1 A novel fully-spliced accessory gene in equine lentivirus with distinct Rev-responsive 2 element

3 Xiangmin Zhang, Jiwei Li, Mengmeng Zhang, Bowen Bai, Weiwei Ma, Yuezhi Lin,
4 Xing Guo, Xue-Feng Wang* and Xiaojun Wang*

5 State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research
6 Institute of Chinese Academy of Agricultural Sciences, Harbin 150069, China

7 * Corresponding author. Email:

8 * wangxuefeng@caas.cn (X-FW); * wangxiaojun@caas.cn (XJ)

9 Author Contributions

10 **Conceptualization:** Xiangmin Zhang, Xue-Feng Wang, Xiaojun Wang.

11 **Funding Acquisition:** Xue-Feng Wang, Xiaojun Wang.

12 **Investigation:** Xiangmin Zhang, Jiwei Li, Mengmeng Zhang, Bowen Bai, Weiwei
13 Ma, Xing Guo, Xue-Feng Wang

14 **Methodology:** Xiangmin Zhang, Jiwei Li.

15 **Formal analysis:** Xiangmin Zhang, Xue-Feng Wang, Yuezhi Lin.

16 **Supervision:** Xue-Feng Wang, Xiaojun Wang.

17 **Validation:** Xiangmin Zhang, Jiwei Li, Mengmeng Zhang, Weiwei Ma.

18 **Writing-Original Draft Preparation:** Xiangmin Zhang, Xue-Feng Wang, Xiaojun
19 Wang.

20 **Writing-Review & Editing:** Xiangmin Zhang, Xue-Feng Wang, Xiaojun Wang.

21 **Running title:** A novel Rev-responsive element in EIAV

22 **Abstract**

23 All lentiviruses encode an accessory protein Rev, whose main biological function is to
 24 mediate the nuclear export of unspliced and incompletely spliced viral transcripts by
 25 binding to a viral cis-acting element (termed the Rev-responsive element, RRE) that is
 26 located within env-encoding region. Equine infectious anemia virus (EIAV) is a
 27 member of the Lentivirus genus in the Retroviridae family, and is considered an
 28 important model for the study of lentivirus pathogenesis. Here, we identified a novel
 29 transcript from the EIAV genome that encodes a viral protein, named Mat, with
 30 unknown function. The transcript *mat* is fully spliced and comprises parts of the
 31 coding regions of MA and TM. Interestingly, the expression of Mat depends on Rev
 32 and the CRM1 pathway. Rev could specifically bind to Mat mRNA to promote its
 33 nuclear export. We further identified that the first exon of Mat mRNA, which is
 34 located within the Gag-encoding region, acts as an unreported RRE. Altogether, we
 35 identified a novel fully spliced transcript *mat* with an unusual RRE, which interacts
 36 with Rev for nuclear export through the CRM1 pathway. Our findings may help to
 37 expand the understanding of gene transcription and expression of lentivirus.

38

39 **Author summary**

40 In lentiviruses, the nuclear export of viral transcripts is an important step in
 41 controlling viral gene expression. Generally, the unspliced and incompletely spliced
 42 (that is, intron-containing) transcripts are exported via the CRM1-dependent export
 43 pathway in a process mediated by the viral Rev protein by binding to the Rev-
 44 responsive element (RRE) located within the Env-coding region. However, the
 45 completely spliced (intronless) transcripts are exported via an endogenous cellular
 46 pathway which is Rev independent. Here we identified a novel intronless transcript
 47 from EIAV and demonstrated that it encodes a viral protein, termed Mat. Interestingly,
 48 we have determined that the expression of Mat depends on Rev, and identified that the
 49 first exon of Mat mRNA could specifically bind to Rev and be exported to the
 50 cytoplasm, which suggests that first exon of Mat mRNA is a second RRE of EIAV.
 51 These findings update the EIAV genome structure and highlight the diversification of
 52 post-transcriptional regulation patterns in EIAV, and provides important insights into
 53 the Rev-dependent nuclear export of completely spliced transcripts in lentiviruses.

54 **Introduction**

55 Equine infectious anemia virus (EIAV) is a member of the Lentivirus genus of the
 56 Retroviridae family and primarily infects *equids*. Unlike other lentiviral infections
 57 characterized by chronic degenerative progression, EIAV infection in horses usually
 58 leads to an acute disease course [1]. Moreover, after approximately one year of
 59 infection, most infected animals are able to control viral replication and progress to
 60 inapparent carriers [2]. An inapparent carrier can still recrudescence disease under stress
 61 or immune suppression [1, 3]. These phenomena indicate that infected horses have
 62 gained immunologic control over the EIAV. Thus, the EIAV system provides a good
 63 model for the study of lentivirus replication and lentiviral vaccine development [4].

64 Lentiviruses, including the human immunodeficiency virus type 1 (HIV-1), simian
 65 immunodeficiency virus (SIV) and EIAV, like all retroviruses, are RNA viruses that
 66 replicate via a DNA intermediate (called proviral DNA) which is integrated into the
 67 host genomic DNA. In addition to the three structural proteins (Gag, Pol and Env)
 68 common to retroviruses, the lentiviral genome also encodes several accessory proteins.
 69 Current knowledge indicates that EIAV encode at least four accessory proteins (Tat,
 70 S2, Rev and Ttm) [5], while primate lentiviruses (HIV-1 and SIV) encode six
 71 accessory proteins (Tat, Rev, Nef, Vpu, Vif, Vpr/Vpx) [6, 7]. These accessory proteins
 72 are not only directly involved in viral transcriptional and posttranscriptional
 73 regulation (Tat and Rev) [8], but also regulate various stages of viral replication by
 74 interacting with host proteins (e.g., Nef, Vpu, Vif, Vpr, S2) [9-11]. At both ends of the
 75 lentiviral genome are identical long terminal repeats (LTR) [12]. As a viral promoter,
 76 5'LTR regulates viral genome transcription by binding with host RNA polymerase II,
 77 other host proteins (such as cyclinT) as well as virus-encoded proteins including Tat
 78 [8, 13].

79 The lentiviral proviral DNA acts as a transcription template to synthesize viral mRNA
80 in the nucleus. Transcription is initiated at the U3-R border of the 5'LTR and
81 terminates at the R-U5 border of the 3'LTR [14]. A full-length genome mRNA is first
82 transcribed, which is a primary unspliced transcript, and then most of this transcript
83 undergoes alternative splicing to generate multiple spliced transcripts, allowing
84 expression of several viral proteins. In HIV-1, more than 40 transcripts have been
85 found, which are responsible for the expression of the viral proteins [8]. In EIAV, five
86 transcripts have been identified in earlier studies [5, 15, 16], including an unspliced
87 full-length mRNA transcript encoding Gag and Pol protein as well as genomic RNA
88 for encapsidation into progeny virions, an incompletely spliced transcript encoding
89 Env and S2 protein, three fully spliced subgenomic transcripts encoding the accessory
90 proteins (Tat, Rev) and a predicted Ttm protein whose function is unknown. Recently,
91 our lab identified an unreported transcript in EIAV encoding the S4 protein, which
92 was found to facilitate viral release by antagonizing the antiviral activity of the host
93 protein tetherin (manuscripts, under review).

94 In lentiviruses, the synthesis and production of all viral mRNA takes place in the
95 nucleus and the mRNA is then transported to the cytoplasm by different nuclear
96 export pathways [17]. For example, in HIV-1, the unspliced full-length and
97 incompletely spliced transcripts encoding viral structural proteins (Gag, Pol and Env)
98 and some accessory proteins (vif, vpr and vpu), are exported to the cytoplasm
99 utilizing the CRM1-dependent export pathway mediated by the viral Rev protein.
100 However, all the completely spliced transcripts encoding the viral accessory proteins
101 Tat, Rev and Nef, are exported to the cytoplasm by an endogenous cellular pathway
102 used by cellular mRNAs independent of Rev [8, 18]. Other lentiviruses, such as EIAV
103 and FIV, also utilize the Rev-mediated CRM1 pathway to export incompletely spliced

mRNA transcripts [19]. The Rev protein is an accessory protein of all lentiviruses, and interacts with a specific cis-acting element termed the Rev responsive element (RRE) at the viral pre-mRNA [20, 21]. The lentiviral RRE is a highly structured viral RNA sequence that is generally located in the Env-coding region [20]. Although the lentiviral RREs lack sequence homology, they may share specific stem-loop structures. Some studies have identified that the RRE in HIV-1 is about 350 nt and that in EIAV is about 555 nt [20, 22]. The molecular details of the Rev-mediated incompletely-spliced mRNA export have been well characterized [23]. A single Rev molecule first binds to RRE, and subsequently multiple Rev molecules undergo a series of Rev-Rev oligomerizations to form a multimer, which subsequently recruits the nuclear export factor CRM1 to promote the export of the incompletely spliced mRNA into the cytoplasm. Thus, Rev binding and subsequent multimerization are necessary for Rev-RRE complex function. Previous studies have identified the RNA binding domain and multimerization of EIAV Rev [24-27]. Considering that Rev-mediated RNA export is essential for lentiviral replication, the Rev-RRE complex is a potential therapeutic target in lentiviruses.

In the present study, we discovered a novel viral protein in EIAV and named it Mat. The protein is translated from an unreported fully-spliced transcript from the EIAV genome. Interestingly, the expression of Mat depends on the Rev-mediated RNA nuclear export pathway. We provide evidence that Rev binds to Mat mRNA and is required for the nuclear export of Mat mRNA.

Results

Discovery and identification of a novel protein-encoding gene from EIAV

128 According to the genome structure and transcriptional characteristics of EIAV, we
129 designed an upstream primer (p399), and two downstream primers (p8144 and p8081),
130 as shown in Fig 1A. The PCR products were cloned and sequenced, and the sequences
131 were then aligned to the EIAV genome (GenBank accession numbers: GU385361.1).
132 In addition to the S4 and S5 transcripts that have been described in other manuscripts
133 (under review), a novel EIAV mRNA with single-splicing event was also discovered
134 by sequence alignment. The mRNA utilized a 5' splice donor site (SD) at nt863
135 (⁸⁶²TG⁸⁶³|⁸⁶⁴GG⁸⁶⁵) and a 3' splice acceptor site (SA) at nt7850 (⁷⁸⁴⁸AT⁷⁸⁴⁹|⁷⁸⁵⁰AT⁷⁸⁵¹).
136 The transcript includes a 447 bp open reading frame (ORF), and contains two exons,
137 with the first exon (393 bp) partially overlapping with the MA subunit coding region
138 of the Gag protein (in the same reading frame) and the second exon (54 bp) partially
139 overlapping the 3'terminal region of the TM coding region of the Env protein (in a
140 different reading frame). We named the novel encoding gene *mat* and its encoding
141 protein Mat. According to the splicing pattern of the mat transcripts, we designed
142 specific nested PCR primers (forward primers, p555 and p643; reverse primers, p7878
143 and p7855) to amplify part of this transcript sequence (about 250bp) from multiple
144 tissues (lymph gland, testis, liver, kidney, brain, spleen, marrow and heart) of horses
145 infected with EIAV_{LN40} (Fig 1B). The obtained PCR products were sequenced and the
146 junction between the SD⁸⁶³ and SA⁷⁸⁵⁰ sites, namely mat-specific splicing sites was
147 found. We also amplified the mat-specific bands from the equine monocyte-derived
148 macrophages (eMDM) infected with EIAV_{DLV121} and EIAV_{DLV36} *in vitro* (Fig 1C). All
149 these data suggested that there were mat transcripts in tissues or cells infected with
150 EIAV. Furthermore, to detect the expression of Mat in EIAV-infected cells, we
151 generated a mouse monoclonal antibody against the peptide (ENAKSSYISCNNASI)
152 corresponding to the translated product of the second exon of mat. Using this antibody,

we detected an about 20-kDa specific band in eMDM cells infected with EIAV_{DLV121} and EIAV_{DLV36} but not in uninfected cells (Fig 1D), which suggests that the Mat protein is expressed in EIAV-infected cells.

The expression of Mat depends on the nuclear export activity of Rev

To express the novel protein Mat in a eukaryotic expression system, its coding sequences were inserted into a pcDNA3.1 vector. We were unable to detect the expression of Mat when transfected HEK239T cells with the Mat expression plasmid alone. However, when HEK239T were cotransfected Mat with the Rev expression plasmid, Mat proteins could be observed clearly using western blotting (WB) (Fig 2A). This Mat expression pattern was also observed in fetal donkey dermal (FDD) cells (Fig 2B), which are permissive cells for the replication of EIAV [28, 29]. GFP was used here as a specific control to monitor transfection efficiencies. This result suggested that Rev is necessary for the expression of Mat. Furthermore, when HEK239T cells were transfected with different doses of Rev, the expression of Mat increased in a dose-dependent manner (Fig 2C).

It is well known that Rev is a nuclear-cytoplasmic shuttling protein that can facilitate the export of viral mRNAs from the nucleus via the CRM1-mediated pathway [8, 21]. To investigate whether the expression of Mat was related to the nuclear export activity of Rev, HEK239T cells were cotransfected with Mat and Rev plasmids and then were treated with Leptomycin B (LMB), which is an inhibitor of CRM1. As shown in Fig 2D, the levels of Mat expression significantly decreased in the presence of LMB. In addition, a Rev mutant (Rev_{L95D}), that disrupted nuclear export activity [26, 27], was unable to mediate the expression of Mat to the same extent as the wild-type Rev (Fig 2E). These results indicated that Rev-mediated Mat expression depends on the nuclear export activity of Rev through the CRM1-pathway.

178 To assess whether mat mRNAs interact with Rev protein, an RNA-pull down assay
 179 was performed using biotin-labeled mat RNA and streptavidin (STA) beads. WB
 180 analysis with anti-Flag antibody was used to detect whether any Rev protein was
 181 present in the cell lysates in the mat mRNA pulldown. The results suggest that EIAV
 182 Rev is able to specifically bind to mat and EIAV RRE mRNAs (reported RNA export
 183 element in EIAV, used as a positive-binding control) but not to HIV-1 RRE mRNAs
 184 (RNA export element in HIV-1, used as a specific-binding control) (Fig 2F).
 185 Furthermore, we isolated nuclear and cytoplasmic RNAs from HEK293Tcells after
 186 cotransfection with Mat and Rev or empty plasmids, and the RNAs were subsequently
 187 analyzed using quantitative PCR (qPCR). Our results showed that most of the mat
 188 mRNAs were retained in the nucleus in the absence of Rev, whereas the proportion of
 189 mat mRNAs in the cytoplasm increased markedly when cotransfected with Rev (Fig
 190 2G, left). By contrast, the presence of Rev made no difference to the
 191 nucleocytoplasmic ratio of GFP mRNAs (Fig 2G, right). These results imply that Rev
 192 is able to specifically promote the export of mat mRNAs from the nucleus to the
 193 cytoplasm. To evaluate the efficiency of separation, LaminB1 and α -Tubulin protein
 194 were used as controls for nuclear and cytoplasmic proteins respectively (Fig 2G,
 195 bottom). All these results together demonstrate that the expression of Mat depends on
 196 Rev, and Rev is able to promote the transport of mat mRNAs by binding to them.

197 **The first exon of mat is necessary for Rev-mediated expression of** 198 **Mat**

199 The ORF of mat contains two exons, with the second one having only 54 nucleotides.
 200 In order to map the determinants of Mat expression more precisely, we constructed a
 201 Mat-exon1 expression plasmid (Fig 3A). As expected, we found that expression of
 202 Mat-exon1 also depends on Rev, which is consistent with the expression

characteristics of Mat (Fig 3B). To further confirm this phenomenon in multiple EIAV strains, we constructed 6 mat-exon1 expression plasmids with DNA sequences, each corresponding to the N-terminus of the gag gene coding sequence from one of the different strains, as described in the Materials and Methods. Then we cotransfected these mat-exon1 expression plasmids either with Rev or with an empty plasmid. WB analysis demonstrated that in all cases, Mat-exon1 was expressed only in the presence of Rev (Fig 3C). It further confirmed that Rev is necessary for Mat-exon1 expression in the different EIAV strains. We then used an RNA pull-down assay to assess the interaction between mat-exon1 mRNA and Rev. The results demonstrated that mat-exon1 mRNA can also specifically bind to Rev (Fig 3D). In addition, we separated the cytoplasm and nucleus of the transfected cells and quantified the RNAs in the different fractions separately. The qPCR results showed that accumulation of mat-exon1 mRNAs in the cytoplasm increased in the presence of Rev (Fig 3E). These results identify the first exon of mat as being the determinant of Rev-mediated Mat expression, with the first exon demonstrating the same characteristics as the full-length mat mRNAs in terms of protein expression, Rev binding, and nucleocytoplasmic transport of mRNAs.

Different key domains of Rev underlie the mechanism by which the expression of Gag and Mat is regulated

Previous studies have described the four functional domains within Rev: a nuclear export signal (NES), an RNA binding domain (RBD), a nuclear localization signal (NLS) and a functionally unknown non-essential domain (ND) [30]. In order to map the key domain responsible for the Rev mediation of Mat expression, we constructed five Rev mutants, including the aforementioned four mutants with defects in the respective functional domains [31], and a multimerization-defective mutant Rev^{L95D}

that lost the capacity to promote export of the gag-pol mRNAs from the nucleus to the cytoplasm [27] (Fig 4A). A Mat expression vector was cotransfected into HEK293T cells with each Rev mutant separately. After 48 h, the level of Mat protein was assessed using WB and the distribution of mat mRNAs in the cytoplasm and nucleus was analyzed using qPCR as the ratio of nuclear to cytoplasmic mat mRNA. The WB analysis results showed that only the Rev_{WT} and Rev_{mNES} were able to mediate the expression of Mat, while the other four mutants almost completely lost this function (Fig 4B). The qPCR analysis also showed that mat mRNAs were mainly distributed in the cytoplasm in the presence of Rev_{WT} or Rev_{mNES}, while the other Rev mutants showed results similar to the absence of Rev, where mat mRNA was mainly distributed in the nucleus (Fig 4C). These results indicate that the RBD, NLS and ND domains of Rev and their multimerization play an important role in mediating the nuclear export of mat mRNAs.

In addition, a Rev-dependent EIAV Gag-Pol expression vector, pGP-RRE (described in the materials and methods), was transfected into HEK293T cells in the presence of Rev_{WT} or different Rev mutants. The levels of Gag protein were assessed using WB and the distribution of gag mRNA in the cytoplasm and nucleus was analyzed using qPCR. No Gag protein was detected in the absence of Rev or in the presence of Rev_{L95D} (Fig 4D, lanes 1 and 7). Compared to those in the presence of Rev_{WT}, very little Gag protein was detected in the presence of Rev_{mNES}, Rev_{mRBD} or Rev_{mNLS} (Fig 4D, lanes 3, 4 and 5), but none of Gag proteins expression level showed remarkable changes in the presence of Rev_{ΔND} compared with in the presence of Rev_{WT} (Fig 4D, lane 6). The qPCR analysis also showed that gag mRNAs were mainly distributed in the cytoplasm in the presence of Rev_{WT} or Rev_{ΔND}, while the presence of other Rev mutants gave results similar to those obtained in absence of Rev, with the gag mRNAs

being mainly distributed in the nucleus (Fig 4E). Consistent with the results of previous studies [30], these results also indicate that in addition to ND, other functional domains of Rev play an important role in mediating the nuclear export of gag-pol mRNAs.

Taken together, these results show that the expression levels of Mat or Gag protein are related to the accumulation of the mRNAs in the cytoplasm. Interestingly, we found that the export of mat and gag mRNAs from the nucleus relied on different function domains of Rev, indicating that Rev mediates the expression of Gag and Mat through different mechanisms.

The first exon of Mat cannot replace the known RRE to mediate EIAV Gag/Pol expression

As described above, the first exon of mat can be used as an RNA export element to bind to Rev to mediate the expression of Mat. Therefore, in order to test whether the mat-exon1 can replace the known RRE to mediate the expression of Gag/Pol, a series of EIAV Gag/Pol constructs was prepared: the pGP contained only the EIAV gag/plo-coding sequences; pGP-RRE, pGP-MA and pGP-3×MA were based on pGP with an EIAV RRE (22), or one copy of mat-exon1, or three copies of mat-exon1 inserted respectively downstream of the gag/pol coding sequence (Fig 5A). HEK293T cells were transfected with these constructs in the presence or absence of Rev, and the levels of Gag protein in the cells were analyzed using WB. Gag expression could only be detected in cells transfected with pGP-RRE in the presence of Rev (Fig 5B, lane 4), while we could not detect Gag expression in cells transfected with other constructs with or without Rev. Therefore, these results indicate that the Mat first exon sequence within the Gag/pol-coding sequence was unable to mediate the expression of Gag/Pol under the effect of Rev (Fig 5B, lane 2), and the first exon of mat cannot replace RRE

to mediate the expression of Gag/Pol (Fig 5B, lanes 6 and 8). Taken together, these results show that single or multimer copies of mat-exon1 cannot replace RRE function in the EIAV Gag/Pol expression vector.

Discussion

Lentivirus transcription takes place in the nucleus, and the various transcripts produced by transcription usually undergo nuclear export via two pathways for gene expression [8]. In the case of HIV-1, nuclear export of the completely spliced transcripts is mediated by cellular mRNA transporter proteins, while nuclear export of the unspliced full-length and incompletely spliced transcripts is mediated by Rev through binding to RRE via the CRM1 pathway. In EIAV, only one RRE has been found to date, and this is located in the Env-coding region [22, 32]. In this report, we found a novel fully-spliced transcript in EIAV, encoding a 20kDa protein, and named it Mat. Interestingly, we demonstrated that the expression of Mat depends on Rev, that Rev can increase the nuclear export of Mat mRNAs, and that Rev can specifically bind to Mat mRNAs. Furthermore, the first exon of Mat mRNAs (Mat-exon1 mRNA) has been identified as a key region for Rev-mediated nuclear export. These results indicate that, in addition to the reported previously RRE located in the Env-coding region in EIAV, there is a second RRE in the Mat coding region. These findings have updated our knowledge of EIAV genome structure and its gene transcription pattern, as illustrated in Fig 6.

In lentiviruses, alternative splicing is a common phenomenon that promotes the production of various transcripts encoding different viral proteins. Previous studies have shown that five major transcripts are produced from the EIAV genome [15, 16], encoding three structural proteins, three accessory proteins, and a protein with

unknown function, Ttm. Except for the unspliced full-length transcripts, all spliced transcripts share a common 5'-splice donor site. We recently found an unreported transcript in EIAV encoding a protein that antagonizes the antiviral activity of the host protein Tetherin (manuscripts, under review). In this study, we once again report a novel EIAV transcript, which is produced by a single splicing event, with a 5'-splice donor site that is different from other reported splicing transcripts. This suggests that EIAV alternative splicing patterns are complex. Importantly, we have confirmed that this transcript encodes a novel viral protein, here designated Mat. Although the biological function of Mat is unclear at present, our research results have further updated our knowledge of the genome structure of EIAV and of its patterns of gene transcription.

Cis-acting repressing sequence (CRS) is a type of RNA sequence element that regulates gene expression by inhibiting the nuclear export of mRNAs [33]. It is widely present in a variety of viruses including HIV-1 [34-36], EIAV [37], HBV [38] and HPV [39], as well as in cellular transcripts. Although no uniformly recognizable sequence commonality was detected in CRS, it has been determined that a variety of host proteins can specifically bind to CRS and restrict RNA in the nucleus [40-42]. In HIV-1, several CRS have been identified in multiple regions such as gag, pol and env genes and it has been demonstrated that these inhibitory sequences can prevent nuclear export of unspliced and partially spliced transcripts, but the supply of Rev overcomes this inhibition [34, 36, 43]. A previous study by Rosin-Arbesfeld et al. also found that a CRS exists in the center of the EIAV genome (the C-terminus of the Pol coding region), and also demonstrated that this CRS can inhibit the nuclear export of viral transcripts [37]. In this study, we found that the mat or mat-exon1 gene constructs showed poor expression and cytoplasmic RNA accumulation in the absence of Rev,

327 suggesting the existence of a potential CRS in the first exon of the mat gene.
 328 Interestingly, a CRS has been identified at the N-terminus of the HIV-1 gag gene in a
 329 previous study [35]. In fact, the first exon of the mat gene is also located at the N-
 330 terminus of the EIAV gag gene. However, the gag genes of HIV-1 and EIAV lack
 331 sequence homology. Importantly, the expression of Mat or Mat-exon1 was rescued by
 332 co-expression with Rev, indicating that Rev may overcome the inhibitory effect of the
 333 mat-exon1 gene, which also suggested that mat-exon1 sequences contain an element
 334 enabling Rev-mediated nuclear export, such as an RRE. These results suggest that
 335 both a CRS and RRE may exist simultaneously in the mat-exon1, or the two may
 336 overlap with each other. Notably, the overlap of a CRS and RRE has been
 337 documented HIV-1 [43].

338 Interestingly, we found that the expression of Mat is dependent on Rev. Rev is an
 339 accessory protein necessary for lentiviral replication, and its biological function is to
 340 mediate the transport of the unspliced and incompletely spliced viral mRNAs from the
 341 nucleus to the cytoplasm. Mechanistically, Rev binds to an RRE located within the
 342 Env-coding region, then polymerization occurs, and the polymerized protein
 343 facilitates the export of the Rev-RNA complex into the cytoplasm through a CRM1-
 344 dependent pathway. We found that Rev could specifically bind to mat mRNAs, and
 345 that mutation of the Rev multimerization site disrupted its ability to mediate Mat
 346 expression. In addition, the use of the CRM1 inhibitor LMB significantly decreased
 347 Mat expression levels. These results demonstrate that the expression of Mat depends
 348 on the nuclear export activity of Rev. Previous studies have identified the RRE of
 349 EIAV that is located in a 555 nt region near the 5' end of env gene [22, 32]. In this
 350 study, we found that Rev could bind to Mat-exon1 mRNAs and mediate their nuclear
 351 export, and this phenomenon has been verified in 5 other EIAV strains. This indicates

that the first exon (393 nt) of mat mRNAs is the second known RRE of EIAV. We named it RRE-2, with the reported RRE located in the Env coding region defined as RRE-1. In HIV-1, an additional RRE was also found in the packaging signal located in 5'-untranslated region (5'UTR) of the genome [44, 45]. Here this HIV-1 RRE is termed 5'RRE. Mutation of 5'RRE led to reduced nuclear export of viral genomic RNA (gRNA) and decreased gRNA packaging efficiency but did not affect the production of viral proteins and the export of viral particles, however, replication of the virus was severely disrupted [45]. Rev is an accessory protein encoded by mouse mammary tumor virus (MMTV) and is a homolog of the Rev protein in lentivirus. MMTV, a member of the genus Betaretrovirus, harbors two Rev-responsive elements (RREs): a 5'RRE located at the 5'UTR of gRNA that is required for nuclear export of unspliced RNA, and a 3' RRE located at env-U3 junction region that is needed for expression of both unspliced and spliced transcripts [46]. Therefore, EIAV is different from HIV-1 and MMTV in that its two RREs are located in the coding region sequence. We noticed that RRE-2 plays the nuclear export function of Rev-RRE only in the context of the mat transcript, but not in the context of gag/pol mRNA. However, RRE-1 is able to mediate transport of both Env-encoding transcripts and Gag/Pol-encoding transcripts. Possible explanations for this difference include that the nuclear export ability of Rev-RRE-2 complex is weaker than that of Rev-RRE-1, or that the structure of RRE-2 may be more susceptible to interference from adjacent sequences.

We also observed that the nuclear export functions of Rev-RRE-1 and Rev-RRE-2 are dependent on different amino acid sites and regions of Rev. This suggests that the molecular mechanisms by which Rev mediates nuclear export of diverse target RNAs may differ, although to date no similar phenomenon has been reported in lentiviruses.

We speculate that this may be related to the three-dimensional structure of Rev or to the involvement of host factors [20, 27, 47].

In summary, we identified a novel protein, Mat, from EIAV, and demonstrated that the expression of Mat depends on Rev, and that the first exon of mat mRNA acts as a Rev binding element to bind Rev and mediate Mat expression via CRM1 pathway. These findings update the EIAV genome structure and highlight the diversification of post-transcriptional regulation patterns in EIAV, and provides important insights into the Rev-dependent nuclear export of completely spliced transcripts in lentiviruses.

Materials and Methods

Plasmids

The Mat expression plasmid, pcDNA-Mat-HA, was constructed by inserting the mat gene sequence (GenBank accession numbers: OM417133) into a pcDNA3.1 (+) vector and fusing HA-tags to the C-terminal. To construct a codon-optimized Mat expression plasmid, pcDNA-optMat-HA, we cloned the sequence corresponding to mat-exon1 from the codon-optimized Gag plasmid and inserted it before the mat-exon2 sequence. The plasmid pcDNA-Mat-exon1_{FDDV}-HA was constructed by cloning the first exon of mat gene from pCMV3-8 [48] and inserting it into pcDNA3.1-HA. The other 5 Mat-exon1 expression plasmids were synthesized according to the following EIAV strains genome sequence: EIAV_{UK} (GenBank accession numbers: AF016316.1; nt450-836), EIAV_{Miyazaki-2011-A} (GenBank accession numbers: JX003263.1; nt450-836), EIAV_{Cornwall} (GenBank accession numbers: MH580898.1; nt231-617), EIAV_{DE} (GenBank accession numbers: KM247554.1; nt114-500) and EIAV_{NC-Dagmar} (GenBank accession numbers: MH820164.1; nt100-486). The Rev expression plasmid, VR-Rev-Flag, was constructed by inserting the EIAV Rev gene

sequence into a VR1012 vector and fusing Flag-tags to the C-terminal. VR-Rev_{HIV}-Flag was constructed by substituting the HIV-1 Rev gene sequence (Genbank accession numbers: AF324493.2) into VR-Rev-Flag. A series of Rev-mutant plasmids were generated using site-directed mutagenesis on the template plasmid VR-Rev-Flag. The complete EIAV gag-pol gene was PCR amplified from the molecular clone pCMV3-8 and inserted into the pEGFP-N1 vector (Clontech, USA) to construct the plasmid pGP. The Rev-responsive element (RRE) was cloned from pCMV3-8, and inserted into the EcoR1 and Not1 sites of pGP to construct the pGP-RRE plasmid. The mat-exon1 sequence was cloned as a EcoR1-Not1 fragment into pGP, which obtained plasmids containing a single mat-exon1(pGP-MA) as well as multimer copies of the mat-exon1(pGP-3MA). All constructed plasmids were verified by sequencing.

Cell culture and transfection

Human embryonic kidney epithelium 293T (HEK293T) cells and Hela cells were cultured in DMEM-high glucose (Sigma-Aldrich, USA) supplemented with 10% fetal bovine serum (Sigma-Aldrich, USA) and 1% penicillin-streptomycin (Gibco, USA). Equine monocyte-derived macrophages (eMDM) cells were isolated from fresh heparinized horse whole blood as previously described [49], and cultured in RPMI 1640 (Sigma-Aldrich, USA) supplemented with 20% donor equine serum (HyClone, USA) and 40% newborn bovine serum (Ausbio, Australia). Fetal donkey dermal (FDD) cell cultures were maintained in Minimal Essential Medium (α -MEM, Gibco, USA) supplemented with 10% fetal bovine serum (Sigma-Aldrich, USA) and 1% penicillin-streptomycin (Gibco, USA) [50]. All cells were maintained in a humidified incubator at 37°C containing 5% CO₂. Cells were transfected with the indicated plasmids using PolyJet DNA transfection

reagent (SignaGen, USA), following the manufacturer's instructions. EGFP-expression construct (pEGFP-N1) was added to all transfections to monitor transfection efficiency. All transfections were repeated at least three times.

Amplification and identification of mat

For the amplification of mat, polyadenylated RNAs were prepared from frozen horse tissues infected with EIAV_{LN40} using a Trizol method [51]. The tissues were homogenized in liquid nitrogen, and 1ml of Trizol reagent was used to incubate each sample. RNAs were precipitated with 0.5ml isopropyl alcohol and then washing with 1ml 75% ethanol to obtain purified total RNAs. For further identification of the mat transcript, an EIAV virulent strain (EIAV_{DLV36}) or vaccine strain (EIAV_{DLV121}) at equivalent titer was used to infect eMDM cells cultured in T25 flasks [49]. At 4 days post-infection, cells were harvested and total RNAs were extracted using a Bio-fast simply P RNA extraction kit (Bioer, China).

About 1µg of RNAs were reverse transcribed into cDNAs using a PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Japan) following the manufacturer's protocol. EIAV transcripts were amplified with Kod FX Neo polymerase (Toyobo, Japan) in polymerase chain reactions (PCR). Nested primer pairs included: p399 (5'-GGACAGCAGAGGAGAACTTACAG-3'; forward primer), p8144 (5'-AAGGGACTCAGACCGCAGAATCT-3'; reverse primer) / p8081 (5'-TAAAAACAGGAASTTAACGCGTCAC-3'; reverse primer) were used to amplify EIAV transcripts. Nested primer pairs: p555 (5'-TCAAAAGCTAACTAATGGTAAC-3'; forward primer) / p7878 (5'-ATTGAGGCATTGTTACATGAGA-3'; forward primer), p643 (5'-CAATTAAGGGACGTCATTCCAT-3'; reverse primer) / p7855 (5'-GTAGCTGGATTAGCATTTTCC-3'; reverse primer) were used to identify mat-

specific transcripts. The amplification conditions were as follows: 2min at 98°C; 35 cycles of 30s at 98°C, 30s at 56°C and 10-30s at 72°C; and a final extension step of 10min at 72°C. PCR products were separated using electrophoresis on 1% agarose gels in 1×Tris -acetate-EDTA, and were visualized by staining with ethidium bromide. The amplified fragments were purified using a gel extraction kit (Vazyme, China) and cloned into a pMD-18T vector (Takara, Japan).

Western blotting

The expression of the indicated plasmids was analyzed using western blotting as previously described [52]. In brief, cells were harvested 48 hpt and resuspended in lysis buffer (150mM Tris-HCl [PH 7.6], 50mM NaCl, 5mM EDTA, and 1%Triton X-100). Total protein extracts were separated by SDS-PAGE (Genscript, USA) and then blotted to NC membranes (Millipore, Germany). The membrane was subsequently blocked for 2 h with 5% skim milk (Biosharp, China) in PBS at room temperature. PBS containing 5% skim milk was used to dilute primary and secondary antibodies. After washing, the membrane was analyzed on the LI-COR Odyssey Imaging System (LI-COR, USA). All experiments were performed at least in triplicate.

Antibodies

Mat proteins in eMDM cells infected with EIAV_{DLV121} and EIAV_{DLV36} were detected using a mouse anti-Mat monoclonal antibody (mAb), which was produced by immunization of mice with a Keyhole Limpet Hemocyanin (KLH) carrier protein containing a polypeptide corresponding to the C-terminal 15 amino-acids (ENAKSSYISCNNASI) of Mat. Gag proteins were analyzed using a mouse anti-p26 mAb (9H8), kept in our lab [53]. The primary antibodies purchased from commercial sources are listed below: mouse anti-HA mAb (Sigma, USA), mouse anti-FLAG mAb

(Sigma, USA), mouse anti-GFP mouse mAb (Affinity Bioscience, USA), mouse anti-LaminB1 mAb (Proteintech, China), mouse anti- α tubulin (Abcam, UK), and mouse anti- β actin mAb (Sigma, USA). Goat anti-mouse IRD800-conjugated mAb (Sigma, USA) was used as secondary antibody for western blotting.

Cell fractionation and quantitative PCR analysis

For the preparation of nuclear and cytoplasmic RNAs, HEK293T cells were harvested 36 hpt and purified using a PARIS protein and RNA isolation kit (Invitrogen, Thermo Fisher Scientific, USA) according to the manufacturer's protocols. Equal volume RNAs were subjected to reverse transcription using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Japan). Distribution of mat mRNAs or mat-exon1 mRNAs in transfected cells was assessed using SBYR-Green (Takara, Japan)-based quantitative PCR analysis on Aligent Mx3005P. The primers used are as follows: Mat-forward (5'-AAGGGACGTCATTCCATTGTT-3') and Mat-reverse (5'-ATTTGTAAGCCCATCTTAACG-3'); GFP-forward (5'-GCAAGCTGACCCTGAAGTTCATC-3') and GFP-reverse (5'-GTCTTGTAGTTGCCGTCGTCCTT-3'); Mat-exon1-forward (5'-AGTTAGAGAAGGTGACGGTAC-3') and Mat-exon1-reverse (5'-CAACAATGGAATGACGTCCCT-3'); Gag-forward (5'-CGATGCCAAATCCTCATTAG-3') and Gag-reverse (5'-CTGATCAAAAGCAGGTTCCATCT-3').

In vitro transcription and RNA-pull down assay

Template plasmids were digested using NotI (Thermo Scientific, USA) to generate linearized DNAs, which were subsequently transcribed to produce copies of the mRNAs using an *in vitro* transcription kit (New England Biolabs, UK) according to the manufacturer's protocol. The equimolar mRNAs were labeled using biotin with an

RNA 3' End Desthiobiotinylation Kit (Thermo scientific, USA). The labeled mRNAs were then captured using streptavidin magnetic beads (Thermo scientific, USA). After washing three times with Protein-RNA Binding Buffer (20mM Tris [pH 7.5], 50mM NaCl, 2mM MgCl₂, 0.1% Tween-20 Detergent), 100µg of the total protein extracted from transfected cells was added to the RNA-bound beads and incubated for 1h on a roller at 4°C. Beads were then washed five times with Washing Buffer (20mM Tris [pH 7.5], 10mM NaCl, 0.1% Tween-20 Detergent), and the beads containing RNA-binding protein complexes were harvested and analyzed with Western blotting.

Acknowledgements

The authors thank Yandong Tang for their constructive suggestions in this study. This study was supported by grants from the National Natural Science Foundation of China (32170169 to X. W, 32172831 and 31672578 to XF.W), the Natural Science Foundation of Heilongjiang Province of China (C2017076 to XF.W), and postdoctoral scientific research developmental fund of Heilongjiang Province (LBH-Q20059 to XF.W).

References

1. Leroux C, Cadore JL, Montelaro RC. Equine Infectious Anemia Virus (EIAV): what has HIV's country cousin got to tell us? Vet Res. 2004;35(4):485-512. Epub 2004/07/09. doi: 10.1051/vetres:2004020 V4012 [pii]. PubMed PMID: 15236678.
2. Hammond SA, Li F, McKeon BM, Sr., Cook SJ, Issel CJ, Montelaro RC. Immune responses and viral replication in long-term inapparent carrier ponies inoculated with

- 521 equine infectious anemia virus. J Virol. 2000;74(13):5968-81. Epub 2000/06/14.
522 PubMed PMID: 10846078; PubMed Central PMCID: PMC112093.
- 523 3. Kono Y, Hirasawa K, Fukunaga Y, Taniguchi T. Recrudescence of equine
524 infectious anemia by treatment with immunosuppressive drugs. Natl Inst Anim Health
525 Q (Tokyo). 1976;16(1):8-15. Epub 1976/01/01. PubMed PMID: 177894.
- 526 4. Craig JK, Montelaro RC. Lessons in AIDS vaccine development learned from
527 studies of equine infectious, anemia virus infection and immunity. Viruses.
528 2013;5(12):2963-76. Epub 2013/12/10. doi: 10.3390/v5122963
529 v5122963 [pii]. PubMed PMID: 24316675; PubMed Central PMCID: PMC3967156.
- 530 5. Cook RF, Leroux C, Issel CJ. Equine infectious anemia and equine infectious
531 anemia virus in 2013: a review. Vet Microbiol. 2013;167(1-2):181-204. Epub
532 2013/11/05. doi: 10.1016/j.vetmic.2013.09.031
533 S0378-1135(13)00470-7 [pii]. PubMed PMID: 24183747.
- 534 6. German Advisory Committee Blood SAOPTbB. Human Immunodeficiency Virus
535 (HIV). Transfus Med Hemother. 2016;43(3):203-22. Epub 2016/07/13. doi:
536 10.1159/000445852. PubMed PMID: 27403093; PubMed Central PMCID:
537 PMCPMC4924471.
- 538 7. Ayinde D, Maudet C, Transy C, Margottin-Goguet F. Limelight on two HIV/SIV
539 accessory proteins in macrophage infection: is Vpx overshadowing Vpr?
540 Retrovirology. 2010;7:35. Epub 2010/04/13. doi: 10.1186/1742-4690-7-35. PubMed
541 PMID: 20380700; PubMed Central PMCID: PMCPMC2867959.
- 542 8. Karn J, Stoltzfus CM. Transcriptional and posttranscriptional regulation of HIV-1

543 gene expression. Cold Spring Harb Perspect Med. 2012;2(2):a006916. Epub
544 2012/02/23. doi: 10.1101/cshperspect.a006916. PubMed PMID: 22355797; PubMed
545 Central PMCID: PMCPMC3281586.

546 9. Colomer-Lluch M, Ruiz A, Moris A, Prado JG. Restriction Factors: From
547 Intrinsic Viral Restriction to Shaping Cellular Immunity Against HIV-1. Front
548 Immunol. 2018;9:2876. Epub 2018/12/24. doi: 10.3389/fimmu.2018.02876. PubMed
549 PMID: 30574147; PubMed Central PMCID: PMCPMC6291751.

550 10. Strebel K. HIV accessory proteins versus host restriction factors. Curr Opin Virol.
551 2013;3(6):692-9. Epub 2013/11/20. doi: 10.1016/j.coviro.2013.08.004
552 S1879-6257(13)00140-5 [pii]. PubMed PMID: 24246762; PubMed Central PMCID:
553 PMC3855913.

554 11. Chande A, Cuccurullo EC, Rosa A, Ziglio S, Carpenter S, Pizzato M. S2 from
555 equine infectious anemia virus is an infectivity factor which counteracts the retroviral
556 inhibitors SERINC5 and SERINC3. P Natl Acad Sci USA. 2016;113(46):13197-202.
557 doi: 10.1073/pnas.1612044113. PubMed PMID: WOS:000388970100083.

558 12. Frech K, Brack-Werner R, Werner T. Common modular structure of lentivirus
559 LTRs. Virology. 1996;224(1):256-67. Epub 1996/10/01. doi: 10.1006/viro.1996.0527.
560 PubMed PMID: 8862420.

561 13. Taube R, Fujinaga K, Irwin D, Wimmer J, Geyer M, Peterlin BM. Interactions
562 between equine cyclin T1, Tat, and TAR are disrupted by a leucine-to-valine
563 substitution found in human cyclin T1. J Virol. 2000;74(2):892-8. Epub 2000/01/07.
564 PubMed PMID: 10623752; PubMed Central PMCID: PMC111610.

- 565 14. Pedersen FS. Retroviral Replication2001.
- 566 15. Beisel CE, Edwards JF, Dunn LL, Rice NR. Analysis of multiple mRNAs from
567 pathogenic equine infectious anemia virus (EIAV) in an acutely infected horse reveals
568 a novel protein, Ttm, derived from the carboxy terminus of the EIAV transmembrane
569 protein. J Virol. 1993;67(2):832-42. Epub 1993/02/01. PubMed PMID: 8419648;
570 PubMed Central PMCID: PMC237437.
- 571 16. Noiman S, Yaniv A, Sherman L, Tronick SR, Gazit A. Pattern of transcription of
572 the genome of equine infectious anemia virus. J Virol. 1990;64(4):1839-43. Epub
573 1990/04/01. PubMed PMID: 2157066; PubMed Central PMCID: PMC249326.
- 574 17. Leblanc J, Weil J, Beemon K. Posttranscriptional regulation of retroviral gene
575 expression: primary RNA transcripts play three roles as pre-mRNA, mRNA, and
576 genomic RNA. Wiley Interdiscip Rev RNA. 2013;4(5):567-80. Epub 2013/06/12. doi:
577 10.1002/wrna.1179. PubMed PMID: 23754689; PubMed Central PMCID:
578 PMCPMC3810403.
- 579 18. Booth DS, Cheng Y, Frankel AD. The export receptor CRM1 forms a dimer to
580 promote nuclear export of HIV RNA. Elife. 2014;3:e04121. Epub 2014/12/09. doi:
581 10.7554/eLife.04121. PubMed PMID: 25486595; PubMed Central PMCID:
582 PMCPMC4360530.
- 583 19. Otero GC, Harris ME, Donello JE, Hope TJ. Leptomycin B inhibits equine
584 infectious anemia virus Rev and feline immunodeficiency virus rev function but not
585 the function of the hepatitis B virus posttranscriptional regulatory element. J Virol.
586 1998;72(9):7593-7. Epub 1998/08/08. PubMed PMID: 9696859; PubMed Central

587 PMCID: PMC110012.

588 20. Rausch JW, Le Grice SF. HIV Rev Assembly on the Rev Response Element
589 (RRE): A Structural Perspective. *Viruses*. 2015;7(6):3053-75. Epub 2015/06/16. doi:
590 10.3390/v7062760. PubMed PMID: 26075509; PubMed Central PMCID:
591 PMCPMC4488727.

592 21. Cochrane A. HIV-1 Rev function and RNA nuclear-cytoplasmic export. *Methods*
593 *Mol Biol*. 2014;1087:103-14. Epub 2013/10/26. doi: 10.1007/978-1-62703-670-2_9.
594 PubMed PMID: 24158817.

595 22. Lee JH, Culver G, Carpenter S, Dobbs D. Analysis of the EIAV Rev-responsive
596 element (RRE) reveals a conserved RNA motif required for high affinity Rev binding
597 in both HIV-1 and EIAV. *PLoS One*. 2008;3(6):e2272. Epub 2008/06/05. doi:
598 10.1371/journal.pone.0002272. PubMed PMID: 18523581; PubMed Central PMCID:
599 PMCPMC2386976.

600 23. Jayaraman B, Crosby DC, Homer C, Ribeiro I, Mavor D, Frankel AD. RNA-
601 directed remodeling of the HIV-1 protein Rev orchestrates assembly of the Rev-Rev
602 response element complex. *Elife*. 2014;3:e04120. Epub 2014/12/09. doi:
603 10.7554/eLife.04120. PubMed PMID: 25486594; PubMed Central PMCID:
604 PMCPMC4360532.

605 24. Carpenter S, Dobbs D. Molecular and biological characterization of equine
606 infectious anemia virus Rev. *Curr HIV Res*. 2010;8(1):87-93. Epub 2010/03/10. doi:
607 10.2174/157016210790416424. PubMed PMID: 20210784.

608 25. Harris ME, Gontarek RR, Derse D, Hope TJ. Differential requirements for

609 alternative splicing and nuclear export functions of equine infectious anemia virus
610 Rev protein. Mol Cell Biol. 1998;18(7):3889-99. Epub 1998/06/25. PubMed PMID:
611 9632773; PubMed Central PMCID: PMC108973.

612 26. Umunnakwe CN, Loyd H, Cornick K, Chavez JR, Dobbs D, Carpenter S.
613 Computational modeling suggests dimerization of equine infectious anemia virus Rev
614 is required for RNA binding. Retrovirology. 2014;11:115. Epub 2014/12/24. doi:
615 10.1186/s12977-014-0115-7
616 s12977-014-0115-7 [pii]. PubMed PMID: 25533001; PubMed Central PMCID:
617 PMC4299382.

618 27. Ihm Y, Sparks WO, Lee JH, Cao H, Carpenter S, Wang CZ, et al. Structural
619 model of the Rev regulatory protein from equine infectious anemia virus. PLoS ONE.
620 2009;4(1):e4178. Epub 2009/01/13. doi: 10.1371/journal.pone.0004178. PubMed
621 PMID: 19137065; PubMed Central PMCID: PMC2613556.

622 28. Wang XF, Wang YH, Bai B, Zhang M, Chen J, Zhang X, et al. Truncation of the
623 Cytoplasmic Tail of Equine Infectious Anemia Virus Increases Virion Production by
624 Improving Env Cleavage and Plasma Membrane Localization. J Virol.
625 2021;95(23):e0108721. Epub 2021/09/09. doi: 10.1128/JVI.01087-21. PubMed PMID:
626 34495693; PubMed Central PMCID: PMCPMC8577380.

627 29. Wang XF, Lin YZ, Li Q, Liu Q, Zhao WW, Du C, et al. Genetic Evolution during
628 the development of an attenuated EIAV vaccine. Retrovirology. 2016;13:9. Epub
629 2016/02/05. doi: 10.1186/s12977-016-0240-6. PubMed PMID: 26842878; PubMed
630 Central PMCID: PMCPMC4738788.

- 631 30. Lee JH, Murphy SC, Belshan M, Sparks WO, Wannemuehler Y, Liu S, et al.
632 Characterization of functional domains of equine infectious anemia virus Rev
633 suggests a bipartite RNA-binding domain. J Virol. 2006;80(8):3844-52. Epub
634 2006/03/31. doi: 10.1128/JVI.80.8.3844-3852.2006. PubMed PMID: 16571801;
635 PubMed Central PMCID: PMCPMC1440447.
- 636 31. Ren H, Yin X, Su C, Guo M, Wang XF, Na L, et al. Equine lentivirus counteracts
637 SAMHD1 restriction by Rev-mediated degradation of SAMHD1 via the BECN1-
638 dependent lysosomal pathway. Autophagy. 2020:1-18. Epub 2020/11/12. doi:
639 10.1080/15548627.2020.1846301. PubMed PMID: 33172327.
- 640 32. Belshan M, Park GS, Bilodeau P, Stoltzfus CM, Carpenter S. Binding of equine
641 infectious anemia virus rev to an exon splicing enhancer mediates alternative splicing
642 and nuclear export of viral mRNAs. Mol Cell Biol. 2000;20(10):3550-7. Epub
643 2000/04/25. doi: 10.1128/MCB.20.10.3550-3557.2000. PubMed PMID: 10779344;
644 PubMed Central PMCID: PMCPMC85647.
- 645 33. Ostermann PN, Ritchie A, Ptok J, Schaal H. Let It Go: HIV-1 cis-Acting
646 Repressive Sequences. J Virol. 2021;95(15):e0034221. Epub 2021/05/14. doi:
647 10.1128/JVI.00342-21. PubMed PMID: 33980600; PubMed Central PMCID:
648 PMCPMC8274598.
- 649 34. Schwartz S, Felber BK, Pavlakis GN. Distinct RNA sequences in the gag region
650 of human immunodeficiency virus type 1 decrease RNA stability and inhibit
651 expression in the absence of Rev protein. J Virol. 1992;66(1):150-9. Epub 1992/01/01.
652 doi: 10.1128/JVI.66.1.150-159.1992. PubMed PMID: 1727477; PubMed Central

- 653 PMCID: PMCPMC238270.
- 654 35. Schwartz S, Campbell M, Nasioulas G, Harrison J, Felber BK, Pavlakis GN.
655 Mutational inactivation of an inhibitory sequence in human immunodeficiency virus
656 type 1 results in Rev-independent gag expression. J Virol. 1992;66(12):7176-82. Epub
657 1992/12/01. doi: 10.1128/JVI.66.12.7176-7182.1992. PubMed PMID: 1433510;
658 PubMed Central PMCID: PMCPMC240411.
- 659 36. Schneider R, Campbell M, Nasioulas G, Felber BK, Pavlakis GN. Inactivation of
660 the human immunodeficiency virus type 1 inhibitory elements allows Rev-
661 independent expression of Gag and Gag/protease and particle formation. J Virol.
662 1997;71(7):4892-903. Epub 1997/07/01. doi: 10.1128/JVI.71.7.4892-4903.1997.
663 PubMed PMID: 9188551; PubMed Central PMCID: PMCPMC191719.
- 664 37. Rosin-Arbesfeld R, Yaniv A, Gazit A. Suboptimal splice sites of equine infectious
665 anaemia virus control Rev responsiveness. J Gen Virol. 2000;81(Pt 5):1265-72. Epub
666 2000/04/18. doi: 10.1099/0022-1317-81-5-1265. PubMed PMID: 10769069.
- 667 38. Roy D, Bhanja Chowdhury J, Ghosh S. Polypyrimidine tract binding protein
668 (PTB) associates with intronic and exonic domains to squelch nuclear export of
669 unspliced RNA. FEBS Lett. 2013;587(23):3802-7. Epub 2013/10/23. doi:
670 10.1016/j.febslet.2013.10.005. PubMed PMID: 24145297.
- 671 39. Tan W, Felber BK, Zolotukhin AS, Pavlakis GN, Schwartz S. Efficient expression
672 of the human papillomavirus type 16 L1 protein in epithelial cells by using Rev and
673 the Rev-responsive element of human immunodeficiency virus or the cis-acting
674 transactivation element of simian retrovirus type 1. J Virol. 1995;69(9):5607-20. Epub

675 1995/09/01. doi: 10.1128/JVI.69.9.5607-5620.1995. PubMed PMID: 7637007;
676 PubMed Central PMCID: PMCPMC189417.

677 40. Suh D, Seguin B, Atkinson S, Ozdamar B, Staffa A, Emili A, et al. Mapping of
678 determinants required for the function of the HIV-1 env nuclear retention sequence.
679 Virology. 2003;310(1):85-99. Epub 2003/06/06. doi: 10.1016/s0042-6822(03)00073-4.
680 PubMed PMID: 12788633.

681 41. Levesque K, Halvorsen M, Abrahamyan L, Chatel-Chaix L, Poupon V, Gordon H,
682 et al. Trafficking of HIV-1 RNA is mediated by heterogeneous nuclear
683 ribonucleoprotein A2 expression and impacts on viral assembly. Traffic.
684 2006;7(9):1177-93. Epub 2006/09/28. doi: 10.1111/j.1600-0854.2006.00461.x.
685 PubMed PMID: 17004321.

686 42. Xiao H, Wyler E, Milek M, Grewe B, Kirchner P, Ekici A, et al. CRNKL1 Is a
687 Highly Selective Regulator of Intron-Retaining HIV-1 and Cellular mRNAs. mBio.
688 2021;12(1). Epub 2021/01/21. doi: 10.1128/mBio.02525-20. PubMed PMID:
689 33468685; PubMed Central PMCID: PMCPMC7845644.

690 43. Brighty DW, Rosenberg M. A cis-acting repressive sequence that overlaps the
691 Rev-responsive element of human immunodeficiency virus type 1 regulates nuclear
692 retention of env mRNAs independently of known splice signals. Proc Natl Acad Sci U
693 S A. 1994;91(18):8314-8. Epub 1994/08/30. doi: 10.1073/pnas.91.18.8314. PubMed
694 PMID: 8078879; PubMed Central PMCID: PMCPMC44596.

695 44. Gallego J, Groatorex J, Zhang H, Yang B, Arunachalam S, Fang J, et al. Rev
696 binds specifically to a purine loop in the SL1 region of the HIV-1 leader RNA. J Biol

- 697 Chem. 2003;278(41):40385-91. Epub 2003/07/10. doi: 10.1074/jbc.M301041200.
- 698 PubMed PMID: 12851400.
- 699 45. Greatorex JS, Palmer EA, Pomerantz RJ, Dangerfield JA, Lever AML. Mutation
700 of the Rev-binding loop in the human immunodeficiency virus 1 leader causes a
701 replication defect characterized by altered RNA trafficking and packaging. J Gen
702 Virol. 2006;87(Pt 10):3039-44. Epub 2006/09/12. doi: 10.1099/vir.0.81658-0.
703 PubMed PMID: 16963763.
- 704 46. Akhlaq S, Panicker NG, Philip PS, Ali LM, Dudley JP, Rizvi TA, et al. A cis-
705 Acting Element Downstream of the Mouse Mammary Tumor Virus Major Splice
706 Donor Critical for RNA Elongation and Stability. J Mol Biol. 2018;430(21):4307-24.
707 Epub 2018/09/05. doi: 10.1016/j.jmb.2018.08.025. PubMed PMID: 30179605;
708 PubMed Central PMCID: PMC6193448.
- 709 47. Suhasini M, Reddy TR. Cellular proteins and HIV-1 Rev function. Curr HIV Res.
710 2009;7(1):91-100. Epub 2009/01/20. doi: 10.2174/157016209787048474. PubMed
711 PMID: 19149558.
- 712 48. Wang XF, Bai BW, Lin YZ, Qi T, Du C, Song MX, et al. High-Efficiency Rescue
713 of Equine Infectious Anemia Virus from a CMV-Driven Infectious Clone. Virol Sin.
714 2019;34(6):725-8. PubMed PMID: ISI:000500088400013.
- 715 49. Lin YZ, Cao XZ, Li L, Jiang CG, Wang XF, Ma J, et al. The pathogenic and
716 vaccine strains of equine infectious anemia virus differentially induce cytokine and
717 chemokine expression and apoptosis in macrophages. Virus Res. 2011;160(1-2):274-
718 82. Epub 2011/07/26. doi: 10.1016/j.virusres.2011.06.028. PubMed PMID: 21782860.

719 50. Jiang CG, Gao X, Ma J, Lin YZ, Wang XF, Zhao LP, et al. C-terminal truncation
720 of the transmembrane protein of an attenuated lentiviral vaccine alters its in vitro but
721 not in vivo replication and weakens its potential pathogenicity. *Virus Res.*
722 2011;158(1-2):235-45. Epub 2011/05/05. doi: 10.1016/j.virusres.2011.04.007.
723 PubMed PMID: 21539871.

724 51. Lin Y, Wang XF, Wang Y, Du C, Ren H, Liu C, et al. Env diversity-dependent
725 protection of the attenuated equine infectious anaemia virus vaccine. *Emerg Microbes*
726 *Infect.* 2020;9(1):1309-20. Epub 2020/06/12. doi: 10.1080/22221751.2020.1773323.
727 PubMed PMID: 32525460; PubMed Central PMCID: PMC7473056.

728 52. Bai B, Wang XF, Zhang M, Na L, Zhang X, Zhang H, et al. The N-glycosylation
729 of Equine Tetherin Affects Antiviral Activity by Regulating Its Subcellular
730 Localization. *Viruses.* 2020;12(2). Epub 2020/02/23. doi: 10.3390/v12020220.
731 PubMed PMID: 32079099; PubMed Central PMCID: PMC7077275.

732 53. Hu Z, Chang H, Ge M, Lin Y, Wang X, Guo W, et al. Development of antigen
733 capture ELISA for the quantification of EIAV p26 protein. *Appl Microbiol Biotechnol.*
734 2014;98(21):9073-81. Epub 2014/09/27. doi: 10.1007/s00253-014-6078-8. PubMed
735 PMID: 25256618.

736

Figure legends

Fig 1. Discovery and identification of a novel protein-encoding gene from EIAV.

(A) A schematic representation of the EIAV genome is displayed at the top of the figure. Known ORFs are shown as boxes with different colors. The position of RRE is depicted with a heavy line. The splicing pattern of mat transcripts is shown below the genomic structure. Rectangles represent exons of genes. Vertical arrows represent splice site of mRNA. Horizontal arrows show the locations of oligonucleotide primers for amplification of cDNAs. (B) Identification of the mat gene from multiple tissues (lymph gland, testis, liver, kidney, brain, spleen, marrow and heart) of horses infected with EIAV_{LN40}. PC, mat sequence inserted into pMD18T-vector as a positive control. NC, negative control. (C) Identification of the mat gene from eMDM cells infected with EIAV_{DLV36} and EIAV_{DLV121}. (D) Identification of Mat from eMDM cells infected with EIAV_{DLV36} and EIAV_{DLV121}. The Mat-specific antiserum (described in Materials and Methods) was used for the immunodetection of Mat protein. An anti-EIAV p26 antiserum was used to monitor the effect of infection. Hela cells transfected with empty plasmids were used as a negative control (lane 1). Hela cells transfected with a codon-optimized Mat expression plasmid were used as a positive control (lane 2). eMDM cells isolated from healthy horses were used as a specific control (lane 5).

Fig 2. The expression of Mat depends on Rev. A Mat constructs was transiently cotransfected into HEK293T cells (A) or FDD cells (B) with Rev expression plasmids or empty plasmids. Mat expression was analyzed with WB using a mouse anti-HA antibody 48h post transfection (hpt); β -actin expression was detected using an actin-specific antiserum to monitor unity of blotting and equal loading of the gel; GFP expression was quantified using a GFP-specific antiserum to monitor transfection efficiencies. (C) An Mat constructs was transfected into HEK293T cells, along with

increasing amounts of a Rev expression plasmid. (D) LMB inhibited Rev-independent expression of Mat. HEK293T cells were cotransfected with Mat constructs and Rev expression plasmids. At 12 hpt, cells were treated with leptomycin B (0nM [-] or 50nM [+], as indicated). At 24 h post drug treatment, cells were harvested and subjected to WB. (E) No Mat protein expression was detected when cotransfected with Mat constructs and Rev_{L95D} expression plasmids. Rev_{L95D} is a mutant of Rev without the ability to mediating mRNA export to the cytoplasm. VR1012 was used as an empty vector control. (F) Interaction between mat mRNA and Rev protein was determined using an RNA-pull down assay. Cell lysates from HEK293T transfected with Rev expression plasmids were incubated with biotin-labeled or unlabeled RNA transcribed *in vitro*. The biotin-labeled mat mRNA, EIAV RRE or HIV-1 RRE were pulled down with streptavidin (STA)-conjugated magnetic beads, followed by WB of co-purified proteins using anti-Flag antibody. Levels of pulled-down Rev are shown in a bar chart. (G) Distribution of cytoplasmic and nuclear mat or GFP RNAs in HEK293T cells at 36 hpt was analyzed using qPCR with primers specific to mat or GFP mRNA. Ratios of cytoplasmic to nuclear mRNAs were then calculated and are presented. LaminB1 and α -tubulin were detected using WB to assess the efficacy of cell fractionation. All of the experiments were performed three times, and a representative result is shown.

Fig 3. Mapping of the determinants of Rev-mediated expression of Mat. (A)

Schematic diagram of the mat ORF coding sequence with 447 nucleotides, which contains two exons (exon-1:1-393nt; exon-2:394-447nt). The exon-1 sequence was inserted into pcDNA-HA to express the Mat-exon1 protein. (B) HEK293T cells were cotransfected Mat or mat-exon1 expression constructs with Rev or empty plasmids. Mat expression was analyzed with WB using a mouse anti-HA antibody; GFP, derived

from pEGFP-N1, was quantified using a GFP-specific antiserum to monitor transfection efficiencies. (C) HEK293T cells were cotransfected each of six different Mat-exon1 plasmids, which were derived from six EIAV strains as indicated, with Rev or empty plasmids. Mat-exon1 proteins were detected 48 hpt with WB. (D) An RNA-pull down assay was performed to verify the interaction of mat-exon1 mRNA with Rev protein. The Rev pulled down by mat mRNA acted as a positive control. Biotin-unlabeled mRNAs were used to indicate the specificity of this assay. (E) Distribution of cytoplasmic and nuclear mat-exon1 mRNAs in HEK293T cells at 36 pht was analyzed using qPCR. The subcellar distribution of mat mRNAs was analyzed as a control.

Fig 4. The key domains of Rev mediating the expression of Gag and Mat are

different. (A) Schematic diagram of wild-type EIAV Rev, which contains four functional domains: NES (nuclear export domain; aa 31-55); RBD (RNA-binding domain; aa76-88 or aa159-163); ND (Non-essential domain; aa122-148); NLS (nuclear localization domain; aa159-163). The structures of Rev five mutants are: Rev_{mNES}, the mutation of Rev from L36, W41, L45, L49 to G; Rev_{mRBD}, the mutation of Rev from RRDSW to AADAA; Rev_{mNLS}, the mutation of Rev from KRRRK to KAAAK; Rev_{ΔND}, the 27 amino acid (aa122-148) deletion of ND domain and Rev_{L95D}, described in the materials and methods. WB analysis of Mat (B) or Gag (D) expression levels when a Mat expression plasmid or a Rev-dependent EIAV Gag/Pol expression construct (pGP-RRE) was transfected into HEK293T cells with either an empty vector VR1012 (as a negative control) or a series of Rev expression plasmids, as indicated. Distribution of cytoplasmic and nuclear mat (C) or gag (E) mRNAs in HEK293T cells at 36 pht was analyzed using qPCR when Mat or pGP-RRE expression plasmids were transfected into HEK293T cells with a series of Rev

812 expression plasmids. All of the experiments were performed three times, and a
813 representative result is shown.

814 **Fig 5. Functional identification of Mat first exon used as an RNA export element.**

815 (A) Schematic diagram of a series of plasmids used for the identification of the
816 function of the RNA export element. pGP: contains only EIAV gag/plo-coding
817 sequences; pGP-RRE: with EIAV RRE inserted downstream of the gag/pol coding
818 sequence; pGP-MA: with one copy of mat-exon1(MA) inserted downstream of the
819 gag/pol-coding sequence; pGP-3×MA with three copies of mat-exon1(MA) inserted
820 downstream of the gag/pol-coding sequence. (B) WB analysis of indicated plasmids
821 cotransfected with Rev or empty vector. All of the experiments were performed
822 three times, and a representative result is shown.

823 **Fig 6. The EIAV proviral genome structure and transcriptional patterns.** The
824 previously identified Rev-response element (RRE) was named RRE-1 (blue line), and
825 is located in the env-coding region. The unspliced full-length mRNA transcripts
826 encoding Gag and Pol proteins and an incompletely spliced transcript encoding Env
827 and S2 proteins are exported to the cytoplasm in a pathway mediated by the binding
828 of Rev to RRE-1. The transcripts encoding Tat or Rev protein are exported to the
829 cytoplasm via an endogenous cellular pathway used by cellular mRNAs independent
830 of Rev. The transcripts encoding Mat protein are exported to the cytoplasm via a
831 pathway mediated by the binding of Rev to the first exon of Mat mRNA. The first
832 exon of mat therefore has the function of RRE, and is named here RRE-2 (red line).

Figure 2

Figure 4

A

EIAV proviral DNA

mRNAs Transcripts

B

D

C

Figure 1

A **B**

Figure 5

Figure 6

Figure 3