-transfected E14TG2a cells were pre-treated with 50  $\mu$ M GSK-LSD1 for 24 h before infection with EMCV. Both the levels of viral RNA (**i**) and the VP1 protein (**j**) were analyzed. **k–m** RNase H1 inhibition of virus infection relies on its enzyme activity. The wild type and an enzymatically inactive mutant (E186Q) of mouse RNase H1 were transfected into E14TG2a cells followed by virus infection. The levels of viral RNA (**k**), the VP1 protein (**l**) and the viral titers (**m**) were analyzed. Data are representative of three independent experiments (**a–c, e, h, j, l**). The graphs represent means  $\pm$  SD from three (**d, g, i, k**), four (**f**) or five (**m**) independent replicates measured in triplicate. Statistics were calculated by the two-tailed unpaired Student's *t*-test.

RNase H1 is essential for the antiviral function of vDNA in undifferentiated mESCs  
DNA/RNA hybrid is a special structure that can be recognized and cleaved by RNase H, leading to degradation of the RNA strand.<sup>42,43</sup> Thus, we tested whether the endogenous RNase H1 is involved in the antiviral function of vDNA in mESCs. As shown in Fig. 4a, the

DNA/RNA hybrid in the infected mESCs was colocalized with RNase H1 in the cytoplasm. We used RNAscope, a highly sensitive RNA in situ hybridization assay,<sup>44</sup> to investigate the localization of viral RNA. The results showed that a portion of the viral RNA was colocalized with RNase H1 in the cytoplasm (Fig. 4b). Furthermore, the vDNA could be immunoprecipitated by the antibody of

RNase H1 in virus-infected mESCs (Fig. 4c; Supplementary information, Fig. S11a). Taken together, these results demonstrated that the endogenous RNase H1 was recruited to the heteroduplex formed by vDNA and viral RNA.

Next, we knocked down the expression of RNase H1 by siRNAs before virus infection to evaluate whether RNase H1 can regulate the virus infection. The results showed that depletion of RNase H1 significantly promoted the infection of both EMCV and MHV in mESCs, without affecting the expression of ISGs (Fig. 4d–f; Supplementary information, Fig. S11b–i). The viral production was also increased in the RNase H1-depleted primary mESCs (Supplementary information, Fig. S4f, g). However, in the differentiated mESCs and MEFs, depletion of RNase H1 showed minimal effects on virus infection (Supplementary information, Fig. S12a–d). To further investigate the relationship between the RTase and RNase H1-mediated antiviral process, we treated the RNase H1-depleted mESCs with AZT or GSK-LSD1 before virus infection. As shown in Fig. 4g, h, inhibition of RTase activity by AZT significantly increased virus infection and RNase H1 knockdown failed to further promote the virus proliferation after AZT treatment, indicating that RNase H1 exerted its antiviral function downstream of endogenous RTase. In contrast, RNase H1 depletion could partially eliminate the inhibition of GSK-LSD1 (Fig. 4i, j). This result implies that there may be other factors that also contribute to the inhibition of viral replication downstream of DNA/RNA hybrid formation, or the DNA/RNA hybrid itself may have some degree of inhibitory effect on viral replication. Furthermore, overexpression of RNase H1 inhibited virus infection, whereas the E186Q mutant of RNase H1, which is deficient in the RNase H catalytic activity,<sup>39</sup> had no effects on virus infection in mESCs (Fig. 4k–m). As expected, the overexpression of RNase H1 could not influence the virus infection in differentiated mESCs and MEFs (Supplementary information, Fig. S12e–h). Altogether, these results demonstrated that the active RNase H1 participates in the restriction of RNA virus replication downstream of endogenous RTase in mESCs.

## DISCUSSION

Unlike somatic cells, mammalian ESCs are refractory to the antiviral IFN signaling. However, ESCs are resistant to viral infection and the mechanism remains largely unknown. In this study, we revealed a previously uncharacterized antiviral mechanism in mESCs. After virus infection, the endogenous RTase reverse transcribes viral RNA into vDNA, resulting in the formation of DNA/RNA hybrid composed of vDNA and viral RNA. The heteroduplex then recruits RNase H1 and the latter hydrolyzes viral RNA in the hybrid, consequently leading to the suppression of viral replication. We named this type of antiviral mechanism as ERASE (endogenous RTase/RNase H-mediated antiviral system), implying that the invading RNA virus can be erased by ERASE (Fig. 5a).

Although synthetic antisense DNA oligonucleotides have been used in suppression of mRNA expression and virus replication,<sup>45</sup> our results present the first evidence that the duo of endogenous RTase and RNase H plays a role in natural antiviral defense by producing the vDNA complementary to the viral genome and subsequently cleaving the viral RNA. Thus, ERASE represents another type of nucleic acid-based antiviral defense, which is an addition to the previously known nucleic acid-based antiviral mechanisms including RNAi and CRISPR/Cas antiviral systems (Fig. 5b). Mechanistically, the ERASE in mESCs defends against RNA virus infection by DNA-mediated RNA cleavage, while the CRISPR/Cas system restricts DNA bacteriophages by RNA-guided DNA cleavage and RNAi prevents RNA virus infection mainly in plants and invertebrates by RNA-guided RNA cleavage.<sup>1,2</sup>

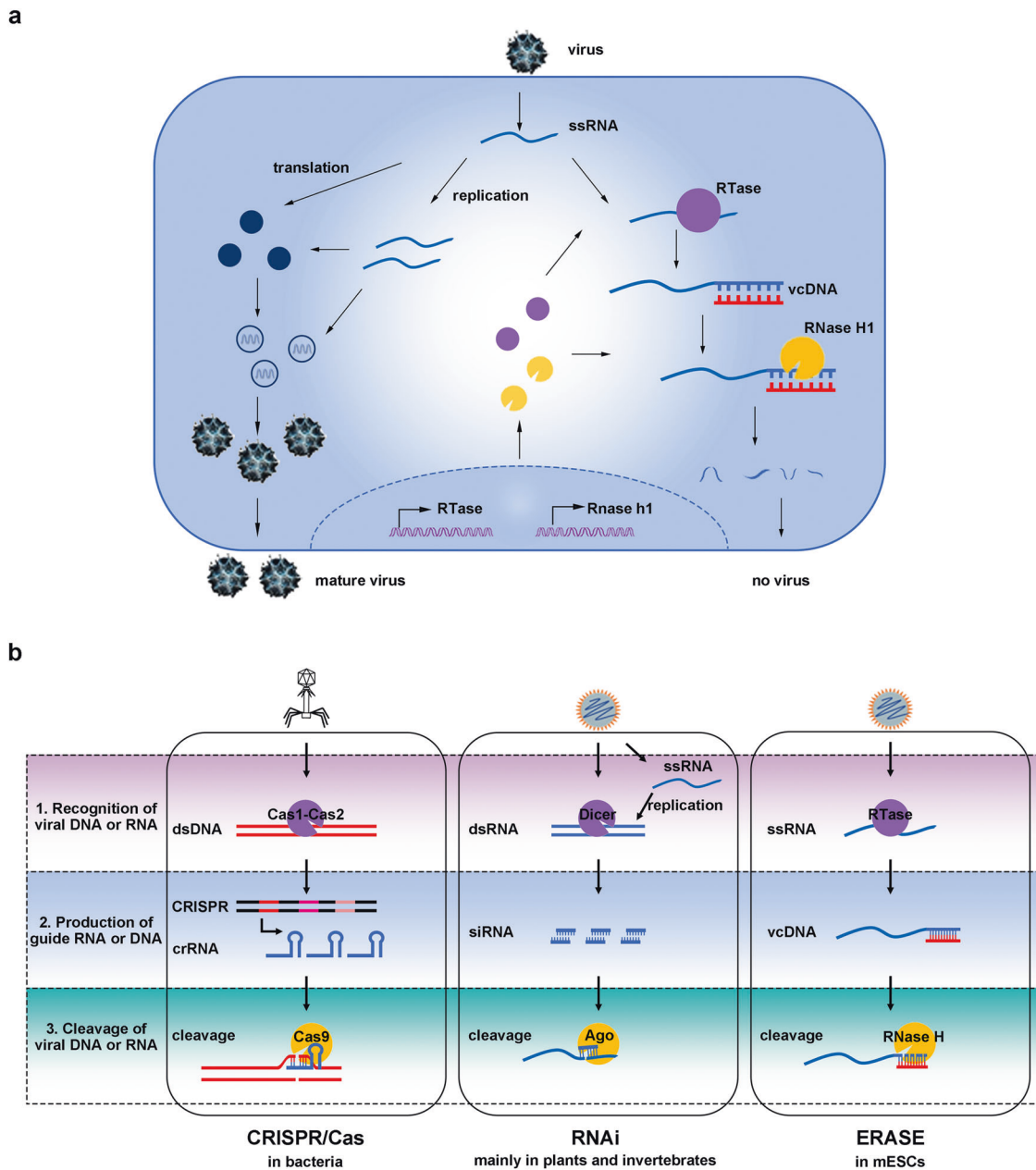
The first step of the ERASE mechanism is reverse transcription of viral RNA into vDNA by the endogenous RTases encoded by the endogenous retroelements, which constitute up to 37% and 43% of the mouse and human genomes, respectively.<sup>17,46,47</sup> The retroelements are composed of two types of elements: the LTR

retrotransposons or ERVs and the non-LTR retrotransposons, including long interspersed nuclear elements (LINEs, e.g., LINE-1) and short interspersed nuclear elements (SINES, e.g., ALU sequence).<sup>17</sup> Most of these retroelements are mutated during evolution, but there still remain some intact open reading frames in ERVs and LINEs which encode active RTase, such as LINE-1, MusD and IAP in mouse genome.<sup>26</sup> In this study, we found that MusD, but not IAP or LINE-1, could inhibit the replication of EMCV in mESCs (Supplementary information, Fig. S2l–q). However, as many types of RTase-encoding ERVs and LINEs exist in mammalian cells and are dynamically regulated,<sup>48</sup> the RTase sources may vary in different cells or at different stages of embryonic development. As reverse transcription of LINE-derived RTase mainly occurs in the nucleus while that of ERV-derived RTase occurs in cytoplasm,<sup>17</sup> there would be more chance for ERV-derived RTase to contribute to the ERASE in defense against the cytoplasmically replicating RNA viruses. In this study, we also showed that the RTase-based ERASE could not suppress the replication of DNA virus that replicates in the nucleus (Supplementary information, Fig. S3a–d). The specified types of RTases and detailed mechanism of the target RNA selection and reverse transcription priming are to be investigated in the future.

Our results showed that the RTase activity in somatic cells and differentiated mESCs were much lower than that in undifferentiated mESCs (Fig. 1b, c), and minimal vDNA was detected in virus-infected somatic cells and differentiated mESCs (Fig. 2b, c). The RTase activity is generated from endogenous retroelements and their activation could trigger IFN response in somatic cells.<sup>49</sup> Continuous activation of the IFN pathway will inhibit cell proliferation and lead to cell death. The abnormal activation of ERVs in somatic cells are often accompanied by the emergence of cancer, autoimmune diseases or neurological diseases.<sup>50</sup> Thus, in normal conditions, ERASE system and IFN pathway may not be able to coexist in the same cells. In somatic cells, the ERVs are tightly repressed by DNA methylation<sup>26</sup> and IFN pathway is a timely and efficient antiviral pathway for somatic cells to restrict virus infection. On the contrary, in mESCs, where the IFN pathway is generally deficient,<sup>6</sup> the ERASE mechanism may serve to protect cells against virus infection.

Considering ubiquitous expression of RNase H1 in host cells,<sup>51</sup> the activation of ERASE mechanism may mostly rely on the expression level of endogenous RTases, which are tightly regulated in mammalian cells.<sup>26</sup> Thus, the level of ERV activation and RTase expression may act as an indicator for the activity of ERASE. In this study, we showed that mESCs contain a relatively high level of RTase activity and ERASE activity. One important question is whether ERASE also exists in other mammalian ESCs, such as human ESCs (hESCs)? In our experimental settings, we could not show the ERASE activity in long-term cultured hESCs, which is analogous to mEpiSCs derived from the mouse post-implantation epiblast with typically “primed” pluripotency, in clear contrast with mESCs.<sup>52</sup> However, in the porcine induced pluripotent stem cells (iPSCs), a 10–100-fold higher transcription of porcine ERV elements are observed than that in porcine fetal fibroblasts.<sup>53</sup> Therefore, porcine iPSCs may possess the ERASE function as mESCs. More work is needed to characterize the ERASE activity in other mammalian species in the future.

Previous studies showed that vDNA from non-retroviral RNA viruses is important for the host survival and viral tolerance in insects.<sup>54–56</sup> However, the RTase activity and vDNA in insect cells are involved in amplifying the RNAi effect,<sup>54,55</sup> whereas the ERASE in mESCs is RNAi independent (Supplementary information, Fig. S8) and relies on the RNase H activity to cleave the viral RNA strand in the DNA/RNA hybrid (Fig. 4). Interestingly, a recent study showed that orally infected *Drosophila* can successfully eliminate the invading virus even after Ago-2 or Dcr-2 deletion,<sup>57</sup> suggesting that other antiviral mechanisms independent of RNAi may be involved in viral clearance in insects. As RNase H proteins are among the most ancient and abundant proteins in eukaryotic organisms,<sup>58</sup> it is tempting to assume that the ERASE pathway



**Fig. 5 The proposed mechanistic model of ERASE and its comparison with RNAi and CRISPR/Cas antiviral mechanisms.** **a** Working model of ERASE in mESCs. When facing RNA virus infection, the endogenous RTase catalyzes the production of vcDNA from the viral RNA. The vcDNA forms DNA/RNA hybrid structures with viral RNA. Then, the hybrids recruit endogenous RNase H1 to hydrolyze the viral RNA and inhibit virus replication. **b** The nucleic acid-based antiviral defense generally consists of three important steps: the recognition of viral nucleic acids, the generation of guide RNA or DNA, and the cleavage of viral nucleic acids. CRISPR/Cas is RNA-guided DNA cleavage by Cas enzymes to prevent DNA virus infection in bacteria. RNAi is RNA-guided RNA cleavage by Ago to defend against RNA virus infection mainly in plants and invertebrates. ERASE, in this report, is DNA-mediated RNA cleavage by RNase H to restrict RNA virus infection in mESCs. The ERASE activity in other mammalian species awaits future investigation.

may also be involved in the control of virus infection in insects and other organisms. The existence of non-retroviral RNA virus-derived sequences in the genomes of a variety of organisms, including fungi, plants, insects and mammals,<sup>36,59</sup> may be the fossil evidences of the ERASE to help ancestor hosts to defend against virus infection, and the production of vcDNA by endogenous RTase may represent a common antiviral strategy in eukaryotic organisms. Furthermore, these stable and heritable viral elements also implied that the events of virus infection and the ERASE defense happened most probably in germ cells or early stem cells.

Prokaryotes also encode retroelements with RTase, which have been characterized into three different types: retrons, group II introns and diversity-generating retroelements.<sup>60</sup> Intriguingly, during the preparation of this manuscript, two recent studies reported the antiviral function of prokaryotic retroelements.<sup>61,62</sup> Furthermore, the enzymatic activities of these retroelement-encoded RTases were shown to be essential for the antiviral defense.<sup>61</sup> Further studies are needed to investigate whether the RNase H encoded by these retroelements also participates in the antiviral system of bacteria.



Together, our study unraveled a new type of nucleic acid-based antiviral mechanism in mESCs, which consists of three major steps: recognition of viral RNA, synthesis of vcDNA by reverse transcription of viral RNA, and cleavage of viral RNA by RNase H (Fig. 5a). Thus, the ERASE mechanism represents a new addition to the previously known nucleic acid-based antiviral systems such as RNAi and CRISPR/Cas (Fig. 5b). The role of endogenous RTases as well as RNase H in the defense against RNA virus infection in mESCs implies that the viral DNA derived from non-retroviral RNAs may have a function in antiviral defense. It will be interesting to study antiviral functions of RTases and retroelements in a broad range of organisms.

## MATERIALS AND METHODS

### Plasmids and constructs

The plasmid pCMV-rNJ08, which contains a full-length cDNA copy of EMCV strain NJ08 (GenBank Accession No. HM641897), was a kind gift from Prof. Ping Jiang.<sup>63</sup> The pLentiCRISPR v2 was a gift from Prof. Feng Zhang (Addgene, #52961).<sup>64</sup> The three RTase-containing plasmids, pCMV-L1-neo<sup>RT</sup> (LINE-1), pCMV-IAP-neo<sup>TNF</sup> (IAP) and pCMV-MusD-neo<sup>TNF</sup> (MusD) were initially constructed by Prof. Thierry Heidmann lab.<sup>21,22,65,66</sup> The full-length mouse RNase H1 sequence was amplified from the cDNA of E14TG2a cells with PrimeSTAR Max DNA Polymerase (Takara, R045A) and cloned into the pEASY-T vector (Transgen, CB101-01). The enzymatically inactive mutant, RNase H1 (E186Q), was generated by introducing a glutamic acid (E)-to-glutamine (Q) mutation at amino acid 186 site of RNase H1 using site-directed mutagenesis PCR with PrimeSTAR GXL DNA Polymerase (Takara, R050A). Both wild-type and mutant RNase H1 were subcloned into the lentiviral vector pHAGE-CMV-Flag and validated by sequencing. All the primers are listed in Supplementary information, Table S1.

### Cells and cell culture

mESCs (E14TG2a and D3), BHK21 and HEK293T cells were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). Rat lung epithelial cells (L2) were kindly provided by Prof. Rong Ye (Shanghai Medical School of Fudan University). The primary mouse stem cells from embryos were isolated from 3.5-day C57BL/6 mouse embryos. MEFs were primarily isolated from 13.5-day ICR mouse embryos.

E14TG2a and D3 cells were cultured on tissue culture plates coated with 0.2% gelatin (Sigma-Aldrich, G1890). Cells were maintained in stem cell medium: KnockOut<sup>TM</sup> DMEM (Thermo Fisher Scientific, 10829018) supplemented with 15% KnockOut<sup>TM</sup> Serum Replacement (Thermo Fisher Scientific, A3181502), 2 mM Glutagro (Corning, 25-015-CI), 0.1 mM Non-essential Amino Acids (NEAA, Corning, 25-025-CIR), 0.055 mM  $\beta$ -mercaptoethanol (Thermo Fisher Scientific, 21985023), 1 mM sodium pyruvate (Corning, 25-000-CIR), 1000 U/mL LIF (Millipore, ESG1107), 3  $\mu$ M CHIR99021 (Selleck, S1263), 1  $\mu$ M PD0325901 (Selleck, S1036), 1% penicillin/streptomycin (Corning, 30-002-CI). mESCs were passaged every three days with 0.05% trypsin-EDTA (Corning, 25-053-CI). For induction of differentiation, mESCs were cultured in the differentiation medium for 7 days on 0.2% gelatin-coated tissue culture plates. The differentiation medium is high-glucose DMEM (Corning, 10-013-CVR) with 15% fetal bovine serum (FBS, Gibco, 10099141), 0.1 mM NEAA, 2 mM Glutagro, 1 mM sodium pyruvate, 1% penicillin/streptomycin, and 0.055 mM  $\beta$ -mercaptoethanol. BHK-21, HEK293T and L2 cells were maintained in somatic cell medium: high-glucose DMEM supplemented with 10% FBS and 1% penicillin/streptomycin. All cells were grown at 37 °C in a humidified incubator with 5% CO<sub>2</sub>.

**Isolation of MEFs.** The experimental procedures were reviewed and approved by the Institutional Animal Care and Use Committee of Sun Yat-sen University. The 13.5-day-pregnant ICR mice were

obtained from the Laboratory Animal Center of Sun Yat-sen University. Primary MEFs were derived from the 13.5-day embryos by following a protocol detailed elsewhere.<sup>67</sup> MEFs were cultured in the somatic cell medium. A fraction of MEFs were used in the viral infection experiments, and the rest of MEFs were treated with 10  $\mu$ g/mL mitomycin C (Selleck, S8146) for 3 h and subsequently used as feeder cells supporting the growth of the primary mESCs.

**Characterization of the primary mouse stem cells from embryos.** The primary mESCs, ESC-18 and ESC-22, were generated from 3.5-day embryos of C57BL/6 female mice in the lab of Profs. Junjiu Huang and Guang Shi by following their previously published procedures.<sup>68</sup> Before the viral infection experiments, the pluripotency of the cells was characterized. Primary ESCs were cultured on mitomycin C-treated MEF feeder cells in the stem cell medium. Their colony morphology was examined and the alkaline phosphatase activity was measured using a BCIP/NBT Alkaline Phosphatase Color Development Kit (Beyotime, C3206) according to the manufacturer's instructions. The pluripotency was further evaluated by immunofluorescent staining of pluripotent markers, SSEA-1 (EMD Millipore, MAB4301C3) and Oct4 (EMD Millipore, MAB4419A4). Their differentiation potentials into three germ layers were evaluated by the formation of embryoid bodies using a suspension culture method with the differentiation medium, as described previously.<sup>69</sup> All the virus infection experiments with the primary ESCs were performed within ten passages.

### Viruses

The EMCV virus, a member of the family *Picornaviridae*, was derived from the infectious clone pCMV-rNJ08.<sup>63</sup> The plasmid was transfected into BHK21 cells with Lipofectamine 2000 reagent following the manufacturer's protocol. Supernatants were harvested when most of cells exhibited cytopathic effects and passaged twice. Virus titers were determined by plaque assay on BHK21 cells as previously described<sup>63</sup> and the viral stocks were stored at -80 °C.

MHV strain A59, which belongs to the genus *Betacoronavirus* of the family *Coronaviridae*, is kindly provided by Prof. Rong Ye (Shanghai Medical School of Fudan University).

HSV-1, a DNA virus belonging to the family *Herpesviridae*, is kindly provided by Pei Xu (School of Medicine, Sun Yat-sen University).

For generation of lentivirus, the pLentiCRISPR v2 or pHAGE-CMV-Flag plasmids containing target sequences were co-transfected with psPAX2 and pMD2.G into the 293T cells. The supernatants were collected and concentrated by PEG8000 (Sigma-Aldrich, St. Louis, MO, USA). The viral stocks were then stored at -80 °C.

### Titration of infectious virus (plaque assay)

The plaque assay to measure virus titers was performed as previously described.<sup>70</sup> Briefly, the supernatants were collected after virus infection and were 10-fold serially diluted with fresh medium. The diluted samples were added to a confluent monolayer of BHK21 cells (EMCV), Vero cells (HSV-1) or L2 cells (MHV) plated on 24-well plates. After incubation for 2 h at 37 °C and 5% CO<sub>2</sub>, the supernatants were removed and the cells were overlaid with 1.3% methylcellulose medium. Three days later, the cells were stained with crystal violet (0.2%) overnight and plaques were counted. The virus infectious titers were calculated based on the plaque forming units (pfu) per mL.

### Drug treatment

Small molecules, AZT (Sigma-Aldrich, 2169-25MG) and GSK-LSD1 (Selleck, S7574), were used to modulate the endogenous RTase activity in mESCs. To select the proper concentration for the subsequent experiments, mESCs were seeded in 96-well plates at a density of 20,000 cells/well and treated with drugs at different

concentrations (0, 1, 5, 10, 50, 100, 200, 500, 1000  $\mu\text{M}$ ) for 24 h. Cell viability was tested by using Cell Counting Kit-8 (CCK-8, Bimake, B34302) as described previously.<sup>71</sup> According to the results, mESCs were treated with AZT at 100  $\mu\text{M}$  and GSK-LSD1 at 50  $\mu\text{M}$  in the subsequent experiments.

One day before drug treatment, cells were seeded in 12-well plates at a density of  $3 \times 10^5$  cells/well. Then, cells were pre-treated with 100  $\mu\text{M}$  AZT for 4–6 h or 50  $\mu\text{M}$  GSK-LSD1 for 24 h, followed by virus infection and other experiments.

#### siRNA knockdown

The siRNAs targeting mouse RNase H1 and IFITM1-3 were designed and synthesized by RiboBio Technology (Guangzhou, China).

Control siRNA, 5'-UUCUCCGAACGUGUACGUTT-3'

RNase H1 siRNA-1, 5'-GAGAGTCAGTCGTTGTCTA-3'

RNase H1 siRNA-2, 5'-GAGTTGGAGTGAGTGCAAA-3'

IFITM1 siRNA, 5'-GGUCUGGUCCUGUCAAU-3'

IFITM2 siRNA, 5'-GCCACAAUUCUCAAGCCUU-3'

IFITM3 siRNA, 5'-GGAUCGGUUCUGUCAGAA-3'

The transfection was carried out with Lipofectamine RNAiMAX (Thermo Fisher Scientific, 13778150). Forty-eight hours after transfection, mESCs were treated with drugs and infected with EMCV, followed by qPCR and immunoblotting assays.

#### Generation of *Dicer*-knockout mESCs and MAVS-knockout 293T cells

CRISPR/Cas9 technology was used to generate gene-knockout cell lines.<sup>72,73</sup> First, sgRNA sequences targeting mouse *Dicer* or targeting human MAVS were designed using CRISPR DESIGN (<http://crispr.mit.edu/>). The sgRNAs targeting *Dicer* (exon3; 5'-GCTCAGGGGAGACGTT-CATCG-3'), and MAVS (exon4; 5'-GGCCACCATCTGATTCTT-3') were selected and cloned into pLentiCRISPR v2 plasmid (Addgene, #52961) using the *BsmBI* restriction enzyme. After lentivirus production and transduction, cells were selected with 1  $\mu\text{g}/\text{mL}$  puromycin for 1 week and then cells were seeded in 96-well plates at about one cell per well. After enough expansion, single cell colonies with successful gene knockout were confirmed by immunoblotting. *Dicer*-knockout E14TG2a cell lines were further validated by qRT-PCR test of miR-16.

#### RTase activity assay

The cell pellets were washed with cold PBS and lysed with CHAPS lysis buffer (10 mM Tris-HCl, pH 7.5, 1 mM  $\text{MgCl}_2$ , 1 mM EGTA, 5 mM  $\beta$ -mercaptoethanol, 0.5% CHAPS, and 10% glycerol) supplemented with 1 $\times$  Complete Protease Inhibitor Cocktail (Roche, 11873580001) for 30 min on ice. Cell extracts were clarified by centrifugation at 13,000 rpm for 10 min at 4°C. The protein concentrations were measured by BCA Protein Assay Kit (Pierce, 23225). Endogenous RTase activity was tested by using product-enhanced reverse transcriptase (PERT) assay as described previously with some modifications.<sup>74</sup> In brief, bacteriophage MS2 RNA (Roche, 10165948001) was used as template in a 25  $\mu\text{L}$  reaction containing 10  $\mu\text{g}$  or 1  $\mu\text{g}$  of cell extracts, 10 ng of MS2 RNA, 20 U of RNasin RNase inhibitor (Promega, N251B), 30 pmol of MS2 reverse primer, 0.2 mM dNTP, 1 $\times$  reverse transcription buffer (Promega) and nuclease-free water. The mixture was incubated for 5 min at 25°C, 30 min at 37°C, followed by 5 min at 85°C in a PCR instrument (Bio-Rad). Negative control was set up by omitting protein extract. One unit of M-MLV reverse transcriptase (Promega, M170B) was used as positive control. The MS2 reverse-transcribed DNA (cDNA) in samples was quantitated by qPCR with MS2 specific primers (Supplementary information, Table S1). The relative RTase activities between samples were compared based on the production of MS2 cDNA.

For in vitro RTase inhibition assays, AZTTP, instead of AZT, was used because AZT does not directly inhibit the reaction of reverse transcription. Upon entering the cells, AZT is converted to its mono-, di- and tri-phosphate forms (AZTMP, AZTDP and AZTTP) by cellular nucleotide kinases. AZTTP is the active form that

competitively inhibits RTase activity and interrupts the DNA chain elongation.<sup>75</sup> Thus, increasing concentrations of AZTTP (TriLink Biotechnologies, N-4009), instead of AZT, was used in the in vitro assay to validate the function of AZT to inhibit RTase activity.

#### RNA extraction and qRT-PCR

Total RNA was extracted with TRIzol reagent (Thermo Fisher Scientific, 15596026) and reverse-transcribed into cDNA using PrimeScript RT Master Mix (Takara, RR036A) or PrimeScript RT Reagent Kit with gDNA Eraser (Takara, RR047A) according to the manufacturer's instructions. qRT-PCR was performed with PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, A25742) on the ABI QuantStudio5 (Applied Biosystems, USA). The relative abundance of target RNAs was normalized to the housekeeping gene *GAPDH*. The primer sequences were presented in Supplementary information, Table S1.

For detection of miRNAs,<sup>76</sup> the isolated total RNA was reverse transcribed with miRNA-specific stem-loop primers (miR-16 RT: GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGA TACGACCGC-CAAT; U6 RT: GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGA TACGACAAAATATG). Then, standard qRT-PCR was performed with U6 snRNA for normalization.

#### Immunoblotting

Whole cell lysates were extracted with RIPA buffer supplemented with protease inhibitors (Roche). Immunoblotting was performed as described in our previous report.<sup>77</sup> The membrane was probed with the indicated antibodies to the following proteins: GAPDH (Proteintech, 10494-1-AP),  $\beta$ -actin (Proteintech, 20536-1-AP), lamin B1 (Proteintech, 12987-1-AP), RNase H1 (Proteintech, 15606-1-AP), Flag tag (Sigma-Aldrich, F1804), histone H3 (Abcam, ab1791), H3K4me2 (Abclonal, A2356), Dicer (Abcam, ab14601), Oct4 (Cell Signaling Technology, 2840), Nanog (Cell Signaling Technology, 4903), STING (Cell Signaling Technology, 13647) and MAVS (Cell Signaling Technology, 3993). Rabbit polyclonal antibody specific for EMCV VP1 was developed using full-length recombinant VP1 protein by Abclonal Technology (Wuhan, China). The protein signals were visualized with Clarity<sup>TM</sup> Western ECL substrate (Bio-Rad, 1705061) by Tanon-5200 Multi Gel Imaging System (Tanon Science & Technology, Shanghai, China). Band intensity was analyzed using ImageJ software (National Institutes of Health, USA; <http://rsb.info.nih.gov/ij/>) and was shown below each sample.

#### Flow cytometry analysis

E14TG2a cells pretreated with drugs were infected with EMCV (MOI = 1) for 24 h. Cells were trypsinized and fixed by 1% formaldehyde in PBS for 15 min at room temperature. After being permeabilized with 0.1% Triton X-100 for 15 min, cells were blocked in 3% BSA for 30 min. Then, cells were sequentially incubated with primary antibody (anti-VP1) and secondary antibody conjugated with Alexa Fluor 555 at room temperature. The proportion of virus-positive cells was acquired using the CytoFLEX flow cytometer (Beckman Coulter, Brea, CA, USA). Data analysis was performed with FlowJo v10 software.

#### Immunofluorescent staining

Glass coverslips were coated with 20  $\mu\text{g}/\text{mL}$  poly-L-ornithine (Sigma-Aldrich, P4957) for 1 h and then 10  $\mu\text{g}/\text{mL}$  of laminin (Corning, 354232) for another 2 h at room temperature. ESCs were seeded onto the pre-coated coverslips overnight and infected with EMCV (MOI = 1) for 24 h. BHK21 cells were seeded onto coverslips overnight and infected with EMCV (MOI = 1) for 6 h. Cells were fixed with 4% paraformaldehyde (PFA) in PBS at room temperature for 15 min and permeabilized by 0.2% Triton X-100 for 20 min. After blocking with 3% bovine serum albumin (BSA; Abcone, B24726) for 30 min, cells were incubated overnight at 4°C with the indicated primary antibodies: anti-DNA/RNA hybrid S9.6 antibody (Kerafast, ENH001), anti-dsRNA J2 antibody (English and

Scientific Consulting, 10010200), anti-RNase H1 (Proteintech, 15606-1-AP), anti-VP1 antibody (ABclonal Technology, 7794). The washed cells were stained with Alexa Fluor 488- or 555-conjugated secondary antibodies for 1 h and subsequently with DAPI for 5 min at room temperature. The coverslips were then mounted using Vectashield Mounting Medium (Vector Laboratories, H-1000). Slides were imaged under a Zeiss LSM 880 Confocal Microscope (Zeiss) with a 63× oil objective using ZEN microscope software (Carl Zeiss Microimaging).

For RNase H digestion, the permeabilized coverslips were incubated with 5 U of RNase H (Thermo Scientific, EN0201) in 150  $\mu$ L digestion buffer (20 mM Tris-HCl, pH 7.8, 40 mM KCl, 8 mM MgCl<sub>2</sub>, 1 mM DTT) for 1 h at room temperature. Cells were washed twice with PBS, followed by antibody incubation.

#### DNA FISH

To investigate the presence and subcellular localization of EMCV vDNA, DNA FISH assays were performed using a Ribo<sup>TM</sup> Fluorescent in situ Hybridization Kit (RiboBio Technology, C10910) according to the manufacturer's protocols. FISH probes targeting the negative strand of EMCV genome sequence were synthesized by RiboBio Technology (Guangzhou, China). Briefly, cells were fixed with 4% PFA in PBS for 15 min and permeabilized by 0.5% Triton X-100 for 10 min. Then, cells were digested with 200  $\mu$ g/mL RNase A and 100 U/mL RNase H for 1 h at 37 °C. After two washes with PBS, cells were blocked with Pre-Hybridization buffer in the kit for 30 min at 37 °C, followed by overnight incubation with cy3-labeled probes in Hybridization buffer of the kit. On the next day, the slides were washed sequentially with Washing buffer I, II, III and PBS. Then, cells were fixed again with 4% PFA for 10 min and immunofluorescent staining was performed as above with the indicated antibodies. Images were acquired using Zeiss LSM 880 Confocal Microscope (Zeiss) with a 63× oil objective.

#### RNA in situ hybridization with RNAscope

E14TG2a cells were seeded on poly-L-ornithine/laminin-coated Millicell<sup>®</sup> EZ chamber slides (Millipore, PEZGS0416). After viral infection for 24 h, EMCV RNA was detected using RNAscope<sup>®</sup> Multiplex Fluorescent V2 Assay kit (Advanced Cell Diagnostics, 323100) according to the manufacturer's protocols. Briefly, viral infected cells were washed once with PBS and fixed with 4% PFA in PBS for 30 min. Cells were permeabilized with ethanol and pretreated with RNAscope<sup>®</sup> Protease III (Advanced Cell Diagnostics, 322340) for 10 min at room temperature, then hybridized with EMCV RNA probes (custom designed by Advanced Cell Diagnostics) according to RNAscope<sup>®</sup> manual. After hybridization, cells were blocked with 0.5% BSA in PBS and immunofluorescent staining was performed with anti-RNase H1 antibody (Proteintech, 15606-1-AP) as described above. After immunostaining, cells were mounted on a clean glass slide with DAPI Fluoromount-G mounting media (Southern Biotech, 0100-20). Fluorescence images of RNase H1 and EMCV RNA signal were acquired under a Zeiss LSM880 confocal microscope using a 63× oil-immersion objective lens (Plan-Apochromat 63×, 1.4 NA, Carl Zeiss). Confocal images were acquired at 1.0 AU pinhole, 1024 × 1024 image size, a zoom factor of 2 (pixel size 70 nm). Single stack image was used to acquire fluorescence signals of an optical slice with 640 nm thickness. For the colocalization analysis, the ZEN colocalization tool was used (ZEN software, Carl Zeiss).

#### Cellular fractionation

The cytoplasmic and nuclear fractions of mESCs were prepared as described previously with some modifications.<sup>71</sup> Briefly, EMCV-infected cells were collected and washed with cold PBS. Cell pellets were resuspended in hypotonic buffer (10 mM HEPES, pH 7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.34 M sucrose, 10% glycerol, 1 mM DTT, 0.1% Triton X-100) and incubated on ice for 10 min, followed by brief vortex. After centrifugation at 1300 rpm for 5 min at 4 °C, the supernatant was kept as cytoplasmic fraction. The pellet was

washed with hypotonic buffer and then lysed with RIPA buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% NP40, 1 mM EGTA, 0.5% sodium deoxycholate, 0.5% SDS) for 15 min on ice as nuclear fraction. Both fractions were clarified by high-speed centrifugation at 13,000 rpm for 10 min. The separation efficiency was tested by immunoblotting with subcellular markers (GAPDH and lamin B1).

#### DNA extraction and vDNA test

Cells were infected with EMCV (MOI = 1) for the indicated time periods. The whole cell lysates, cytoplasmic and nuclear fractions were treated with 100 mg/mL proteinase K (Sigma-Aldrich, P6556) in 55 °C for 1 h. The DNA was purified via phenol-chloroform extraction with the standard procedure. The isolated nucleic acid samples were incubated with RNase A (Thermo Fisher Scientific, EN0531) and RNase H (Thermo Fisher Scientific, EN0201) mix in 37 °C for 1 h. The reactions were stopped by 10 mM EDTA and the nucleic acids were purified by phenol-chloroform extraction again. For DNase treatment, the isolated nucleic acids were further incubated with Turbo DNase (Thermo Fisher Scientific, AM2238) in 37 °C for another 1 h. PCR was performed with mock-treated or DNase-digested samples using GoTaq Green Master mix (Promega, M7123). The primer sequences can be found in Supplementary information, Table S1.

For HSV-1 genomic DNA detection: mESCs were infected with HSV-1 (MOI = 0.1) for 24 h. Cells were collected and total genomic DNA was extracted with the TIANamp Genomic DNA kit (Tiagen Biotech, DP304-03). The relative abundance of HSV-1 genomic DNA was normalized to Tert. The primer sequences were presented in Supplementary information, Table S1.

#### Characterization of vDNA

The cytoplasmic DNA was purified from mESCs after being infected with EMCV (MOI = 1) for 48 h. The DNA samples were incubated with dsDNA restrictive endonuclease, *BsrBI* (New England Biolabs, R0102V) or *MfeI* (New England Biolabs, R0589V), at 37 °C for 30 min, separately. Then, the samples were subjected to PCR using the indicated primer pairs and the products were resolved on agarose gels. The plasmid pCMV-rNJ08 containing the EMCV genome was used as positive control of dsDNA.

Alternatively, the cytoplasmic DNA and plasmid pCMV-rNJ08 were treated with single-strand specific S1 nuclease (Thermo Scientific, EN0321) at 37 °C for 30 min, followed by PCR amplification and agarose gel electrophoresis.

To further investigate the strand polarity of the vDNA, one cycle of unidirectional primer extension was performed with either the positive strand primers or negative strand primers. The extension was carried out using PrimeSTAR Max DNA Polymerase (Takara, R045A) on PCR instrument with three steps: 95 °C for 30 s, 58 °C for 30 s, and 72 °C for 300 s. The products were subjected to enzyme digestion with either *BsrBI* or *MfeI*, followed by PCR amplification and agarose gel electrophoresis.

#### DRIP

The cytoplasmic fractions were prepared as described above. After RNase A digestion at 300 mM NaCl concentration for 1 h, the total nucleic acid was isolated via phenol-chloroform extraction and ethanol precipitation with the standard procedure. The nucleic acid pellet was dissolved in RNase/DNase-free water. Ten percent of samples were reserved as input. Approximately 3  $\mu$ g of S9.6 antibody (Kerafast, ENH001) was incubated with the nucleic acid in binding buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, and 0.1% sodium deoxycholate) overnight at 4 °C. Then, DNA/RNA hybrids bound with S9.6 antibody were immunoprecipitated with Pierce<sup>TM</sup> Protein A/G Magnetic Beads (Thermo Fisher Scientific, #88802) at 4 °C for 2 h. After thorough washing, bound nucleic acids were eluted in the elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.5% SDS) supplemented with 100  $\mu$ g/mL of proteinase K (Sigma-Aldrich, P6556) at 55 °C for 30 min. Both the bound and unbound fractions were further

precipitated by ethanol. For viral RNA detection, mouse normal IgG (Merck Millipore, NI03-100UG) mediated immunoprecipitation was used as negative control. After digestion by Turbo DNase (Thermo Fisher Scientific, AM2238), reverse transcription and qPCR were performed as described above with the indicated primers. The relative abundance of target regions was normalized to input values. For vDNA detection, the uninfected samples were used as negative controls. Half of the isolated nucleic acid samples were incubated with 30 U RNase H (Thermo Fisher Scientific, EN0201) in 37 °C for 2 h. The nucleic acids were purified by phenol-chloroform extraction again before immunoprecipitation with S9.6 antibody. PCR was performed directly.

#### RNase H1 pull-down and vDNA PCR

After infection with EMCV for 48 h (MOI = 1), E14TG2a cells were crosslinked by 1% formaldehyde for 10 min at room temperature. The excess formaldehyde was quenched with 125 mM glycine. Cells were washed with cold PBS and resuspended in polysome lysis buffer (20 mM Tris-HCl, pH 7.5, 85 mM KCl, 1% NP40, 2 mM DTT) supplemented with 1× Complete Protease Inhibitor Cocktail (Roche, 11873580001) and 200 U/mL RNasin RNase inhibitor (Promega, N251B) on ice for 30 min. The cytoplasmic lysates were clarified by centrifugation at 13,000 rpm for 15 min at 4 °C. One fifth volume of 5× binding buffer (250 mM Tris-HCl, pH 7.5, 750 mM NaCl, 5 mM EDTA, 5% Triton X-100, and 0.5% sodium deoxycholate) was added and mixed thoroughly. Ten percent of samples were reserved as input. The remaining lysates were divided into two aliquots and incubated with 2 µg of anti-RNase H1 (Proteintech Group, 15606-1-AP) or normal rabbit IgG (Cell Signaling Technology, #2729), respectively. After overnight incubation, Pierce<sup>TM</sup> Protein A/G Magnetic Beads (Thermo Fisher Scientific, #88802) were added and incubated for another 2 h at 4 °C. After thorough washing with binding buffer, a fraction of beads and input were boiled using Laemmli sample buffer for immunoblotting. RNase H1 precipitates were eluted using elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.5% SDS) supplemented with 100 µg/mL of proteinase K (Sigma-Aldrich, P6556) at 55 °C for 30 min. Nucleic acid purification was performed using AMPure XP Beads (Beckman Coulter, A63880) according to the manufacturer's protocol. The DNA was dissolved in 20 µL nuclease-free water and PCR was performed directly.

#### Enrichment and sequencing of vDNA

The cytoplasmic fractions were separated and nucleic acids were extracted as described above. The customized panel of probes that cover the full-length antisense EMCV genomic sequence was obtained from iGeneTech Bioscience (Beijing, China). The extracted nucleic acids were then hybridized with the probes using the TargetSeq enrichment kit (iGeneTech Bioscience, 324091-V3) according to the manufacturer's instructions. Hybridized vDNA was captured by Hydrophilic Streptavidin Magnetic Beads (New England Biolabs, S1421) and then eluted with nuclease-free water on a magnetic separator immediately after heating at 98 °C for 5 min. The enriched vDNA was fragmented by a brief sonication. The ssDNA libraries were prepared using the Accel-NGS Methyl-Seq DNA Library kit (Swift Biosciences, 30024) following the manufacturer's instructions. The libraries were sequenced on Illumina Novaseq 6000 using the paired-end sequencing strategy. The DNA library preparation, high-throughput sequencing and data analysis were conducted by Seqhealth Technology Co., Ltd (Wuhan, China).

#### Statistical analysis

Data were displayed as the means ± SD. All the statistical analyses were performed with GraphPad Prism 7.0 software using the two-tailed unpaired Student's *t*-test for two-group comparisons and one-way ANOVA with Tukey's post hoc tests for multiple-group comparisons. If *P* value is < 0.05, the difference was considered significant.

#### DATA AVAILABILITY

The vDNA sequencing raw data have been deposited in the Sequence Read Archive (SRA) database under accession number PRJNA596036. All relevant data and materials that are included in this study are available from corresponding author upon reasonable request.

#### ACKNOWLEDGEMENTS

We thank Prof. Ping Jiang (Nanjing Agricultural University, Nanjing, China) for providing the plasmid pCMV-rNJ08 carrying the cDNA clone of EMCV, Prof. Rong Ye (Fudan University, Shanghai, China) for providing MHV and L2 cells and Prof. Shan Cen (Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing, China) for providing the plasmids carrying the retro-elements. The work is supported by the National Natural Science Foundation of China (81620108020 and 32041002 to D.G.; 31800151 to J.W.; 81803568 to F.X.), Guangdong Zhujiang Talents Program (to D.G.), Shenzhen Science and Technology Program (KQTD20180411143323605 and JSGG20200225150431472 to D.G.; JCYJ20170818162249554 to F.X.) and National Ten-thousand Talents Program (to D.G.).

#### AUTHOR CONTRIBUTIONS

D.G. conceived and supervised the study; J.W. and D.G. designed the experiments, analyzed the data and wrote the manuscript. J.W., C.W., F.X., L.C., and W.Z. conducted all the major experiments. G.S. and J.H. derived the primary mESCs. L.G., P.L., Y.Z., H.J., M.L., L.B., Y.J., P.H., H.P., and C.L. participated in some of the experiments or helped with reagents and discussions.

#### ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41422-021-00524-7>.

**Competing interests:** The authors declare no competing interests.

#### REFERENCES

- Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170 (2010).
- Guo, Z., Li, Y. & Ding, S. W. Small RNA-based antimicrobial immunity. *Nat. Rev. Immunol.* **19**, 31–44 (2019).
- Yan, N. & Chen, Z. J. Intrinsic antiviral immunity. *Nat. Immunol.* **13**, 214–222 (2012).
- tenOever, B. R. The evolution of antiviral defense systems. *Cell Host Microbe* **19**, 142–149 (2016).
- Burke, D. C., Graham, C. F. & Lehman, J. M. Appearance of interferon inducibility and sensitivity during differentiation of murine teratocarcinoma cells in vitro. *Cell* **13**, 243–248 (1978).
- Wang, R. et al. Mouse embryonic stem cells are deficient in type I interferon expression in response to viral infections and double-stranded RNA. *J. Biol. Chem.* **288**, 15926–15936 (2013).
- Wang, R. et al. Antiviral responses in mouse embryonic stem cells: differential development of cellular mechanisms in type I interferon production and response. *J. Biol. Chem.* **289**, 25186–25198 (2014).
- Hong, X. X. & Carmichael, G. G. Innate immunity in pluripotent human cells: attenuated response to interferon-beta. *J. Biol. Chem.* **288**, 16196–16205 (2013).
- Wash, R. et al. Permissive and restricted virus infection of murine embryonic stem cells. *J. Gen. Virol.* **93**, 2118–2130 (2012).
- Swartzendruber, D. E., Friedrich, T. D. & Lehman, J. M. Resistance of teratocarcinoma stem cells to infection with simian virus 40: early events. *J. Cell. Physiol.* **93**, 25–30 (1977).
- Gonczol, E., Andrews, P. W. & Plotkin, S. A. Cytomegalovirus replicates in differentiated but not in undifferentiated human embryonal carcinoma cells. *Science* **224**, 159–161 (1984).
- Wolf, D. & Goff, S. P. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* **458**, 1201–1204 (2009).
- Wu, X. et al. Intrinsic immunity shapes viral resistance of stem cells. *Cell* **172**, 423–438 (2018).
- Maillard, P. V. et al. Antiviral RNA interference in mammalian cells. *Science* **342**, 235–238 (2013).
- Witteveldt, J., Knol, L. I. & Macias, S. MicroRNA-deficient mouse embryonic stem cells acquire a functional interferon response. *Elife* **8**, e44171 (2019).
- Johnson, W. E. Origins and evolutionary consequences of ancient endogenous retroviruses. *Nat. Rev. Microbiol.* **17**, 355–370 (2019).
- Kassiotis, G. & Stoye, J. P. Immune responses to endogenous retroelements: taking the bad with the good. *Nat. Rev. Immunol.* **16**, 207–219 (2016).

18. Wang, J. et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**, 405–409 (2014).
19. Kunarso, G. et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634 (2010).
20. Lu, X. et al. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.* **21**, 423–425 (2014).
21. Dewannieux, M., Dupressoir, A., Harper, F., Pierron, G. & Heidmann, T. Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells. *Nat. Genet.* **36**, 534–539 (2004).
22. Ribet, D., Dewannieux, M. & Heidmann, T. An active murine transposon family pair: retrotransposition of “master” MusD copies and ETn trans-mobilization. *Genome Res.* **14**, 2261–2267 (2004).
23. Grow, E. J. et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **522**, 221–225 (2015).
24. Matsui, T. et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927–931 (2010).
25. Rowe, H. M. et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237–240 (2010).
26. Rowe, H. M. & Trono, D. Dynamic control of endogenous retroviruses during development. *Virology* **411**, 273–287 (2011).
27. Macfarlan, T. S. et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
28. Contreras-Galindo, R., Dube, D., Fujinaga, K., Kaplan, M. H. & Markovitz, D. M. Susceptibility of human endogenous retrovirus type K to reverse transcriptase inhibitors. *J. Virol.* **91**, e01309–17 (2017).
29. Yang, B. X. et al. Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell* **163**, 230–245 (2015).
30. Macfarlan, T. S. et al. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* **25**, 594–607 (2011).
31. Shi, Y. et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
32. Sheng, W. et al. LSD1 ablation stimulates anti-tumor immunity and enables checkpoint blockade. *Cell* **174**, 549–563 (2018).
33. Sun, L. J., Wu, J. X., Du, F. H., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–791 (2013).
34. Klenerman, P., Hengartner, H. & Zinkernagel, R. M. A non-retroviral RNA virus persists in DNA form. *Nature* **390**, 298–301 (1997).
35. Geuking, M. B. et al. Recombination of retrotransposon and exogenous RNA virus results in nonretroviral cDNA integration. *Science* **323**, 393–396 (2009).
36. Horie, M. et al. Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature* **463**, 84–87 (2010).
37. Shimizu, A. et al. Characterisation of cytoplasmic DNA complementary to non-retroviral RNA viruses in human cells. *Sci. Rep.* **4**, 5074 (2014).
38. Boguslawski, S. J. et al. Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J. Immunol. Methods* **89**, 123–130 (1986).
39. Wu, H., Lima, W. F. & Crooke, S. T. Investigating the structure of human RNase H1 by site-directed mutagenesis. *J. Biol. Chem.* **276**, 23547–23553 (2001).
40. Schonborn, J. et al. Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. *Nucleic Acids Res.* **19**, 2993–3000 (1991).
41. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chedin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* **45**, 814–825 (2012).
42. Wahba, L., Amon, J. D., Koshland, D. & Vuica-Ross, M. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* **44**, 978–988 (2011).
43. Stein, H. & Hausen, P. Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. *Science* **166**, 393–395 (1969).
44. Wang, F. et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **14**, 22–29 (2012).
45. Crooke, S. T., Witztum, J. L., Bennett, C. F. & Baker, B. F. RNA-targeted therapeutics. *Cell Metab* **27**, 714–739 (2018).
46. Mouse Genome Sequencing Consortium. et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
47. Lander, E. S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
48. Baillie, G. J., van de Lagemaat, L. N., Baust, C. & Mager, D. L. Multiple groups of endogenous betaretroviruses in mice, rats, and other mammals. *J. Virol.* **78**, 5784–5798 (2004).
49. Chiappinelli, K. B. et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell* **162**, 974–986 (2015).
50. Garcia-Montojo, M., Doucet-O’Hare, T., Henderson, L. & Nath, A. Human endogenous retrovirus-K (HML-2): a comprehensive review. *Crit. Rev. Microbiol.* **44**, 715–738 (2018).
51. Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* **276**, 1494–1505 (2009).
52. Tesar, P. J. et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
53. Godehardt, A. W. et al. Comparative gene expression profiling of pig-derived iPSC-like cells: Effects of induced pluripotency on expression of porcine endogenous retrovirus (PERV). *Xenotransplantation* **25**, e12429 (2018).
54. Goic, B. et al. RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*. *Nat. Immunol.* **14**, 396–403 (2013).
55. Goic, B. et al. Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nat. Commun.* **7**, 12410 (2016).
56. Tassetto, M., Kunitomi, M. & Andino, R. Circulating immune cells mediate a systemic RNAi-based adaptive antiviral response in *Drosophila*. *Cell* **169**, 314–325 (2017).
57. Mondotte, J. A. et al. Immune priming and clearance of orally acquired RNA viruses in *Drosophila*. *Nat. Microbiol.* **3**, 1394–1403 (2018).
58. Moelling, K., Broecker, F., Russo, G. & Sunagawa, S. RNase H as gene modifier, driver of evolution and antiviral defense. *Front. Microbiol.* **8**, 1745 (2017).
59. Kondo, H., Chiba, S. & Suzuki, N. Detection and analysis of non-retroviral RNA virus-like elements in plant, fungal, and insect genomes. *Methods Mol. Biol.* **1236**, 73–88 (2015).
60. Zimmerly, S. & Wu, L. An unexplored diversity of reverse transcriptases in bacteria. *Microbiol. Spectr.* **3**, MDNA3-0058-2014 (2015).
61. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
62. Millman, A. et al. Bacterial retrons function in anti-phage defense. *Cell* **183**, 1551–1561 (2020).
63. Fang, P. et al. Construction and characterization of an infectious cDNA clone of encephalomyocarditis virus from pigs in China. *Arch. Virol.* **160**, 805–809 (2015).
64. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
65. Esnault, C., Maestre, J. & Heidmann, T. Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.* **24**, 363–367 (2000).
66. Li, X. et al. The MOV10 helicase inhibits LINE-1 mobility. *J. Biol. Chem.* **288**, 21148–21160 (2013).
67. Jain, K., Verma, P. J. & Liu, J. Isolation and handling of mouse embryonic fibroblasts. *Methods Mol. Biol.* **1194**, 247–252 (2014).
68. Huang, J. et al. Association of telomere length with authentic pluripotency of ES/iPS cells. *Cell Res.* **21**, 779–792 (2011).
69. Kurosawa, H. Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *J. Biosci. Bioeng.* **103**, 389–398 (2007).
70. Hou, P. et al. A novel selective autophagy receptor, CCDC50, delivers K63 polyubiquitination-activated RIG-I/MDA5 for degradation during viral infection. *Cell Res.* **31**, 62–79 (2021).
71. Wu, J., He, Z., Wang, D. L. & Sun, F. L. Depletion of JMJD5 sensitizes tumor cells to microtubule-destabilizing agents by altering microtubule stability. *Cell Cycle* **15**, 2980–2991 (2016).
72. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
73. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
74. Pyra, H., Boni, J. & Schupbach, J. Ultrasensitive retrovirus detection by a reverse transcriptase assay based on product enhancement. *Proc. Natl. Acad. Sci. USA* **91**, 1544–1548 (1994).
75. Kedar, P. S. et al. Mechanism of HIV reverse transcriptase: enzyme-primer interaction as revealed through studies of a dNTP analogue, 3'-azido-dTTP. *Biochemistry* **29**, 3603–3611 (1990).
76. Chen, C. et al. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **33**, e179 (2005).
77. He, Z. et al. JMJD5 (Jumonji Domain-containing 5) associates with spindle microtubules and is required for proper mitosis. *J. Biol. Chem.* **291**, 4684–4697 (2016).