Insights into the nuclear export of murine leukemia virus intron-containing RNA

Lucie Pessel-Vivares, Laurent Houzet, Sébastien Lainé, Marylène Mougel

Insights into the nuclear export of murine leukemia virus intron-containing RNA

Lucie Pessel-Vivares[1,2], Laurent Houzet[3], Sébastien Lainé[1], and Marylène Mougel[1]*

[1] CNRS, UM, CPBS, 1919 Route de Mende, Montpellier, France
[2] Department of Infectious Diseases, King's College London, London SE1 9RT, United Kingdom
[3] Inserm U1085-IRSET, Université de Rennes 1, Structure Fédérative Recherche Biosit, Campus de Beaulieu, Rennes, France
*corresponding authors

Key words: RNA export; nucleocytoplasmic trafficking; intron-containing RNA**;** intron retention; *cis*-acting RNA element; NXF1/Tap; MLV; gammaretrovirus; post-transcriptional gene regulation; virus replication.

Acknowledgements: This work was supported by institutional grants from the CNRS and the University of Montpellier. LPV was supported by fellowships from La Ligue contre le Cancer and the Ministère de l'Enseignement Supérieur et de la Recherche. Many thanks to C. Swanson and M. Ferrer for critical reading of the manuscript.

Running title: MLV nuclear export

ABSTRACT

The retroviral genome consists of an intron-containing transcript that has essential cytoplasmic functions in the infected cell. This viral transcript can escape splicing, circumvent the nuclear checkpoint mechanisms and be transported to the cytoplasm by hijacking the host machinery. Once in the cytoplasm, viral unspliced RNA acts as mRNA to be translated and as genomic RNA to be packaged into nascent viruses. The murine leukemia virus (MLV) is among the first retroviruses discovered and is classified as simple *Retroviridae* due to its minimal encoding capacity. The oncogenic and transduction abilities of MLV are extensively studied, whereas surprisingly the crucial step of its nuclear export has remained unsolved until 2014. Recent work has revealed the recruitment by MLV of the cellular NXF1/Tap-dependent pathway for export. Unconventionally, MLV uses of Tap to export both spliced and unspliced viral RNAs. Unlike other retroviruses, MLV does not harbor a unique RNA signal for export. Indeed, multiple sequences throughout the MLV genome appear to promote export of the unspliced MLV RNA. We review here the current understanding of the export mechanism and highlight the determinants that influence MLV export. As the molecular mechanism of MLV export is elucidated, we will gain insight into the contribution of the export pathway to the cytoplasmic fate of the viral RNA.

Introduction

Murine leukemia viruses (MLV) were among the first mammalian retroviruses discovered more than 60 years ago and are today among the best characterized retroviruses. The study of retroviruses has a broad impact on diverse domains of biology, medicine and biotechnologies. Since most of the retroviruses are oncogenic, study of their pathogenic effects at the cellular and molecular levels has led to a better understanding of fundamental mechanisms involved in

cellular growth control and carcinogenesis.[1, 2] MLVs induce leukemia in mice and provide a useful model to investigate the genetic basis of hematopoietic cancer contributing to a better understanding of hematopoiesis. Like many retroviruses, MLVs induce tumors by proviral insertional mutagenesis, corresponding to the integration of viral DNA near cellular proto-oncogenes and subsequent alteration of their expression. Analysis of integration sites led to the identification of many genes that regulate cell growth and differentiation and are involved in tumor formation.[3-6] Further understanding of the retrovirus life-cycle has brought new concepts in molecular biology and provided new powerful tools for biotechnology, cancer research and gene therapy. The discovery of the reverse transcription, the copying of RNA to DNA, in retrovirus replication cycle had a strong impact on biotechnology. Indeed, the use of MLV reverse transcriptase in the laboratory to generate DNA copies from RNA provided a remarkable tool for analysis of RNA molecules. Finally, in-depth knowledge of retrovirus molecular structure and full replication cycle allowed to use them as backbone for vectors to generate efficient and specific gene delivery systems. Vectors derived from MLV are widely used for cell transduction in laboratory [7] and their use in gene therapy trials has been successful in most of the patients. However, adverse effects related to insertional mutagenesis were observed in 10% of the patients and work is in progress to improve their biosafety.[8-10]

Considered as a prototypic retrovirus, MLV has been extensively studied and has a major impact on understanding the molecular organization and life-cycle of retroviruses. The genome of all retroviruses is a capped and polyadenylated RNA about 8-kb long and flanked by two long terminal repeats (LTR) containing regulatory elements for transcription and polyadenylation. Like all simple retroviruses, MLV genomic RNA (gRNA), also called full-length RNA (FL RNA), contains only the three major coding domains common to all retroviruses: *gag*, *pol* and *env* encoding for structural, enzymatic and receptor binding viral proteins, respectively (Fig. 1), whereas complex retroviruses, such as the human immunodeficiency virus (HIV), contain additional coding domains corresponding to accessory and regulatory proteins. In addition to coding capacity, *cis*-acting elements present on the FL RNA are directly involved in many steps of the virus replication cycle, such as splicing, packaging or reverse transcription, through their interaction with cellular or viral *trans*-acting factors. Finally, only a subset of the FL RNA will be used to produce viral spliced RNAs. Simple retroviruses are characterized by only one spliced RNA resulting from the removing of a single intron and generating the *env* mRNA, whereas complex retroviruses produce many other spliced RNA coding for accessory proteins, with at least 46 spliced RNA for HIV.[11] However, although MLV belongs to simple retroviruses, a second MLV singly spliced mRNA called SD' RNA and generated from an alternative splice donor site was discovered. Unlike *env* mRNA, SD' RNA can be packaged, reverse transcribed and integrated in the host genome.[12]

In contrast to cellular RNA, the replication cycle of a retrovirus requires unspliced intron-containing viral RNA to be exported from the nucleus. Once in the cytoplasm, this FL RNA will be either translated to provide the viral Gag and Gag-Pol polyproteins, or packaged as a genome into assembling virions. To this dual fate correspond two distinct pools of unspliced viral RNA in the cytoplasm of MLV infected cells, while a unique pool of a bifunctional RNA is used for translation and packaging in the case of HIV.[13] As opposed to retroviral FL RNA, the presence of unspliced cellular RNA in the cytoplasm is rare since splicing is a prerequisite for nuclear export of mRNA and intron retention targets RNA to degradation by cellular quality control mechanisms. Thus, retrovirus FL RNA needs to escape from degradation and then find a strategy to reach the cytoplasm. For example, HIV Rev regulatory protein binds the Rev response element (RRE) present in FL RNA and interacts with the nuclear export receptor CRM1 to exit the nucleus *via* the nuclear pore complex (NPC). The simple Mason Pfizer Monkey virus *(*MPMV*)* contains a *cis*-acting RNA motif, the constitutive transport element (CTE), which mediates the export of the

FL RNA through the direct binding of the nuclear export factor NXF1, also called Tap. Studies of retroviral mechanisms to export their unspliced gRNA have also led to a better understanding of their host nuclear export pathways. Surprisingly, whereas MLV molecular organization and life-cycle have been extensively studied, the issue of the export of its unspliced FL RNA was resolved only very recently by several teams simultaneously, closing a major gap in the knowledge of MLV replication cycle. Here we will focus on the main similarities and differences between the *cis*-acting elements and *trans*-acting factors reported as contributing to MLV export and we will discuss the apparent complexity of MLV export pathway.

Overview of mammalian mRNA export

The nucleocytoplasmic transport of mRNA is closely related to its nuclear maturation (capping, splicing, cleavage and polyadenylation) since a mRNA maturation defect prevents its export and leads to RNA degradation.[14] Interestingly, mRNA export is saturable, indicating the involvement of specific limiting factors. Export of the majority of the mature mRNA is dependent on the highly conserved Tap factor. Tap shuttles between nucleus and cytoplasm by interacting with the nucleoporins (NUP) in the nuclear pore *via* its C-terminus while its N-terminus binds the RNA cargo. Other factors are recruited to reinforce these weak interactions. The heterodimerization of Tap with the NTF2-related export protein-1 (NXT1), also called p15, induces a conformational switch that strengthens Tap-NUPs interaction within the nuclear pore. In parallel, the conserved transcription-export (TREX) complex, mainly including the THO complex, UAP56, Aly and CIP29, is recruited to the newly synthesized transcripts during splicing. This RNA/TREX complex finally uses the Tap/p15 heterodimer as an export factor to pass through the nuclear pores.[15] RNA helicases, such as Dbp5, help mRNAs to exit the pores on the cytoplasmic side (Fig.2A).[16] Recent microscopy approaches focusing on single mRNA molecules in living cells have provided valuable insights into the mRNA nucleocytoplasmic translocation. The kinetic of mRNA export was spatially resolved and showed that mRNA molecules pass rapidly through the central channel of nuclear pores, and that docking on the nuclear side and then waiting for release into the cytoplasm are the most time consuming steps (totaling 180 +/- 10 ms). Some mRNA molecules (10%) sit for seconds at nuclear pores without gaining exit, suggesting that RNA could be screened for quality at this point (for review see ref. [17]).

In contrast to splicing-dependent mRNA export, little is known about how natural mRNAs with retained introns are exported in mammalian cells. A substantial majority of genes contains introns and over 90% of human genes are alternatively spliced. Expression of genes with persistence of introns has been reported in several cancers [18] and could participate in adaptive regulation.[19] In addition, natural intronless cellular mRNAs such as histone H2A, HSPB3, IFNα1 and IFNβ1 constitute model systems. Their export is blocked by knockdown of TREX components (UAP56 and THOC2) or Tap (for review see ref. [20]) but the mechanisms used to bypass nuclear export restrictions remain unclear. However, most of the studies done with intronless viral mRNAs (i.e. herpes simplex virus, influenza A virus), reveal that these mRNAs bind to specific cellular or viral proteins that recruit the classic mRNA export machinery (for review see ref. [21]). The unspliced RNA genome of retroviruses also provides valuable information on intron-containing mRNA export. Especially, studies of HIV and MPMV led to the understanding of the function of cellular export factors, CRM1 and Tap, respectively. In addition to the role of the Tap/p15 complex in the export of bulk transcripts, it is now well established that Tap/p15 promotes export of intron-containing mRNA through binding to the well-conserved CTE motif present in both viral and cellular mRNAs.[22]

MLV requires Tap/NXF1-dependent pathway to export its viral RNA

Little is known about viral RNA localization in cells infected with replication-competent MLV. Localization of unspliced viral RNA in mouse NIH3T3 cells chronically infected with Mo-MLV was determined by fluorescence *in situ* hybridization (FISH) and by cell fractionation experiments. These results showed that the majority of the unspliced RNA remained within the nucleus, suggesting that the nucleus constitutes an MLV RNA reservoir.[23] Unspliced RNAs are divided into two distinct RNA pools according to its dual fate in the cytoplasm: it can be targeted to the ribosomes to be translated as mRNA or be driven at the plasma membrane (PM) to be packaged as gRNA in the newly formed viruses. These two RNA populations (mRNA, gRNA) can be distinguished by Actinomycin D treatment and measurement of their different half-lives.[13, 24, 25] Indeed, the virion precursor gRNA has a half-life of 3-4 hours although the unspliced mRNA is stable for a period of 12-24 hours,[25] suggesting that the two pools of RNA could form different ribonucleoprotein complex (RNP) according to their fate. Whether the RNP are already formed in the nucleus and predetermine the cytoplasmic fate of the RNA remains to be demonstrated. The export pathway used by MLV RNA was identified very recently, in 2014.[26-30] Different complementary approaches such as RNA silencing, dominant-negative mutant of Tap and RNA competitor harboring four copies of the CTE sequence have been used either in human or murine cells lines, infected or transfected with MLV. All results concur that the Tap pathway is strictly required for MLV expression (Fig. 2). Indeed, over-expression of Tap protein increased the expression of MLV chimeric mRNA [26] and blocking the Tap pathway induced a drastic decrease in Gag and Env protein levels.[27, 29] The effect of Tap pathway inhibition on RNA export was very difficult to monitor due to the rapid degradation of the RNA retained in the nucleus.[27, 29] Thus, detection of nuclear retention of blocked unspliced RNA was made only possible when using a very sensitive RNA imaging method with single molecule sensitivity. This microscopy analysis allowed to image unspliced RNA from 8 hours after Tap-inhibition. [27] Furthermore, the interaction between Tap and the MLV RNAs was observed by co-immunoprecipitation experiments. Spliced and unspliced RNA were complexed with the cellular Tap protein in infected human cells [29] or transfected murine cell lines.[27]

In conclusion, in contrast to HIV-1, spliced and unspliced MLV RNA take the same route to reach the cytoplasm in order to be translated. It is not known how MLV discriminates between spliced and unspliced viral RNAs. Also, whether the same pathway is used to export the two different pools of unspliced viral RNA (to be translated or packaged) remains to be determined.

Tap-cofactors involved in MLV RNA export

The export of unspliced MLV RNA mediated by Tap seems to require several co-factors (Fig. 2). Sakuma *et al.* highlighted the requirement of the three factors UAP56, THOC5 and THOC7.[28] THOC5 and THOC7 are two proteins from the THO complex recruited onto the mRNA during splicing and able to recruit UAP56.[31] UAP56 is an RNA helicase of the DEAD-Box family that plays a critical role in nuclear export *via* its ability to recruit the Aly/Ref protein onto the mRNA, which itself is needed to recruit the Tap-p15 heterodimer.[32, 33] Interaction of Aly/Ref and THOC5 with Tap induces a conformational change of Tap resulting in a better interaction with the mRNA.[34] Interestingly, the study by Sakuma *et al.* revealed that although UAP56, THOC5 and THOC7 are required for the export of MLV unspliced RNA, Aly/Ref is dispensable (Fig. 2). The involvement of other proteins of the THO complex (THOC1, THOC2, THOC6) as well as the Cip29 protein has been assessed and the results showed that these proteins are not involved in the export of MLV RNA. Interestingly, the involvement of the co-factors are different for the MLV spliced and unspliced RNA. Indeed, although the UAP56 is necessary for the export of both spliced and unspliced RNA, THOC5 and THOC7 are only needed for the export of unspliced RNA.[28] Thus,

although the MLV spliced and unspliced RNA take the same route to be exported from the nucleus, the RNP formed are different suggesting that some partners could be involved in the balance between export and splicing.

It was also demonstrated that MLV unspliced RNA is sensitive to the presence of the cellular SRp20 protein. SRp20 shuttles between the nucleus and the cytoplasm and promotes nuclear export and translation of cellular mRNAs by recruiting Tap/p15 on mRNAs.[35-37] SRp20 was shown to serve as a bridge between Tap/p15 and the MLV unspliced RNA.[26] Moreover, SRp20 increases the binding of polysomes on the Gag mRNA in a Tap-dependent manner, thus linking nuclear export and translation.[26] These results are apparently not in agreement with the requirement of THOC5 and THOC7 but different binding sites could be present on the MLV unspliced RNA and could independently recruit Tap/p15 *via* some adaptors like SRp20 or THOC proteins (Fig. 2). Moreover, it is not excluded that SRp20 and THOC act in concert in the same export pathway.

MLV *cis*-acting RNA sequences

Retroviral unspliced RNA harbors particular *cis*-regulatory RNA elements that contribute to the recruitment of cellular nuclear export factors. For instance, the RRE present in the *env* sequence of HIV-1 is promoting the nuclear export of the FL RNA *via* the CRM1-dependent pathway. For the simple retrovirus MPMV, the FL RNA export is mediated *via* the binding of the CTE motif to Tap. Altogether, these studies reveal that retroviral *cis*-export elements usually localize at the 3' end of the viral FL RNA. However, the identification of the nuclear export signals of MLV appears more complex since the presence of multiple *cis*-export elements have been reported throughout the RNA, from its 5' to 3' regions (Table 1). Note that the species context used to study MLV should be taken into account since mechanisms of cellular pathways may differ from human to murine cells ([38-40] and personal data).

- *The R stem-loop*. The importance of the R sequence present in the 5' UTR of the MLV SL3 strain was first reported by Cupelli and colleagues. By using a construct containing a chloramphenicol acetyltransferase *(*CAT*)* reporter placed under the control of both SL3 U3 and the first 32 nt of the R region, they showed that the presence of R enhanced CAT activity in mouse cell lines whereas this CAT activity appeared more variable in human cell lines.[41, 42] The first 28 nt of R is predicted to form a stem loop, called RSL. Maintaining the secondary structure of the RSL appears to be important to enhance the CAT activity of the construct.[41] The RSL stem-loop is present in both unspliced and spliced RNA. Importantly, RSL is highly conserved among type-C retroviruses [41, 43] and can also form inside the R sequence of the Moloney strain (Mo-MLV) (Fig. 3A).[43] Indeed, MoMLV RSL can substitute the SL3 RSL activity.[41] Then, the role of this RSL was studied in the presence or the absence of an intronic and exonic sequence by using the same CAT reporter assays and revealed the importance of the R sequence in the expression of intron-containing reporter RNA.[41] The role of the RSL was later supported by experiments conducted with the full-length SL3 MLV. Deletion or mutation of the RSL in this viral context impaired the cytoplasmic accumulation of unspliced gRNA in murine and human cells lines.[44] Interestingly, all these studies showed that mutation or deletion of the RSL impacted Gag production, suggesting that the export driven by R targets the unspliced RNA to the translation machinery.

- *The Psi region*. The Psi region is mainly known for its role in RNA dimerization [45-47] and RNA packaging in new viral particles.[48-51] Psi is located at the 5' end of the genome, 350-nucleotide downstream of the 5' splice donor site (SD), and thus it is removed during the canonical splicing generating the non-packageable *env* RNA (Fig. 1). The Psi region forms an independent and highly structured domain including four stem-loops (called A, B, C and D) [52-54] (Fig. 3B). In 2005, two independent studies claimed a role for Psi in the nuclear export of FL RNA

(Fig. 3).[23, 55] The A-D stem-loops were deleted independently or in combination in Mo-MLV-based constructs and the RNA localization was monitored by subcellular fractionation followed by RT-qPCR as well as fluorescence microcopy (FISH). Both teams showed that ABCD hairpins promote the nuclear export of the viral transcripts in murine NIH3T3 cells and human HT-Fly cells.[23, 55] The role of Psi in FL RNA export was confirmed by competition experiments in which the over-expressed Psi sequence competed with the unspliced MLV FL transcript for the nuclear export in MLV-infected cells.[23] Altogether, these studies revealed a new function for the four Psi stem-loops in the transport of intron-containing viral RNAs from the nucleus to the cytoplasm.

- *The pol coding sequence*. Several studies have reported the contribution of coding sequences located in the *pol* gene. MLV or XMRV *gag* gene is poorly expressed under the control of the natural LTR or a CMV promoter whereas *gag-pol* gene is efficiently expressed,[26, 30] suggesting a putative role of the *pol* sequence in the expression of viral transcripts. Pilkington *et al*. worked with the complete genome of Mo-MLV or XMRV.[30] They deleted *env* or *pol* separately or *env* and *pol* together and observed in human cells that *pol* is required for Gag production. They defined a region of 1467nt in Mo-MLV (RT and IN) which was called post-transcriptional element (PTE). Insertion of PTE in the intron of poorly expressed reporter genes known to depend on post-transcriptional regulation (i.e. HIV Gag or CAT genes in respectively pNLgag and pDM138 plasmids), enhances the expression of the respective gene reporter. The PTE is a highly complex RNA structure with seven stem-loops (SL1 to SL7). A major role has been ascribed to SL1 and SL7 in the activity of the PTE, suggesting the presence of a bipartite signal for the PTE function (Fig. 3C). The fact that SL2 contains another *cis*-sequence identified by another group,[26] reinforces the importance of the PTE. Although SL2 is dispensable for the function of the XMRV PTE,[30] its role remains to be studied in the context of Mo-MLV.

Bartels and Luban identified two sequences in XMRV and MLV *pol* gene, which independently enhanced the expression of Gag (2232nt-3456nt and 4876nt-5199nt in XMRV) in human cells.[26] These sequences increased i) mRNA stability, ii) mRNA association to the polyribosomes and iii) mRNA nuclear export. From the first sequence, Luban's team predicted a hairpin of 100nt similar to the CTE *cis*-element present in MPMV, and named it γ-CTE (Fig. 3D). In particular, γ-CTE and CTE share a similar AAGACA loop that specifically recruits the cellular Tap factor to promote nuclear export of viral transcripts.[56] The study also shows that both MLV *gag* and *pol* sequences are also sensitive to the presence of SRp20 protein. This cellular factor links nuclear export and translation by increasing the binding of polysomes to Gag mRNA in a Tap-dependent manner.[26]

More recently, another *cis*-element of 70nt was identified in the *pol* gene and called CAE for cytoplasmic accumulation element (Fig. 3).[29] This CAE sequence is conserved among the gamma-retroviruses and is predicted to fold into 3 major stem-loops (Fig. 3E). Nine nucleotides (GGAAAGGAC) of XMRV CAE sequence are essential but not sufficient for the CAE function (Fig. 3E). Both MoMLV and XMRV CAE promote the nuclear export of viral transcript without affecting the stability, splicing or translation efficiency of the transcripts in human cell lines. Indeed, the expression of an HIV Gag construct became Rev independent when the CAE was inserted downstream the *gag* gene. In addition, the study revealed that MLV and XMRV CAE are both able to specifically interact with the nuclear export factor Tap triggering export of both spliced and unspliced viral RNAs.[29]

Finally, a sequence upstream the SA site in the *integrase* gene, formed by 237bp (5119-5355 in MLV), has been identified as a *cis*-element promoting the specific expression of the unspliced RNA in a position-independent manner.[57] This *cis*-element may play a role in FL RNA export, but this hypothesis was not assessed in the study and remains to be elucidated.

- *The U3 region.* A deletion within the U3 region is commonly used in the design of MLV-based vectors to prevent the expression of the full-length RNA by the LTR promoter.[58] A recent

study showed that the nuclear cytosolic levels of RNA transcripts decreased when most of the U3 region was deleted (381 nt).[59] The U3 deletion was performed in constructs containing the 5' LTR, 3' LTR, and Psi domain of MLV and a reporter gene (GFP or neomycin). This effect was partly due to the incorrect 3' end RNA processing but also to a default in unspliced RNA nuclear export. This defect was complemented by transferring the U3 region to another position within the vector, highlighting the presence of a position independent *cis*-acting sequence within the U3 region (Fig. 3).

Discussion

Viral FL RNA metabolism and transport are highly regulated processes relying on cellular pathways. MLV does not encode regulatory proteins unlike complex retroviruses and requires specialized interactions with cellular machinery to ensure splicing, stabilization and export of its viral RNA. In particular, splicing and export are closely linked processes resulting in the presence of a significant amount of unspliced transcripts in the cytosol. Thus, it is important to study the nuclear export of viral RNA while maintaining the intron-containing feature of the RNA (i.e. functional splice sites). Note that several aberrant splice sites have been reported in MLV-based vectors in which splicing was shown to depend on sequences located in *gag* (1183-1649 nt) and *pol* (5140-5400 nt) genes [60, 61]. None of these sequences are present in the different *cis*-elements described as facilitating export of the MLV unspliced RNA. However, it cannot be excluded that in native context of the complete virus, mutating export signals located in intronic region may deregulate splicing and indirectly affect RNA export. Furthermore, studies of HIV replication in murine cells have revealed species-specific defects at the level of splicing,[38] nuclear export or assembly efficiencies.[39, 40] These studies reinforce the notion that post-transcriptional regulation of retroviral gRNA requires specific cellular factors that are not functionally equivalent in murine and human cells. Therefore, it would be important to study MLV nuclear export within murine cell lines (Table 1) and using reporter constructs harboring MLV splice sites. Unfortunately, it is difficult to monitor the transport of the entire replication-competent MLV, since the effects of any mutation/deletion in the sequence of this multifunctional gRNA would make the interpretation ambiguous.

While MLV is among the first retroviruses discovered, we waited until 2014 to find out that MLV recruits the Tap-dependent export pathway, mainly due to the technical difficulties associated with the detection of a nuclear retention of MLV RNA when inactivating the Tap pathway. In contrast, clear and convincing nuclear accumulation of MPMV [62] was reported when inactivating CTE. What makes MLV so special? Actually, blocking the Tap-pathway systematically and specifically resulted into a fast and drastic degradation accounting for undetectable levels of MLV RNA in the nucleus.[27, 29] In addition, while other retroviruses harbor an export signal at their 3'end, MLV has the particularity to contain *cis*-acting signals throughout the molecule, at its ultimate 3' end (U3), at its 5' end (R, Psi) as well as in the coding sequence of *pol* (PTE, γ-CTE, and CAE) (Table 1). This dispersed pattern of *cis*-acting signals could suggest the existence of a conformational determinant but the complete RNA structure has not been solved yet and this hypothesis would require further investigation. Moreover, it is noticeable that Psi and CAE are particularly close to SD and SA splicing sites, respectively (Fig. 3), suggesting that splicing and export, two competitive events, could be regulated in a spatial way. Such competition could be crucial for the stability of the unspliced RNA as shown in HIV and MPMV studies.[63, 64] It might induce the recruitment of different co-factors which would explain the difference of the RNP on the spliced and unspliced RNA.[28] The fact that the same highly structured 5' region, Psi, also has the particularity to serve for both export and packaging functions suggests that the two processes might be intimately related to each other and might influence the cytosolic RNA fate.

Since Psi and PTE determinants are only present in unspliced FL RNA (Fig. 3), they could be recognized by Tap cofactors to specifically export these RNA species into the cytosol. This could explain how MLV discriminates its spliced and unspliced RNAs that are all dependent of the Tap export pathway.

Thus, further studies are required to better understand the relationship between the different *cis*-export signals and the other post-transcriptional regulation steps. To our knowledge, there is no inhibitory sequences identified in MLV RNA (unlike HIV-1 RNA) that could mediate the nuclear retention of the unspliced MLV RNA. Altogether, a better understanding of this particular mechanism might unravel new lines in these complex processes. Finally, export of intron-containing RNA represents a key step of retroviral infection that provides an interesting model to better understand the post-transcriptional regulation of the mammalian genes with persistence of introns leading to their expression in some cancers.

References

1. Braoudaki M, Tzortzatou-Stathopoulou F. Tumorigenesis related to retroviral infections. J Infect Dev Ctries 2011; 5:751-8.
2. Wolff L, Koller R, Hu X, Anver MR. A Moloney murine leukemia virus-based retrovirus with 4070A long terminal repeat sequences induces a high incidence of myeloid as well as lymphoid neoplasms. J Virol 2003; 77:4965-71.
3. Suzuki T, Shen H, Akagi K, Morse HC, Malley JD, Naiman DQ, Jenkins NA, Copeland NG. New genes involved in cancer identified by retroviral tagging. Nat Genet 2002; 32:166-74.
4. Lund AH, Turner G, Trubetskoy A, Verhoeven E, Wientjens E, Hulsman D, Russell R, DePinho RA, Lenz J, Van Lohuizen M. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. Nat Genet 2002; 32:160-5.
5. Kim R, Trubetskoy A, Suzuki T, Jenkins NA, Copeland NG, Lenz J. Genome-based identification of cancer genes by proviral tagging in mouse retrovirus-induced T-cell lymphomas. J Virol 2003; 77:2056-62.
6. Mikkers H, Allen J, Knipscheer P, Romeyn L, Hart A, Vink E, Berns A. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. Nat Genet 2002; 32:153-9.
7. Keswani RK, Pozdol IM, Pack DW. Design of hybrid lipid/retroviral-like particle gene delivery vectors. Mol Pharm 2013; 10:1725-35.
8. Cavazzana-Calvo M, Fischer A. Gene therapy for severe combined immunodeficiency: are we there yet? J Clin Invest 2007; 117:1456-65.
9. Aiyer S, Swapna GV, Malani N, Aramini JM, Schneider WM, Plumb MR, Ghanem M, Larue RC, Sharma A, Studamire B, Kvaratskhelia M, Bushman FD, Montelione GT, Roth MJ. Altering murine leukemia virus integration through disruption of the integrase and BET protein family interaction. Nucleic Acids Res 2014; 42:5917-28.
10. El Ashkar S, De Rijck J, Demeulemeester J, Vets S, Madlala P, Cermakova K, Debyser Z, Gijsbers R. BET-independent MLV-based Vectors Target Away From Promoters and Regulatory Elements. Mol Ther Nucleic Acids 2014; 3:e179.
11. Purcell DF, Martin MA. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J Virol 1993; 67:6365-78.
12. Houzet L, Battini JL, Bernard E, Thibert V, Mougel M. A new retroelement constituted by a natural alternatively spliced RNA of murine replication-competent retroviruses. EMBO J 2003; 22:4866-75.
13. Dorman N, Lever A. Comparison of viral genomic RNA sorting mechanisms in human immunodeficiency virus type 1 (HIV-1), HIV-2, and Moloney murine leukemia virus. J Virol 2000; 74:11413-7.

14. Hocine S, Singer RH, Grunwald D. RNA processing and export. Cold Spring Harb Perspect Biol 2010; 2:a000752.
15. Pante N. Nuclear pore complex structure: unplugged and dynamic pores. Dev Cell 2004; 7:780-1.
16. Montpetit B, Thomsen ND, Helmke KJ, Seeliger MA, Berger JM, Weis K. A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature 2011; 472:238-42.
17. Grunwald D, Singer RH, Rout M. Nuclear export dynamics of RNA-protein complexes. Nature 2011; 475:333-41.
18. Katenkamp K, Berndt A, Hindermann W, Wunderlich H, Haas KM, Borsi L, Zardi L, Kosmehl H. mRNA expression and protein distribution of the unspliced tenascin-C isoform in prostatic adenocarcinoma. J Pathol 2004; 203:771-9.
19. Vitulo N, Forcato C, Carpinelli EC, Telatin A, Campagna D, D'Angelo M, Zimbello R, Corso M, Vannozzi A, Bonghi C, Lucchin M, Valle G. A deep survey of alternative splicing in grape reveals changes in the splicing machinery related to tissue, stress condition and genotype. BMC Plant Biol 2014; 14:99.
20. Grzybowska EA. Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. Biochem Biophys Res Commun 2012; 424:1-6.
21. Harris ME, Hope TJ. RNA export: insights from viral models. Essays Biochem 2000; 36:115-27.
22. Wang B, Rekosh D, Hammarskjold ML. Evolutionary conservation of a molecular machinery for export and expression of mRNAs with retained introns. RNA 2015.
23. Smagulova F, Maurel S, Morichaud Z, Devaux C, Mougel M, Houzet L. The highly structured encapsidation signal of MuLV RNA is involved in the nuclear export of its unspliced RNA. J Mol Biol 2005; 354:1118-28.
24. Levin JG, Grimley PM, Ramseur JM, Berezesky IK. Deficiency of 60 to 70S RNA in murine leukemia virus particles assembled in cells treated with actinomycin D. J Virol 1974; 14:152-61.
25. Messer LI, Levin JG, Chattopadhyay SK. Metabolism of viral RNA in murine leukemia virus-infected cells; evidence for differential stability of viral message and virion precursor RNA. J Virol 1981; 40:683-90.
26. Bartels H, Luban J. Gammaretroviral pol sequences act in cis to direct polysome loading and NXF1/NXT-dependent protein production by gag -encoded RNA. Retrovirology 2014; 11:73.
27. Pessel-Vivares L, Ferrer M, Laine S, Mougel M. MLV requires Tap/NXF1-dependent pathway to export its unspliced RNA to the cytoplasm and to express both spliced and unspliced RNAs. Retrovirology 2014; 11:21.
28. Sakuma T, Tonne JM, Ikeda Y. Murine leukemia virus uses TREX components for efficient nuclear export of unspliced viral transcripts. Viruses 2014; 6:1135-48.
29. Sakuma T, Davila JI, Malcolm JA, Kocher JP, Tonne JM, Ikeda Y. Murine leukemia virus uses NXF1 for nuclear export of spliced and unspliced viral transcripts. J Virol 2014; 88:4069-82.
30. Pilkington GR, Purzycka KJ, Bear J, Le Grice SF, Felber BK. Gammaretrovirus mRNA expression is mediated by a novel, bipartite post-transcriptional regulatory element. Nucleic Acids Res 2015; 42:11092-106.
31. Dufu K, Livingstone MJ, Seebacher J, Gygi SP, Wilson SA, Reed R. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. Genes Dev 2010; 24:2043-53.
32. Gatfield D, Le Hir H, Schmitt C, Braun IC, Kocher T, Wilm M, Izaurralde E. The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila. Curr Biol 2001; 11:1716-21.

33. Zhang L, Shen J, Guarnieri MT, Heroux A, Yang K, Zhao R. Crystal structure of the C-terminal domain of splicing factor Prp8 carrying retinitis pigmentosa mutants. Protein Sci 2007; 16:1024-31.
34. Viphakone N, Hautbergue GM, Walsh M, Chang CT, Holland A, Folco EG, Reed R, Wilson SA. TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. Nat Commun 2012; 3:1006.
35. Ote I, Lebrun M, Vandevenne P, Bontems S, Medina-Palazon C, Manet E, Piette J, Sadzot-Delvaux C. Varicella-zoster virus IE4 protein interacts with SR proteins and exports mRNAs through the TAP/NXF1 pathway. PLoS One 2009; 4:e7882.
36. Huang Y, Steitz JA. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. Mol Cell 2001; 7:899-905.
37. Sanford JR, Gray NK, Beckmann K, Caceres JF. A novel role for shuttling SR proteins in mRNA translation. Genes Dev 2004; 18:755-68.
38. Zheng YH, Yu HF, Peterlin BM. Human p32 protein relieves a post-transcriptional block to HIV replication in murine cells. Nat Cell Biol 2003; 5:611-8.
39. Swanson CM, Puffer BA, Ahmad KM, Doms RW, Malim MH. Retroviral mRNA nuclear export elements regulate protein function and virion assembly. Embo J 2004; 23:2632-40.
40. Sherer NM, Swanson CM, Papaioannou S, Malim MH. Matrix mediates the functional link between human immunodeficiency virus type 1 RNA nuclear export elements and the assembly competency of Gag in murine cells. J Virol 2009; 83:8525-35.
41. Cupelli L, Okenquist SA, Trubetskoy A, Lenz J. The secondary structure of the R region of a murine leukemia virus is important for stimulation of long terminal repeat-driven gene expression. J Virol 1998; 72:7807-14.
42. Cupelli LA, Lenz J. Transcriptional initiation and postinitiation effects of murine leukemia virus long terminal repeat R-region sequences. J Virol 1991; 65:6961-8.
43. Mougel M, Tounekti N, Darlix JL, Paoletti J, Ehresmann B, Ehresmann C. Conformational analysis of the 5' leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. Nucleic Acids Res 1993; 21:4677-84.
44. Trubetskoy AM, Okenquist SA, Lenz J. R region sequences in the long terminal repeat of a murine retrovirus specifically increase expression of unspliced RNAs. J Virol 1999; 73:3477-83.
45. De Tapia M, Metzler V, Mougel M, Ehresmann B, Ehresmann C. Dimerization of MoMuLV genomic RNA: redefinition of the role of the palindromic stem-loop H1 (278-303) and new roles for stem-loops H2 (310- 352) and H3 (355-374). Biochemistry 1998; 37:6077-85.
46. Prats AC, Roy C, Wang P, Erard M, Housset V, Gabus C, Paoletti C, Darlix JL. cis elements and trans-acting factors involved in dimer formation of murine leukemia virus RNA. Journal of Virology 1990; 64:774-83.
47. Rein A. Take two. Nat Struct Mol Biol 2004; 11:1034-5.
48. Mougel M, Zhang Y, Barklis E. Cis-active structural motifs involved in specific encapsidation of Moloney Murine Leukemia Virus RNA. J Virol 1996; 70:5043-50.
49. Mann R, Baltimore D. Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. J Virol 1985; 54:401-7.
50. Mann R, Mulligan RC, Baltimore D. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 1983; 33:153-9.
51. Rein A, Harvin DP, Mirro J, Ernst SM, Gorelick RJ. Evidence that a central domain of nucleocapsid protein is required for RNA packaging in murine leukemia virus. J Virol 1994; 68:6124-9.

52. Tounekti N, Mougel M, Roy C, Marquet R, Darlix JL, Paoletti J, Ehresmann B, Ehresmann C. Effect of dimerization on the conformation of the encapsidation Psi domain of Moloney murine leukemia virus RNA. J Mol Biol 1992; 223:205-20.
53. Miyazaki Y, Garcia EL, King SR, Iyalla K, Loeliger K, Starck P, Syed S, Telesnitsky A, Summers MF. An RNA structural switch regulates diploid genome packaging by Moloney murine leukemia virus. J Mol Biol 2010; 396:141-52.
54. D'Souza V, Melamed J, Habib D, Pullen K, Wallace K, Summers MF. Identification of a high affinity nucleocapsid protein binding element within the Moloney murine leukemia virus Psi-RNA packaging signal: implications for genome recognition. J Mol Biol 2001; 314:217-32.
55. Basyuk E, Boulon S, Skou Pedersen F, Bertrand E, Vestergaard Rasmussen S. The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. J Mol Biol 2005; 354:330-9.
56. Gruter P, Tabernero C, von Kobbe C, Schmitt C, Saavedra C, Bachi A, Wilm M, Felber BK, Izaurralde E. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. Mol Cell 1998; 1:649-59.
57. Hoshi S, Odawara T, Oshima M, Kitamura Y, Takizawa H, Yoshikura H. cis-Elements involved in expression of unspliced RNA in Moloney murine leukemia virus. Biochem Biophys Res Commun 2002; 290:1139-44.
58. Maetzig T, Galla M, Baum C, Schambach A. Gammaretroviral vectors: biology, technology and application. Viruses 2011; 3:677-713.
59. Volkova NA, Fomina EG, Smolnikova VV, Zinovieva NA, Fomin IK. The U3 region of Moloney murine leukemia virus contains position-independent cis-acting sequences involved in the nuclear export of full-length viral transcripts. J Biol Chem 2014; 289:20158-69.
60. Machinaga A, Takase-Yoden S. Polyadenylation of Friend murine leukemia virus env-mRNA is affected by its splicing. Microbiol Immunol 2014; 58:474-82.
61. Yamamoto N, Takase-Yoden S. Friend murine leukemia virus A8 regulates Env protein expression through an intron sequence. Virology 2009; 385:115-25.
62. Hull S, Boris-Lawrie K. Analysis of synergy between divergent simple retrovirus posttranscriptional control elements. Virology 2003; 317:146-54.
63. Schambach A, Wodrich H, Hildinger M, Bohne J, Krausslich HG, Baum C. Context dependence of different modules for posttranscriptional enhancement of gene expression from retroviral vectors. Mol Ther 2000; 2:435-45.
64. Hammarskjold ML, Li H, Rekosh D, Prasad S. Human immunodeficiency virus env expression becomes Rev-independent if the env region is not defined as an intron. J Virol 1994; 68:951-8.

Table 1: Summary of the actors contributing to the MLV nuclear export.

Pathway/Factors	*cis*-acting element	Position	Cellular context	References
Tap	nd	nd	murine	[27]
Tap/p15, SRp20	γ-CTE	2918-3016 (*pol*)	human	[26]
Tap, UAP56, THOC5, THOC7	CAE	5621-5689 (*pol*)	human	[28] [29]
nd	RSL	1-28 (5' LTR)	murine	[41] [42] [44]
nd	Psi	210-374 (5' LTR)	murine	[23] [55]
nd	PTE	2238-3705 (*pol*)	human	[30]
nd	U3	7846-8229 (3' LTR)	murine	[59]

Fig.1

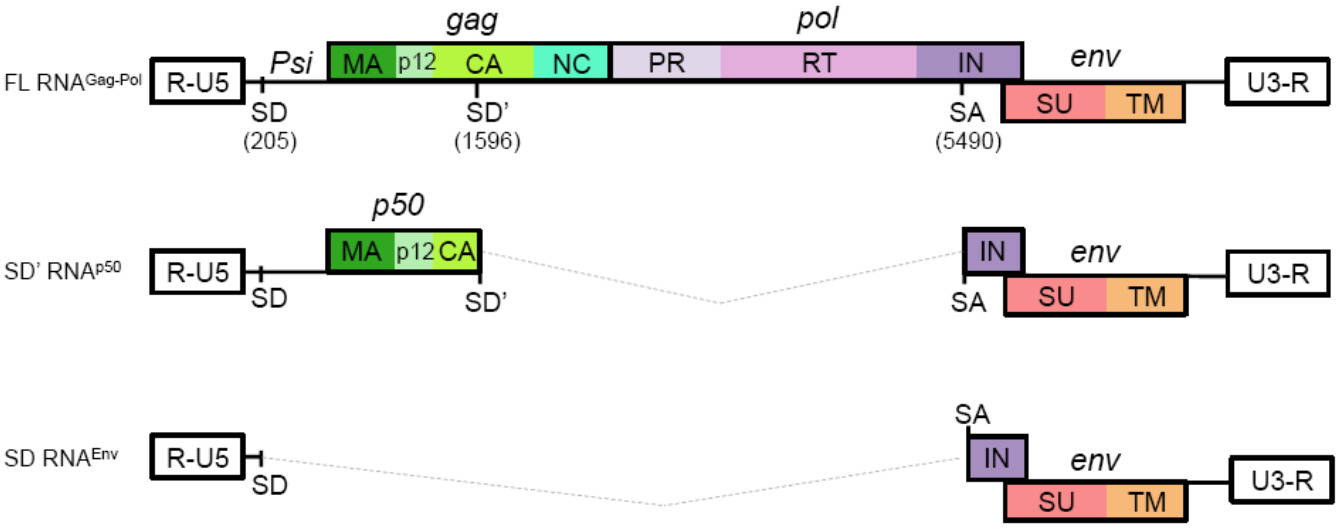
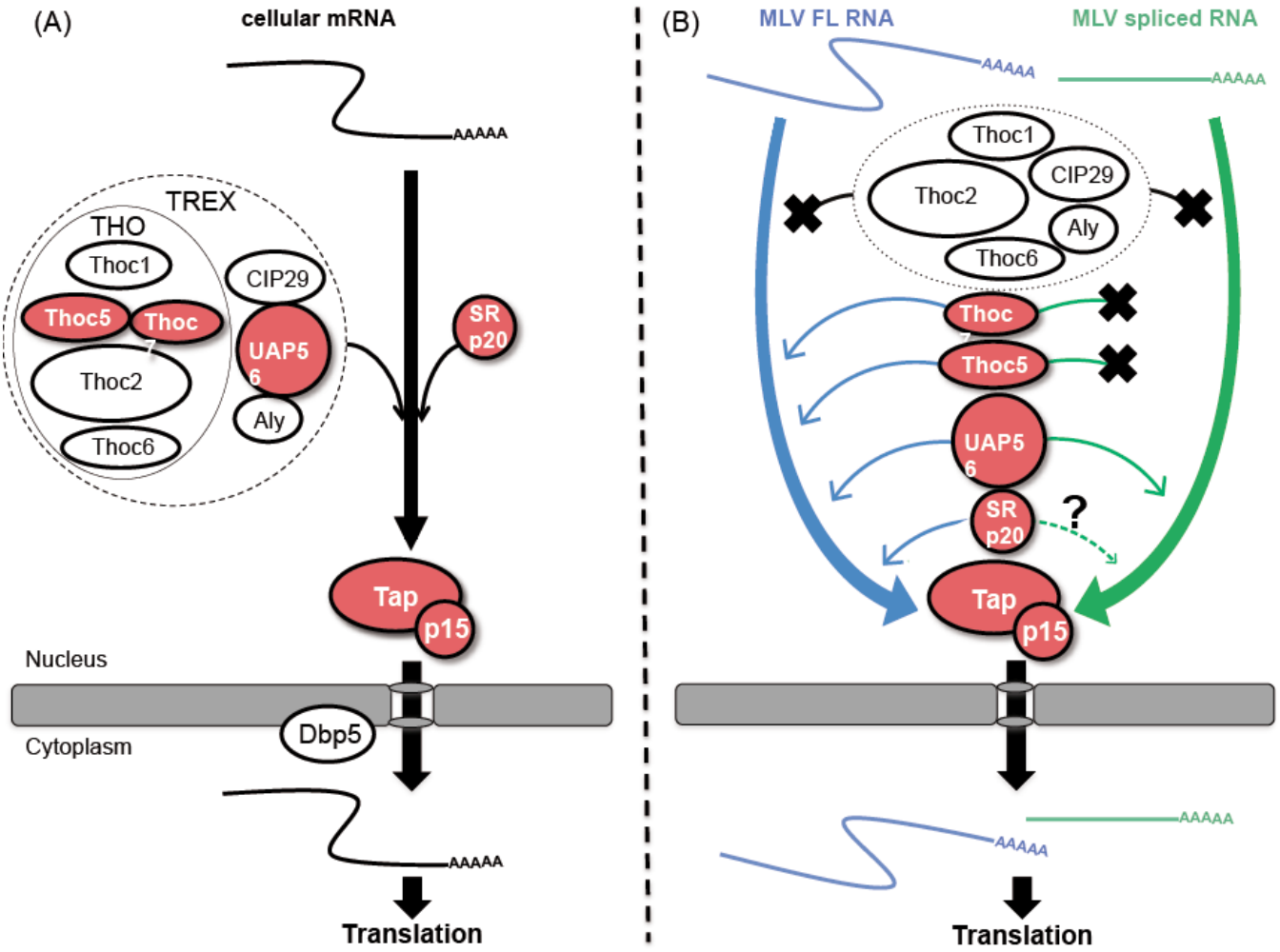


Figure 1: The MLV genome and subgenomic SD' and Env mRNAs.
The primary transcript (8.3kb RNA), named FL RNA$^{Gag-Pol}$, includes the three major genes: gag gene components, including the matrix (MA), p12, capsid (CA), and nucleocapsid (NC), the pol gene components, including the protease (PR), reverse transcriptase (RT) and integrase (IN) as well as the env sequence, coding for the surface (SU) and transmembrane (TM) components of viral envelope glycoprotein. A subset of this FL RNA goes through splicing and generates the SD' RNA^{p50} (4.4Kb) that derives from an alternative donor site (SD') within CA to the canonical splice acceptor site (SA) and codes for the p50 protein. A second fully-spliced RNA is generated, SD RNA^{Env} (3Kb) by splicing between the donor site (SD at position 205) and the acceptor site (SA at position 5490) encoding Env proteins.

Fig.2



Figures 2: MLV FL RNA is exported from the nucleus by Tap. *(A) Export of cellular mRNA via Tap-dependent pathway. The major mRNA export factor Tap, coupled to its heterodimeric partner, p15, is required for formation of an export-competent mRNP including the major adaptors: the TREX/THO complex and the SRp20 protein recruited cotranscriptionally. For simplicity, not all adaptors have been depicted. The factors shared with the MLV RNA export are in red. (B) In contrast with MPMV FL RNA, the interaction between MLV FL RNA and Tap seems indirect and requires the intervention of cellular co-factors (blue pathway). Some subunits from the TREX complex (THOC7, THOC5 and UAP56) and SRp20 proteins (in red) have been identified to promote this interaction. In contrast, only UAP56 has been reported to be involved in export of MLV spliced SD RNA (green pathway). The crosses indicate the dispensable factors for viral RNA export. The question mark indicates that SRp20 contribution to the spliced viral RNA export has not been analyzed yet.*

Fig. 3

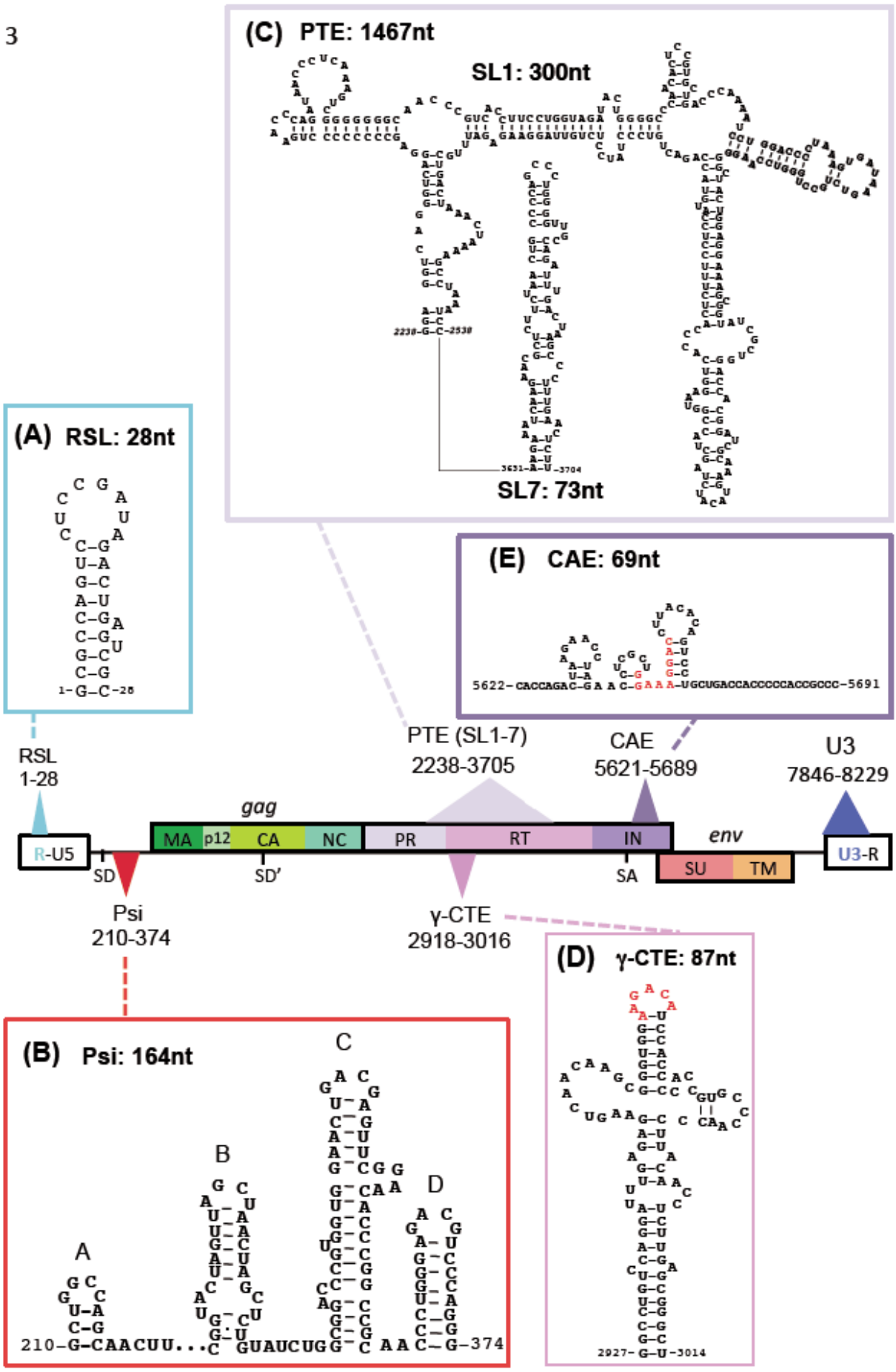


Figure 3: cis-acting motifs on MLV genome. *Different sequences throughout the Mo-MLV RNA have been reported to be involved in FL RNA export. From 5' to 3': The RSL (R stem-loop), a signal made of 28nt folding in one stem-loop as pictured in (A); the Psi encapsidation signal which folds in four stem-loops (B); PTE which covers a large region in pol gene and contains seven stem-loops with the first and seventh motif (SL1 & SL7 depicted in C) of special importance;*

the γ-CTE (D) within the PTE region harbors a short sequence AAGACA (in red) also found in the MPMV CTE and CAE which folds in 3 stem-loops (E), contains a 9nt core sequence (in red) necessary but not sufficient for the function of this motif. At the 3' end the U3 region is required for the specific export of unspliced RNA however no precise motif has been highlighted in this region yet.