THE EVOLUTIONARY INTERPLAY BETWEEN EXOGENOUS AND ENDOGENOUS SHEEP BETARETROVIRUSES

Alessia Armezzani

Thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy

MRC-University of Glasgow Centre for Virus Research
Institute of Infection, Immunity and Inflammation
College of Medicine, Veterinary and Life Sciences

March 2012

Abstract

Retroviruses must integrate their genome into the host DNA as a necessary step of their replication cycle. Normally, retroviruses integrate into somatic cells and are transmitted, from infected to uninfected hosts, as "exogenous" retroviruses. On rare occasions, they can infect germ line cells and become part of the host genome as "endogenous" retroviruses (ERVs), which are transmitted vertically to the offspring and inherited as Mendelian genes. During evolution, most ERVs have accumulated mutations that rendered them defective and unable to produce infectious viral particles. Some ERVs, however, have maintained intact open reading frames for some of their genes, and have been co-opted by the host as they fulfil important biological functions. Sheep betaretroviruses represent a unique model to study the complex evolutionary interplay between host and pathogen in natural settings. In infected sheep, the exogenous and pathogenic Jaagsiekte sheep retrovirus (JSRV) co-exists with the highly related endogenous JSRVs (enJSRVs). The sheep genome harbours at least twenty-seven enJSRV loci and, most likely, the process of endogenization is still occurring. During evolution, one of these enJSRV loci, enJS56A1, has acquired a defective and transdominant Gag polyprotein that blocks the late replication steps of related retroviruses, by a mechanism known as JSRV late restriction (JLR). Interestingly, enJSRV-26, a provirus that integrated in the sheep germ line less than two hundred years ago, possesses the unique ability to escape JLR. In this thesis, the molecular basis of JLR escape was investigated. The main determinant of JLR escape was identified in the signal peptide of enJSRV-26 envelope protein (SP26). A single amino acid substitution in SP26 was found to be responsible for altering its intracellular localization as well as its function as a post-transcriptional regulator of viral gene expression. Interestingly, interference assays demonstrated that enJSRV-26 relies on the presence of the functional signal peptide of enJS56A1 envelope protein (SP56) in order to escape JLR. In addition, the ratio between enJSRV-26 and enJS56A1 Gag polyproteins was found to be critical to elude JLR. Finally, sequence analyses revealed that the domestic sheep has acquired, by genome amplification, several copies of the enJS56A1 provirus, reinforcing the hypothesis that this locus has provided an evolutionary advantage to the host. This study unveils critical aspects of JLR that were previously unknown, and provides new insights on the molecular

mechanisms governing the interplay between endogenous and exogenous sheep betaretroviruses.

Table of Contents

List of Tables

List of Figures

List of Accompanying Material

This thesis includes a copy of the following article:

Armezzani, A., Arnaud, F., Caporale, M., di Meo, G., Iannuzzi, L., Murgia, C. and Palmarini, M. (2011) The signal peptide of a recently integrated endogenous sheep betaretrovirus envelope plays a major role in eluding gag-mediated late restriction. *J Virol*, 85, 7118-7128.

To Fred, for his loving support and continuous enthusiasm

Acknowledgments

«Dicebat Bernardus Carnotensis nos esse quasi nanos, gigantium humeris insidentes, ut possimus plura eis et remotiora videre, non utique proprii visus acumine, aut eminentia corporis, sed quia in altum subvenimur et extollimur magnitudine gigantea[1]».

Grateful thanks are extended to all those who played a smaller or larger part in helping me to complete my thesis. Special mention must go to my principal supervisor, Professor Massimo Palmarini, who gave me the opportunity to join his group and undertake my PhD project, and to Claudio Murgia, for his invaluable assistance and support throughout the bleak months of experimental failures... especially for being critical in turning them into successes!

I will be forever grateful to Lita Murphy, who has made Glasgow and the laboratory a nice place to stay: I greatly miss her warm friendship, humour and cupcakes, and Maxime Ratinier, for his constant encouragement and enthusiasm, and for giving me the possibility to learn from my mistakes!

Special thanks also to Matt Golder, Linda Bellingham, Anne Terry and all the guys of the "Glasgow walking and hillwalking group" for having offered me the possibility to discover the beauties of Scotland!!

I am deeply grateful to my sister and my parents, for their kindness and constant encouragement: without them, I would never have started this beautiful adventure!

Finally, and on a more personal note, I am profoundly thankful to Frédérick Arnaud, for his love and constant support, his bright intelligence and great mentorship: without him (and Virgule!) I would never have ended this beautiful adventure!

[1] «Bernard of Chartres used to say that we [the Moderns] are like dwarves perched on the shoulders of giants [the Ancients], and thus we are able to see more and farther than the latter. And this is not at all because of the acuteness of our sight or the stature of our body, but because we are carried aloft and elevated by the magnitude of the giants» (from Troyan, Scott D., Medieval Rhetoric: A Casebook, London, Routledge, 2004, p. 10).

Author's declaration

I hereby declare that the work presented in this thesis is original, and was carried out by either myself or with due acknowledgement. All additional sources of information have likewise been acknowledged.

This work has not been presented for the award of a degree at any other university, but has been reproduced in part in the following scientific paper:

Armezzani, A., Arnaud, F., Caporale, M., di Meo, G., Iannuzzi, L., Murgia, C. and Palmarini, M. (2011) The signal peptide of a recently integrated endogenous sheep betaretrovirus envelope plays a major role in eluding gag-mediated late restriction. *J Virol*, 85, 7118-7128.

Alessia Armezzani

March 2012

Abbreviations

A	Alanine (one-letter amino acid code)
AIDS	Acquired immunodeficiency syndrome
ALV	Avian leucosis virus
APOBEC	Apolipoprotein B mRNA-editing enzyme
ARM	Arginine-rich motif
ASLV	Avian sarcoma/leucosis virus
BAC	Bacterial artificial chromosome
BLV	Bovine leukaemia virus
BNC	Binucleate cell
BST-2	Bone marrow stromal cell antigen 2
CA	Capsid
cDNA	Complementary DNA
cfr.	Confront
CMV	Cytomegalovirus
Crm1	Chromosome maintenance region 1
CsA	Cyclosporin A
CT	Cytoplasmic tail
CTE	Constitutive transport element
CypA	Cyclophilin A
D	Aspartic acid (one-letter amino acid code)
DAPI	4', 6'-diamidino-2-phenylindole
DIS	Dimerization initiation site
DLS	Dimeric-linkage structure
DNA	Deoxyribonucleic acid
dUTP	Deoxyuridine triphosphate
e.g.	exempli gratia, for example
EDTA	Ethylenediaminetetraacetic acid
EIAV	Equine infectious anaemia virus
ELISA	Gag enzyme-linked immunosorbent assay
Elp	Env leader protein
enJSRV	Endogenous Jaagsiekte sheep retrovirus
ENTV	Enzootic nasal tumour virus
Env	Envelope

ER	Endoplasmic reticulum
ERV	Endogenous retrovirus
ESCRT	Endosomal sorting complexes required for transport
FeLV	Feline leukaemia virus
FIV	Feline immunodeficiency virus
FMLV	Friend MLV
FV	Foamy virus
Fv1	Friend-virus-susceptibility-1
Fv4	Friend-virus-susceptibility-4
Gag	Group-specific antigen
GPI	Glycosylphosphatidylinositol
HA	Hemagglutinin
HCV	Hepatitis C Virus
HERV	Human endogenous retrovirus
HFV	Human foamy virus
HIV	Human immunodeficiency virus
HMG	High mobility group
HNF-3	Hepatocyte nuclear factor-3
HRES	HTLV type I-related endogenous sequences
HTLV	Human T-cell leukaemia virus
i.e.	id est, that is
IAP	Intracisternal A-type particle
IFN	Interferon
IN	Integrase
IP	Immunoprecipitation
IRES	Internal ribosome entry site
JLR	JSRV late restriction
JSE-SP	Signal peptide of JSRV envelope glycoprotein
JSRV	Jaagsiekte sheep retrovirus
K	Lysine (one-letter amino acid code)
LAPC	Lung alveolar proliferating cell
LINE	Long interspersed nucleotide element
LTR	Long terminal repeat
MA	Matrix
MLV	Murine leukaemia virus

MMTV	Mouse mammary tumour virus
M-PMV	Mason-Pfizer monkey virus
mRNA	Messenger RNA
MTOC	Microtubule organizing centre
Mtv	MMTV ERV
MVB	Multi-vescicular body
MVV	Maedi-Visna virus
MYA	Million years ago
NaCl	Sodium chloride
NaF	Sodium fluoride
NC	Nucleocapsid
Nef	Negative regulatory factor
NES	Nuclear export signal
NFI	Nuclear factor I
NLS	Nuclear localization signal
NoLS	Nucleolar localization signal
NP-40	Nonidet P-40, octylphenoxypolyethoxyethanol
OPA	Ovine pulmonary adenocarcinoma
pbs	Primer binding site
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PI(4,5)P2	Phosphatidylinositol 4,5-bisphosphate
PIC	Pre-integration complex
PMSF	Phenylmethylsulfonyl fluoride
Pol	Polymerase
Poly(A)	Polyadenylic acid
PPT	Polypurine tract
PR	Protease
qPCR	Quantitative polymerase chain reaction
R	Arginine (one-letter amino acid code)
Rej	Regulator of JSRV expression
RejRE	Rej-responsive element
RELIK	Rabbit endogenous lentivirus type K
Rem	Regulator of export of MMTV (mRNA)
Rev	Regulator of virion

RmRE	Rem-responsive element
RNA	Ribonucleic acid
rpm	Revolutions per minute
RRE	Rev-responsive element
RSV	Rous sarcoma virus
RT	Reverse transcriptase
RTC	Reverse transcription complex
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SIN	Sindbis virus
SINE	Short interspersed nucleotide element
siRNA	Small interfering RNA
SIV	Simian immunodeficiency virus
SP	Signal peptide
SP18	Signal peptide of the enJSRV-18 envelope protein
SP26	Signal peptide of the enJSRV-26 envelope protein
SP56	Signal peptide of the enJS56A1 envelope protein
SPase	Signal peptidase
SPPase	Signal peptide peptidase
SPRE	Signal peptide-responsive element of enJSRV/JSRV
SRP	Signal recognition particle
SRY	Sex determining region of Y chromosome
SU	Surface
Tap	Tip-associated protein
TAR	Transactivation response element
Tas	Transactivator of spumaretroviruses
Tat	Transactivator of transcription
TBS	Tris buffered saline
TGN	*Trans*-Golgi network
TM	Transmembrane
TRIM	Tripartite motif
Tris-HCl	Tris-hydroxymethyl-aminomethane hydrochloride
tRNA	Transfer RNA
USE	Upstream enhancer element
UTR	Untranslated region
Vif	Virion infectivity factor

Vps	Vacuolar protein sorting
Vpu	Viral protein U
VR	Variable region
vRNA	Viral genomic RNA
W	Tryptophan (one-letter amino acid code)
WDSV	Walleye dermal sarcoma virus
ZAP	Zinc-finger antiviral protein
Ψ	Packaging signal sequence (Psi)

Chapter I

Introduction

1.1 Historic perspectives

The dawn of retrovirology dates back to the beginning of the 20th century, when Henri Vallée and Henri Carré showed that equine anaemia was transmitted by a «*filterable agent*» (Vallée and Carré, 1904), now known as equine infectious anaemia virus (EIAV). Few years later, Vilhelm Ellermann and Oluf Bang reported that chicken leucosis was transmitted by cell-free filtrates (Ellermann and Bang, 1908), and Peyton Rous obtained similar results while studying a solid tumour (sarcoma) of chicken in 1911 (Rous, 1911). We now know that the etiological agent discovered by Ellerman and Bang was the avian leucosis virus (ALV), while the virus isolated by Rous bears today his name (i.e., Rous sarcoma virus, RSV) (Vogt PK, 1997).

For many years, the scientific community did not show great interest for retroviruses, due to their apparent irrelevance to mammals. Moreover, the lack of reliable cell cultures and biochemical techniques rendered the study of "filterable agents" quite inaccurate, costly and time-consuming. These attitudes began to shift when John Bittner (Bittner, 1936) and Ludwik Gross (Gross, 1951) expanded the retroviral paradigm to mammalian hosts, by demonstrating that retroviruses can cause neoplastic disease in mice. Furthermore, the development of quantitative assays to measure viral infectivity by Howard Temin and Harry Rubin (Temin and Rubin, 1958) greatly facilitated virological studies.

In 1970, Howard Temin and David Baltimore independently discovered the reverse transcriptase (Baltimore, 1970; Temin and Mizutani, 1970), for which they were awarded the Nobel Prize in 1975. This finding gave the field its first real molecular tool for the sensitive detection of retroviruses, that were named after this discovery "*retraviruses*" (for *re*verse *tra*nscriptase containing *viruses*) (Dalton et al., 1974). It was only later, and for reasons of linguistic fluidity, that they gained their actual name.

The discovery of the reverse transcriptase was also critical to validate the «*DNA provirus hypothesis*» formulated by Howard Temin in 1960s, which postulated that retroviruses integrate into the DNA of somatic cells (Temin, 1964). However, the notion that retroviruses could also integrate into the host germ

line, and be inherited as Mendelian genes, was still regarded as bizarre. In the late 1960s and early 1970s, ALV, murine leukaemia virus (MLV) and murine mammary tumour virus (MMTV) were found endogenous in the genome of their hosts (Huebner and Todaro, 1969; Payne and Chubb, 1968; Varmus et al., 1972; Weiss and Payne, 1971), while the first description of human endogenous retroviruses (HERVs) was reported in 1970s (Kalter et al., 1973). Today, we know that retroviral genome invasions have occurred throughout the evolution of vertebrates (Jern and Coffin, 2008) and continue to the present day (Arnaud et al., 2007a; Tarlinton et al., 2006).

The realization that retroviruses were involved in oncogenesis in chickens, mice, cats (Jarrett et al., 1964a; Jarrett et al., 1964b) and non-human primates (Kawakami et al., 1972; Theilen et al., 1971) led to postulate that they might have been generally involved in tumorigenesis. This, in turn, generated extensive efforts to identify analogous human viruses that resulted in the isolation of human T-cell leukaemia virus (HTLV) and human immunodeficiency virus (HIV) in 1980s (Barré-Sinoussi et al., 1983; Poiesz et al., 1980). These discoveries galvanized widespread interest in retroviruses, and provided broader insights into several aspects of cellular and molecular biology, including gene regulation, cell growth and immune response.

Over the past decades, scientists have been attracted by retroviruses as important pathogens in humans and animals, and powerful models to study many aspects of both normal and cancer cells. Despite the extraordinary productivity of retrovirology in recent years, questions seem larger and more numerous than ever. The major outlines of the retroviral replication cycle are firmly drawn, but mechanisms of key events, such as viral persistence, are just now coming into view. Important human pathogens have been identified among retroviruses, but strategies for prevention and cure are still desperately needed, and have assumed a greater urgency because of the acquired immunodeficiency syndrome (AIDS) pandemic.

1.2 Retroviral taxonomy

Retroviruses belong to the unique family of Retroviridae, which uses the virally encoded reverse transcriptase (RT) to replicate viral genomic RNA (vRNA) into a

double-stranded DNA intermediate (cDNA). Retroviruses have been shown to infect many vertebrates species, including fish, birds, reptiles and mammals, and cause a wide variety of diseases, such as cancers (leukaemias, mammary carcinomas), immunodeficiencies and arthritis (Vogt PK, 1997). An interesting exception is represented by members of the genus Spumavirus, which can persistently infect wild non-human primates, felines, bovines, equines and small ruminants, with no apparent pathological consequences (Murray et al., 2008).

Retroviruses were originally classified into four groups, named A-type through D-type, according to the morphology and the position of the nucleocapsid core observed by electron microscopy. A-type viruses were defined as non-enveloped cytoplasmic particles, with an electron-lucent centre and one or two concentric electron-dense rings, thereby overall resembling a doughnut. This term is now used to refer to immature particles, such as intracytoplasmic particles formed by some retrotransposons (intracisternal A-type particles, IAPs). B-type viruses, such as Jaagsiekte sheep retrovirus (JSRV) and MMTV, display a round and eccentrically positioned core. Conversely, C-type viruses contain a central and spherical inner core, with ALV being the prototype. Finally, D-type viruses, including Mason-Pfizer monkey virus (M-PMV), have a bar-shaped core (Vogt VM, 1997). Recently, retroviruses have been re-classified by the International Committee on Taxonomy of Viruses into two subfamilies, the Orthoretrovirinae and Spumaretrovirinae[2], and seven genera (Table 1).

Table 1 | Classification of Retroviruses.

Subfamily	Genus	Prototype virus	Genome	Morphology
Orthoretrovirinae	Alpharetrovirus	Avian leucosis virus (ALV)	Simple	C-type
	Betaretrovirus	Mouse mammary tumour virus (MMTV)	Simple	B-type D-type
	Gammaretrovirus	Murine leukaemia virus (MLV)	Simple	C-type
	Deltaretrovirus	Bovine leukaemia virus (BLV)	Complex	C-type
	Epsilonretrovirus	Walleye dermal sarcoma virus (WDSV)	Simple	C-type
	Lentivirus	Human immunodeficiency virus 1 (HIV-1)	Complex	C-type
Spumaretrovirinae	Spumavirus	Human foamy virus (HFV)	Complex	D-type

[2] Note that the Spumaretrovirinae subfamily comprises only the spumavirus genus.

All retroviruses share a similar genetic organization consisting of four open reading frames, including *gag*, *pro*, *pol* and *env*. These genes are normally translated as polyprotein precursors, and subsequently processed by viral or cellular proteases to yield, respectively, internal structural proteins, enzymes, and envelope glycoproteins. Some retroviruses, such as JSRV and M-PMV, contain only the four open reading frames mentioned above and are therefore termed "simple" retroviruses. Conversely, other retroviruses express several genes that encode regulatory and accessory proteins, required for efficient viral replication, nuclear export, transmission and evasion from innate and acquired immunity. These retroviruses that include, among others, lentiviruses and foamy viruses (FVs), are thus referred to as "complex" retroviruses (Vogt PK, 1997).

1.3 Exogenous and endogenous retroviruses

Normally, retroviruses infect host somatic cells, passing horizontally from an infected to an uninfected organism as exogenous retroviruses. Occasionally, retroviruses can infect germ cells, thus leading to integrated proviruses ("endogenous retroviruses", ERVs), which are transmitted vertically to the offspring and inherited as Mendelian genes (Jern and Coffin, 2008). Until recently, it was believed that only simple retroviruses could become endogenous in their host (Weiss, 2006). However, the recent identification of endogenous lentiviruses in the genome of European rabbits (termed RELIK, for rabbit endogenous lentivirus type K) (Katzourakis et al., 2007) and gray mouse lemurs (Gifford et al., 2008), together with the identification of endogenous retroviruses related to spumaviruses (HERV) (Kalter et al., 1973; Ono et al., 1986) and deltaviruses (HTLV type I-related endogenous sequences, HRES) (Perl et al., 1989) in humans, has extended this paradigm to all the members of the Retroviridae family.

Currently, there is no standardised nomenclature for ERVs. HERVs have been classified on the basis of cellular tRNA usage. For example, members of HERV-K group possess a primer binding site (pbs) complementary to a lysine[3]-tRNA (Larsson et al., 1989). However, this method of classification does not faithfully represent phylogenetic relatedness, as often HERVs belonging to the same group display differences in their tRNA usage (Lavie et al., 2004). Conversely,

[3] According to the one-letter amino acid code, lysine is represented by the letter K.

phylogenetic analysis of the Pol polyprotein has proved to be a more useful criterion to classify ERVs, which have been thus divided into three groups: class I, II and III ERVs. Class I ERVs are related to gammaretroviruses, class II to betaretroviruses and class III to alpharetroviruses (Jern et al., 2005).

ERVs share the same genetic structure and organization of their exogenous counterparts. In general, "modern" ERVs exist both as endogenous and exogenous viruses. On the other hand, "ancient" ERVs do not possess any exogenous counterpart, leading to the hypothesis that the process of endogenization is one of the steps that contributes to the extinction of horizontally transmitted infectious retroviruses (Denner, 2010).

The exact mechanisms underlying the expansion of ERVs within the host genome (i.e., the copy number variation) and the host population remain still poorly understood. From a genetic point of view, ERVs are "retrotransposons" with long terminal repeats (LTRs) at each end of their genome. ERVs, together with DNA transposons, long interspersed nucleotide elements (LINEs) and short interspersed nucleotide elements (SINEs), belong to the group of transposable elements, firstly discovered in 1950 by Barbara McClintock (McClintock, 1950) (Fig. 1).

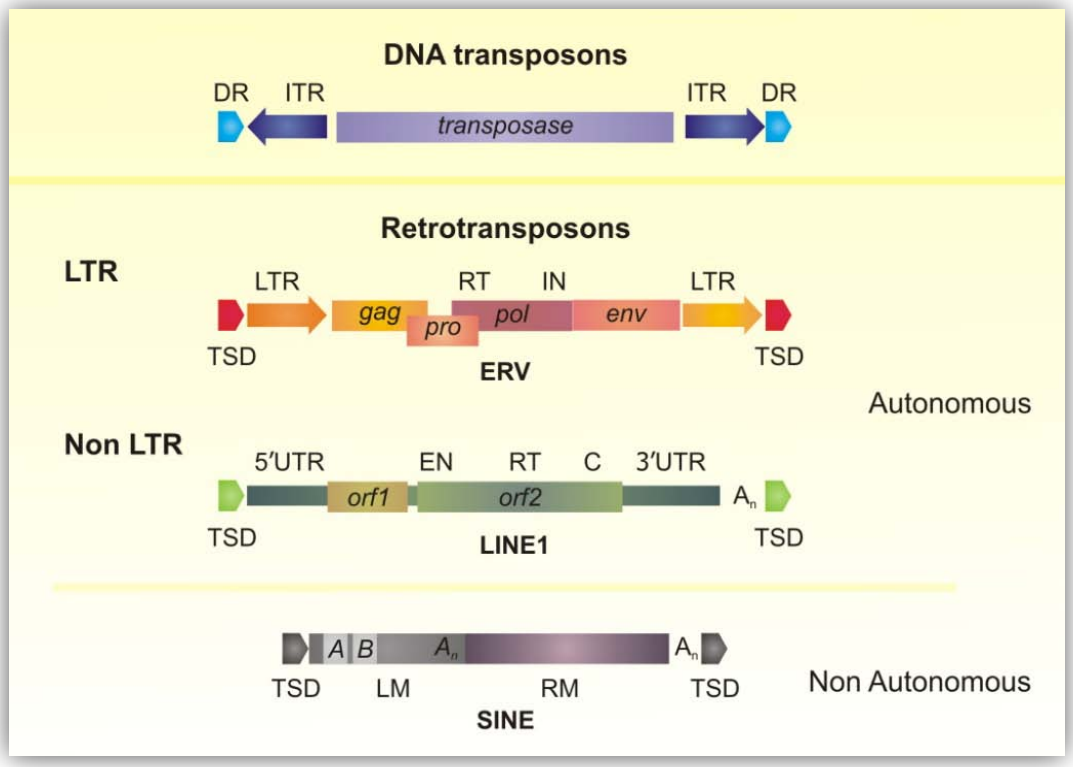


Figure 1| Schematic representation of transposable elements. A/B, A- and B-box Pol III promoter; A_n, poly(A); C, zinc knuckle domain; DR, direct repeat; EN, endonuclease; IN, integrase; ITR, inverted terminal repeat; LM, left monomer; LTR, long terminal repeat; RT, reverse transcriptase; TSD, target site duplication. Figure adapted from (Goodier and Kazazian, 2008).

By definition, retrotransposons move from one genomic location to another one in a "copy and paste" manner, which involves reverse transcription of an RNA intermediate and insertion of its cDNA copy at the new genomic site (Goodier and Kazazian, 2008). It is therefore possible that ERVs may amplify and then "jump" from one locus to another one within a host germ cell. This "intracellular spread" is the mechanism adopted by murine IAPs, which originally derived from retroviruses but lack the *env* gene, and are therefore unable to give rise to infectious particles (Maksakova et al., 2006). Alternatively, ERVs may form viral particles that can essentially act as exogenous retroviruses and re-infect the host germ line, spreading in this way across the host population and, occasionally, across species. Most ERVs have amplified probably using both mechanisms, giving rise to infectious viral particles that can infect the host genome and, perhaps, the germ line of the host population, but rarely crossing the species barrier (Jern and Coffin, 2008).

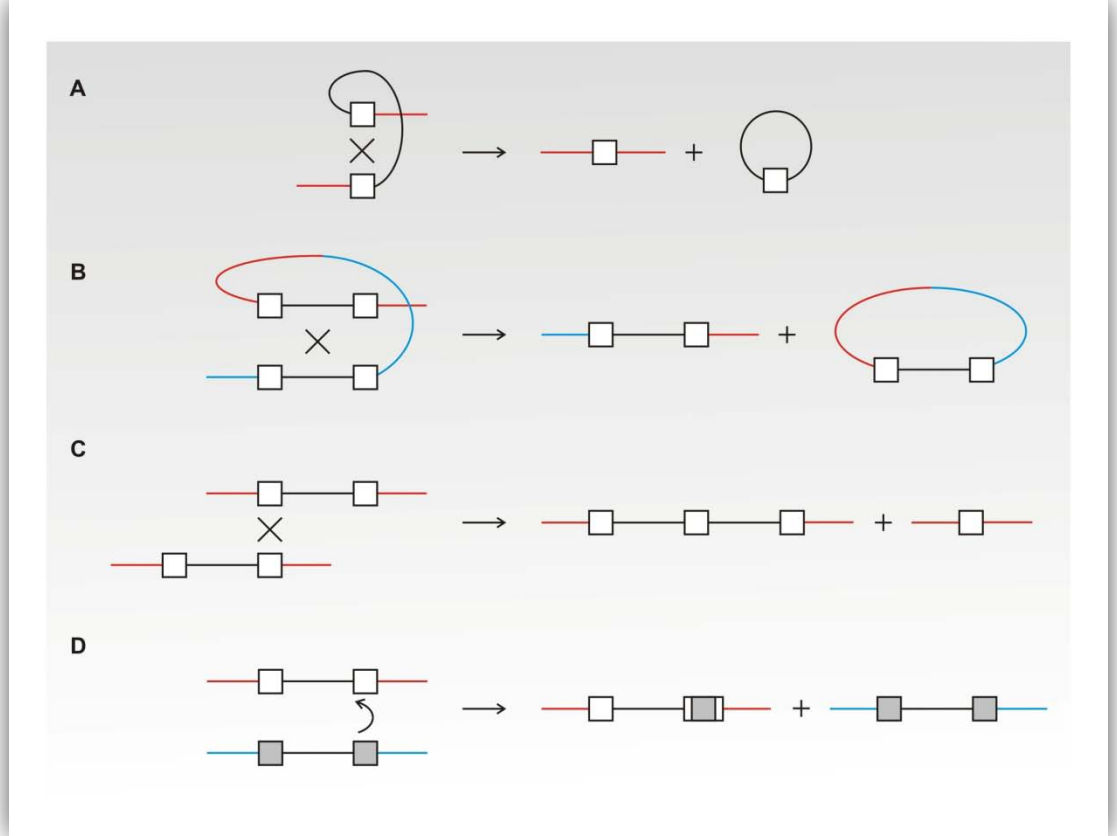


Figure 2| **DNA recombination events involving ERVs.** See text for details. Boxes indicate LTR, while black and coloured lines represent proviral and cellular sequences, respectively. Figure adapted from (Stoye, 2001).

ERVs display a significant tendency to recombine, through different mechanisms. Homologous recombination between the flanking LTRs of a provirus results in the excision of proviral DNA, with consequent formation of "solo LTR" (Fig. 2A). For most ERVs, this relic structure represents the only remnant of their cognate ancestral proviruses embedded in the host genome. Usually, solo LTRs form soon after proviral integration, probably due to the higher recombination rate between identical flanking LTRs that have not accumulated mutations (Belshaw et al., 2007). Similarly, homologous recombination between two proviruses located on the same chromosome leads to significant deletions of both proviral and genetic sequence, with consequent rearrangements of the latter (Fig. 2B). Recombination between the 3' and 5' LTRs of allelic proviruses leads to a tandem provirus sharing, in the middle, a single LTR on one chromatid, and a solo LTR on the other (Fig. 2C). Homologous recombination between proviruses located on the same chromosomes or on different ones can also give rise to nonreciprocal exchange without loss of proviral sequences. This process known as "gene conversion" leads to the exchange of all or part of one proviral sequence to that of the other (Jern and Coffin, 2008; Stoye, 2001) (Fig. 2D).

1.4 Genetic organization of the retroviral genomic RNA

Retroviruses are enveloped viruses that package two copies of positive single-stranded RNA molecules, consisting of four open reading frames flanked by repeated (R) and unique (U) sequences that form the untranslated regions (UTRs). The retroviral genome contains also a polypurine tract (PPT), located just upstream of the 3'UTR, that serves as a primer for the synthesis of the plus-strand DNA during reverse transcription (see paragraph 1.6.2). The open reading frames are always organised in the same order: *gag*, *pro*, *pol* and *env*. Complex retroviruses possess also other genes, located downstream of *pol* and/or *env*, that encode for accessory and regulatory proteins, critical for viral replication and to counteract host defences (e.g., HIV-1 Tat, Vif, Vpr and Vpu; HTLV Tax and FV Tas) (Bannert et al., 2010) (Fig. 3).

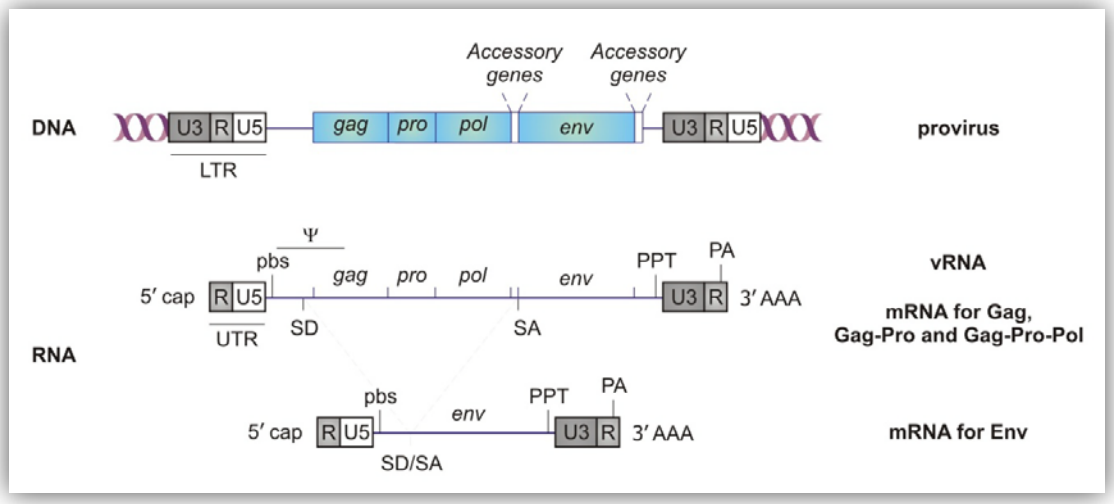


Figure 3| Genetic organization of a typical retrovirus. The proviral DNA (top) is characterized by the presence of a long terminal repeat (LTR) sequence at each end of the genome, abutting cellular sequences (in purple). Accessory genes are usually located as shown. The primary transcript (middle) is a single-stranded RNA molecule that serves both as viral genomic RNA (vRNA) and mRNA for Gag, Pro and Pol proteins. The Env protein, on the other hand, is generated by a double splicing event (bottom). AAA, poly(A) tail; PA, polyadenylation signal; pbs, primer-binding site; PPT, polypurine tract; R, repeat sequence; SA, splice acceptor site; SD splice donor site; U3, 3' unique sequence; U5, 5' unique sequence; UTR, untranslated region; ψ (Psi), packaging signal. Figure adapted from (Gifford and Tristem, 2003).

In retroviruses, the unspliced RNA serves both as a template for viral proteins translation and source of vRNA that is packaged into assembling virions. As such, both RNA species share identical genetic organization (D'Souza and Summers, 2005). In addition, as they are both transcribed by the host RNA polymerase II, they undergo the same post-translational modifications as those of eukaryotic

mRNAs, including the 5' cap structure and poly(A) tail[4] (Balvay et al., 2007) (Fig. 3).

Both 5'UTR and 3'UTR harbour *cis*-acting RNA motifs that are involved in several aspects of the retroviral replication cycle. The 5'UTR contains (i) a poly-adenylation stem loop, which is responsible for the addition of the 3' poly (A) tail, (ii) the pbs, where the cellular tRNA is recruited for the initiation of the reverse transcription, (iii) a dimerization initiation site (DIS), which is involved in the formation of the kissing-loop complex between two vRNAs during packaging, (iv) a major splice donor (SD), used for the production of viral subgenomic mRNAs, and finally (v) a core packaging signal (Psi, Ψ), which selects and directs the vRNA to the viral assembly machinery. Conversely, the 3'UTR is characterized by the presence of RNA motifs that control the termination of transcription, including an upstream enhancer element (USE), essential for efficient polyadenylation, and a poly(A) signal site (Balvay et al., 2007; Bannert et al., 2010).

1.5 Virion structure and viral proteins

Retroviral virions contain two copies of single-stranded vRNA molecules, the Gag polyprotein, the viral enzymes encoded by *pro* and *pol*, cellular RNAs and cellular proteins, such as cyclophilin A (CypA) and apolipoprotein B mRNA-editing catalytic enzyme 3 (APOBEC3) (Onafuwa-Nuga et al., 2006; Strebel et al., 2009).

Upon release, retroviral particles display an "immature" morphology, characterized by an electron-lucent centre, doughnut-shaped. The proteolytic cleavage of the Gag and Gag-Pol polyprotein precursors initiates a series of structural rearrangements that ultimately lead to a "mature" virion, containing an electron-dense core. This process is mediated by the viral protease (PR) encoded by *pro*, during a process referred to as "maturation" (Swanstrom and Wills, 1997).

[4] The 5' cap structure consists in the insertion of a guanosine methylated at position 7, as a first nucleotide of the RNA transcript. The poly(A) tail is a stretch of polyadenylate residues (usually about two hundred residues) that terminates the 3' end of the transcript (Balvay et al., 2007).

Gag is the internal structural protein of all retroviruses, which shapes viral capsid and associates with the vRNA to form viral core. In orthoretroviruses, the Gag polyprotein is processed into at least three mature proteins: nucleocapsid (NC), capsid (CA) and matrix (MA) (Fig. 4). In addition, some retroviruses, such as HIV-1 and RSV, possess spacer peptides between MA and CA, and CA and NC, which have been proposed to influence some steps of viral assembly and particle release. In spumaviruses, by contrast, Gag is not processed in the canonical NC, CA and MA domains, and thus, after release, the extracellular virions retain an "immature" morphology (Bannert et al., 2010).

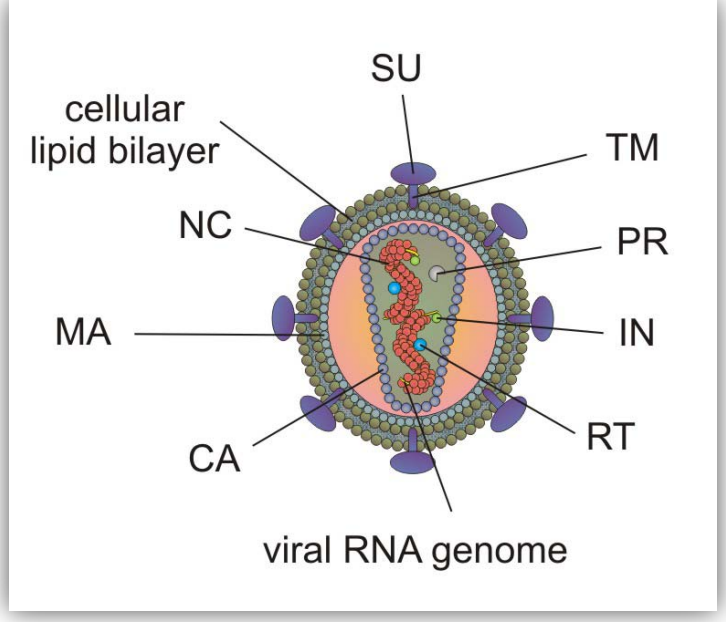


Figure 4| Schematic cartoon of a typical retroviral virion. CA, capsid; IN, integrase; MA, matrix; NC, nucleocapsid; PR, protease; RT, reverse transcriptase; SU, surface; TM, transmembrane. Figure adapted from (Vogt VM, 1997).

The viral membrane derives from the host cell membrane and is decorated with glycoprotein spikes, encoded by the *env* gene, that mediate the attachment and entry into target cells. Env is the only viral protein localized on the surface of infected cells and viral particles, and is the main determinant of viral tropism. Soon after translation from a spliced transcript, the Env precursor protein undergoes co- and post-translational modifications in the endoplasmic reticulum (ER) and Golgi complex, including processing and glycosylation. The proteolytical cleavage is mediated by cellular proteases, and gives rise to two functionally distinct proteins: the surface (SU), which contains the receptor binding domain, and transmembrane (TM), which mediates the early events of fusion between viral and host cell membrane (Bannert et al., 2010) (Fig. 4).

Finally, *pro* and *pol* encode, respectively, the protease (PR), and the reverse transcriptase (RT) and integrase (IN) enzymes (Fig. 4). As mentioned above, PR is responsible for the proteolytical cleavage of Gag, Gag-Pro and Gag-Pro-Pol polyprotein precursors. In this process, PR firstly cleaves itself out and then processes the remaining viral polyproteins. RT and IN are respectively involved in the reverse transcription and integration of vRNA into the host chromosomes. In orthoretroviruses, Pro and Pol proteins are synthesized by RNA frameshifting or termination codon read-through from the Gag-Pro and/or Gag-Pro-Pol polyprotein precursors. The first process depends on the presence of "slippery" (i.e., repeated) sequences and stable RNA secondary structures (RNA pseudoknots) that cause a change in the reading frame. Termination codon read-through is due to the misreading of a stop codon that results in the translation of the downstream gene. By contrast, in spumaviruses, Pro is translated as a Pro-Pol mRNA that is proteolytically cleaved into the IN enzyme and the RT-PR polyprotein (Bannert et al., 2010; Dunn et al., 2002).

1.6 Retroviral replication cycle

Retroviruses replicate through a complex cycle that includes a double-stranded DNA intermediate and can be divided into two overall phases, referred to as "early" and "late" steps (Freed, 2001) (Fig. 5). Although some events occur simultaneously, for simplicity they will be described here as ordered step-wise processes, including:

I. Cell entry
II. Uncoating and reverse transcription
III. Nuclear entry and integration
IV. Transcription, processing and nuclear export
V. Translation of viral proteins
VI. Genome packaging, assembly and budding
VII. Release and maturation

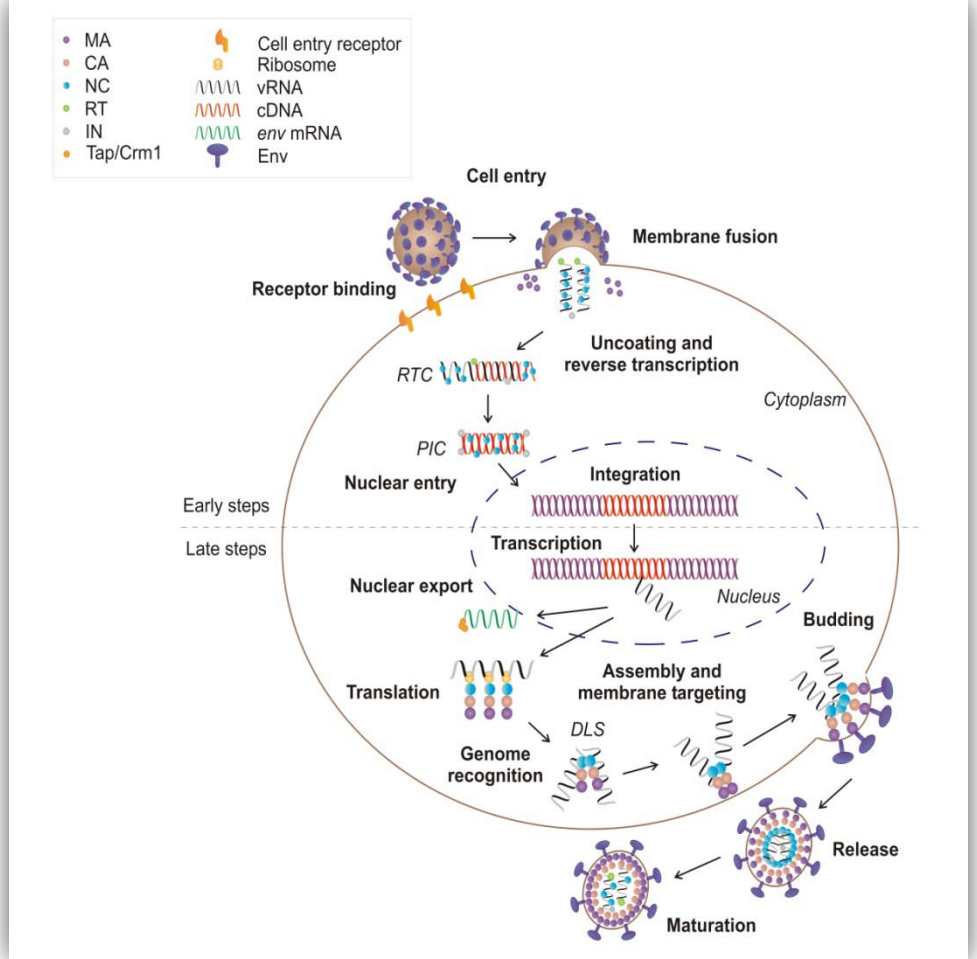


Figure 5| Schematic representation of the replication cycle of a typical orthoretrovirus. See text for details. CA, capsid; cDNA, complementary DNA; DLS, dimer-linkage structure; IN, integrase; MA, matrix; NC, nucleocapsid; PIC, pre-integration complex; RT, reverse transcriptase; RTC, reverse transcription complex; vRNA, viral genomic RNA (unspliced). Figure adapted from (D'Souza and Summers, 2005).

1.6.1 Cell entry

Retroviral entry is a multistep process that involves an initial attachment of the viral particle to the host cell, followed by receptor binding and membrane fusion (Fig. 5). Free extracellular virions attach to host target cells “brushing” the cell surface till they snag-on their cognate receptor. The initial attachment of viral particles can be facilitated by nonspecific interactions with surface proteins of the host cell membrane; however, viruses are taken-up into target cells only after binding to specific receptors on the cell surface (Mothes and Uchil, 2010) that determine viral tropism (Hunter, 1997). Engagement of the cellular receptor triggers conformational changes in the viral Env glycoprotein, which drives the fusion between viral and cellular membrane (Melikyan, 2008) (Fig. 5). In some retroviruses, such as HIV-1, membranes fusion is activated only upon the binding to a co-receptor (Zaitseva et al., 2003). Others, including ALV (Mothes et

al., 2000), MMTV (Redmond et al., 1984; Ross et al., 2002) and JSRV (Bertrand et al., 2008; Côté et al., 2008), require the presence of a cellular receptor followed by exposure to low pH in order to enter target cells. In particular, for JSRV it has been recently proposed that the binding to its entry receptor Hyal2[5] (Rai et al., 2001; Spencer et al., 2003) activates the SU domain of Env which, in turn, dissociates from the TM and inserts into the plasma membrane. Subsequently, a low PH triggers conformational changes in the TM that lead to membranes fusion (Côté et al., 2009).

1.6.2 Uncoating and reverse transcription

Upon membranes fusion, the viral core is propelled into the target cell, exposing to the cytoplasm the vRNA, the NC protein and viral enzymes. At some point after fusion and before integration into the host chromosomal DNA, viral capsids disassemble through a process referred to as "uncoating" and vRNA is reverse transcribed into cDNA (Fig. 5). The process of retroviral uncoating is poorly understood. It is believed that it depends on specific cellular factors and adenosine triphosphate (ATP) molecules (Narayan and Young, 2004), which re-organize the core leading to the formation of a reverse transcription complex (RTC), where synthesis of viral cDNA takes place (Basu et al., 2008; Telesnitsky and Goff, 1997).

In orthoretroviruses, the temporal sequence of the events immediately following viral entry is still a matter of debate and, at present, three main models have been proposed. The first one postulates that viral capsids uncoat gradually after viral entry and remain intact at least for the initiation of reverse transcription. Support for this model comes from the observation that mutant HIV capsids with altered stability are impaired in reverse transcription (Forshey et al., 2002). Conversely, the second model suggests that uncoating occurs at the nuclear pore upon completion of reverse transcription. In support of this hypothesis, there are studies showing that, in absence of viral cDNA, HIV-1 cannot translocate to the nucleus to start the process of integration into the host genome (Arhel et al., 2007). In addition, it seems that reverse transcription can efficiently occur only in intact viral capsids, as they ensure the correct stoichiometry between vRNA and viral enzymes that would otherwise dissociate in the cytoplasm. Finally, a

[5] Hyal2 stands for *hyal*uronidase *2*, a glycosylphosphatidylinositol (GPI)-anchored protein.

recent study supports a third model, whereby uncoating occurs as the reverse transcribing viral genome is transported towards the nucleus, suggesting that there may be an interplay between these two events (Hulme et al., 2011). Interestingly, spumaviruses display a "biphasic DNA synthesis" (Delelis et al., 2003), whereby early reverse transcription occurs during transport towards the nucleus, while late reverse transcription can take place before viral budding from infected cells, thus causing frequent intracellular re-integrations (Rethwilm, 2010).

During reverse transcription, the single-stranded positive vRNA molecule is converted into double-stranded viral cDNA by the retroviral RT enzyme. This is an RNA/DNA-dependent DNA polymerase with an associated RNase H activity that specifically degrades the RNA component of RNA/DNA hybrids. In MLV and MMTV, the RT enzyme functions as a monomeric polypeptide and couples both polymerase and RNase H activities; the lentiviral enzyme, by contrast, is an obligate dimer (Engelman, 2010; Herschhorn and Hizi, 2010). Reverse transcription is controlled by the NC domain of Gag, which drives critical structural rearrangements of the vRNA due to its nucleic acid condensing and chaperoning activities (Muriaux and Darlix, 2010).

Reverse transcription (Fig. 6) begins with the annealing of a cellular tRNA to complementary sequences in the viral pbs. Extension of the nascent minus-strand to the 5' end of the vRNA yields to the minus-strand strong-stop DNA, which includes the R and U5 sequences (Fig. 6A). This step is followed by a first strand transfer that allows annealing of the minus-strand strong-stop DNA to the 3' end of the vRNA (Fig. 6B). Subsequently, minus-strand DNA synthesis resumes and it is accompanied by RNase H degradation of the RNA template (Fig. 6C). The PPT resists RNA digestion and serves as a primer for the plus-strand DNA synthesis towards the 5' end of the minus-strand DNA. The plus-strand DNA is extended until the RT encounters the tRNA primer to form the plus-strand strong-stop DNA (Fig. 6D). At this point, the RNase H removes the tRNA and the PPT and a second strand transfer occurs, resulting in the annealing of the pbs of both DNA strands (Fig. 6E). These DNA strands are then further extended, yielding to a linear double-stranded DNA that contains a copy of U3-R-U5 sequence at each end, known as LTRs (Herschhorn and Hizi, 2010) (Fig. 6F).

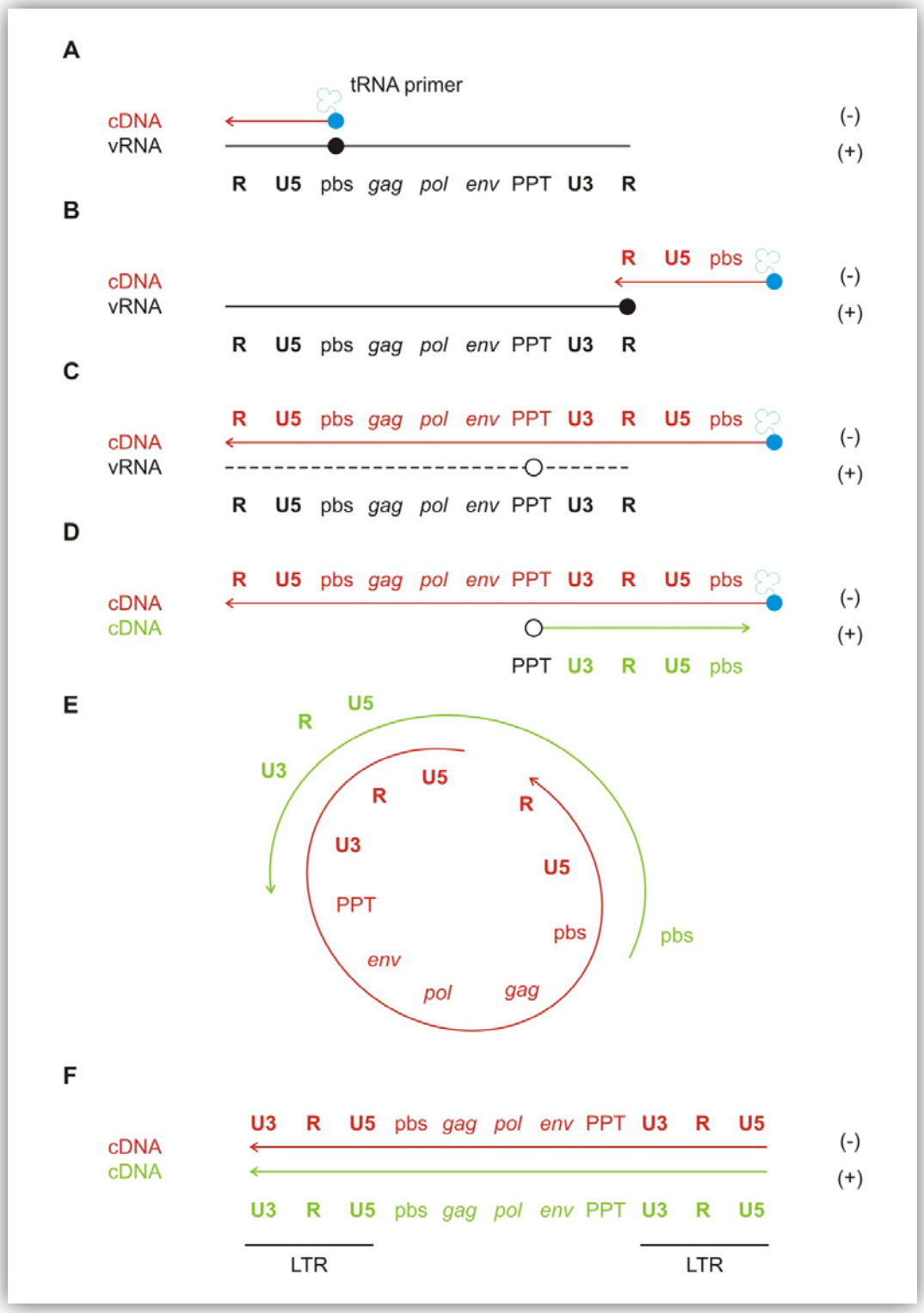


Figure 6| **Reverse transcription.** See text for details. Dashed line represents degraded vRNA by RNase H, while white and blue dots stand for primers for cDNA synthesis. In brackets, symbols (-) and (+) indicate the polarity of nucleic acid molecules. cDNA, complementary DNA; LTR, long terminal repeat; pbs, primer-binding site; PPT, polypurine tract; R, repeat sequence; U3, 3' unique sequence; U5, 5' unique sequence; vRNA, viral RNA. In black, vRNA template; in red, minus-strand cDNA; in green, plus-strand cDNA. Figure adapted from (Sarafianos et al., 2001).

Lentiviruses and spumaviruses contain an additional PPT, known as the central PPT (cPPT), which functions as a second origin of plus-strand DNA synthesis. As a result, a triple-stranded region is formed during reverse transcription in the neo-

synthesized viral genome, referred to as “central DNA flap”, which may be involved in nuclear entry (Engelman, 2010; Stevenson, 2000; Zennou et al., 2000).

1.6.3 Nuclear entry and integration

Once in the cytoplasm, the RTC starts to move towards the nucleus maturing into a pre-integration complex (PIC) (Fig. 5). This is a large nucleoprotein complex that contains the newly synthesized viral cDNA associated with viral and cellular proteins (Suzuki and Craigie, 2007). One of the key components of the PIC is the viral IN enzyme, which accomplishes the integration of the viral cDNA into the host chromosome by processing its 3' ends and mediating DNA strand transfer. During the 3' processing, IN hydrolyses each cDNA adjacent to the conserved CA/GT sequence. This reaction occurs in the cytoplasm and generates 3'-OH groups at the recessed 3' ends that are subsequently used by IN to hydrolyse the host target DNA backbone once it gains access to the nucleus (Li et al., 2011). In spumaviruses, by contrast, the 3' processing is catalyzed by the viral RT enzyme (Juretzek et al., 2004).

Retroviruses, with the exception of lentiviruses, display marked preferences for infecting dividing cells *versus* growth-arrested cells that are nonetheless metabolically active. Lentiviral PICs are actively transported into the nuclei of non-dividing primary cells and growth-arrested cell lines (Lewis et al., 1992; Weinberg et al., 1991). The molecular mechanism underlying this process is still poorly understood. Several reports demonstrate that viral (MA, Vpr, IN, CA) and cellular components (the lens-epithelium-derived growth factor LEDGF/p75, members of the importin-α protein family, tRNA species), as well as viral cDNA structures (i.e., DNA flap), may exert karyophilic activities and promote nuclear import. However, the contribution of each component to this process has not always been conclusively defined (Arhel et al., 2007; Riviere et al., 2010; Suzuki and Craigie, 2007). Other retroviruses, such as MLV (Harel et al., 1981) and JSRV (Murgia et al., 2011), by contrast, require cells to be actively cycling, in order to take advantage of the nuclear membrane dissolution during cell division, thus gaining access to the nucleus. Finally, spumaviruses RTCs display high affinity for the cellular microtubule organizing centre (MTOC), where they localize until the

infected cell is stimulated to divide; upon cell activation, viral disassembly takes place and infection proceeds (Bieniasz et al., 1995; Lehmann-Che et al., 2007).

Following entry into the nucleus, IN engages the viral cDNA as a tetramer and inserts the recessed 3' ends of the molecule into the host genome. Subsequently, cellular enzymes fill-in the resulting single-stranded gaps and 5' overhangs, which initially flank integrated viral cDNA, to finally give rise to a stably integrated provirus (Cherepanov et al., 2011).

Retroviral integration occurs basically in a random fashion throughout the whole genome. However, different retroviruses show predilection for different host DNA sites. Lentiviruses, for instance, preferentially integrate into transcription units of active genes (Schroder et al., 2002), while MLV and spumaviruses target primarily CpG[6] islands (Lewinski et al., 2006; Trobridge et al., 2006).

1.6.4 Transcription, processing and nuclear export

In the proviral phase, retroviruses rely almost entirely on the cellular machinery for gene expression. For this reason, besides encoding for a limited number of proteins, retroviral genomes also contain *cis*-acting sequences, in the viral LTRs, to maximize gene expression by host cell machinery. In particular, the U3 region of the 5'LTR contains *cis*-acting control elements that regulate transcriptional initiation by the cellular RNA polymerase II (Rabson and Graves, 1997). Complex retroviruses, such as HIV, HTLV and spumaviruses, encode their own transcriptional activators (Tat, Tax and Tas, respectively) that enhance viral expression by acting in concert with cellular factors (Nicot et al., 2005; Purcell and Martin, 1993). The U3, R, and U5 regions of the 3'LTR, on the other hand, include *cis*-acting control elements involved in the post-transcriptional processing that occurs at the 3' end of viral RNA transcripts (Rabson and Graves, 1997).

Retroviral transcription generates two mRNAs species, one that undergoes RNA splicing (single or multiple) and the other that remains unspliced and can be translated and/or packaged as vRNA into assembling virions (Fig. 3). In

[6] CpG stands for “cytosine-phosphate-guanine”, to distinguish this linear sequence from cytosine-guanine DNA base pairing. CpG islands are short interspersed DNA sequences, which usually function as sites of transcription initiation. In vertebrates, they can undergo dense methylation, which results in gene silencing. CpG islands can therefore influence local chromatin structure, and regulate gene expression (Deaton and Bird, 2011).

orthoretroviruses, Gag, Gag-Pro and Gag-Pro-Pol polyprotein precursors derive from unspliced transcripts (Goff, 2007), whereas in spumaviruses Pol is synthesized from a separate mRNA (Enssle et al., 1996). The Env precursor derives from a single RNA splicing event that joins sequences upstream of the donor splice site in the 5'UTR and downstream of the acceptor splice site in *pol*, to eliminate *gag*, *pro* and *pol* genes (Swanstrom and Wills, 1997) (Fig. 3). Finally, complex retroviruses such as HIV-1, MMTV and HTLV, possess a third class of mRNAs that is generated by multiple splicing events and encode regulatory and accessory proteins (Bolinger and Boris-Lawrie, 2009).

Retroviral RNA splicing depends on cellular factors, which normally retain intron-containing transcripts in the nucleus. To overcome nuclear sequestration, retroviruses have evolved two elegant strategies to export their intron-containing mRNAs in the cytoplasm: the Crm1[7] and Tap[8] pathways (Cochrane et al., 2006) (Fig. 5). Simple retroviruses, such as M-PMV (Bray et al., 1994) and RSV (Ogert et al., 1996), recruit the Tap pathway, which is normally used for nuclear export of cellular mRNAs. Other viruses, such as influenza, hepatitis B and herpes simplex, were shown to employ the same nuclear export mechanism (Fontoura et al., 2005). This pathway involves a constitutive transport element (CTE) located in the 3'UTR of the viral RNA that interacts specifically with the cellular heterodimer Tap/p15 and mediates nuclear export (Li et al., 2006; Wiegand et al., 2002).

All known lentiviruses, as well as the human oncoretrovirus HTLV, employ the cellular protein Crm1 to export their unspliced and incompletely spliced RNAs in the cytoplasm of infected cells (Pollard and Malim, 1998). Crm1 is a member of the karyopherin β family of transport receptors that normally mediates nuclear export of small nuclear RNAs (snRNAs), ribosomal RNAs (rRNAs) and a small subset of cellular mRNAs (Weis, 2003). The Crm1 pathway is typified by the Rev transactivator protein of HIV-1, which is translated from the fully spliced viral mRNA and is thereby constitutively exported in the cytoplasm. Rev selectively

[7] Crm1 (*c*hromosome *r*egion *m*aintenance *1*) was firstly described in yeast, where its mutation causes abnormal chromosome morphology (Adachi and Yanagida, 1989). A more recent study revealed that Crm1 is an export receptor for leucine-rich nuclear export signals (Fornerod et al., 1997).

[8] Tap stands for Tip-*a*ssociated *p*rotein. Tap was initially identified as a cellular protein interacting with the tyrosine kinase-interacting protein (TIP) of herpesvirus saimiri, and inducing lymphocyte aggregation (Yoon et al., 1997). Experiments conducted in *Xenopus* oocytes established that Tap is the major cellular receptor for nuclear export of bulk mRNA (Bachi et al., 2000).

interacts with a *cis*-acting RNA stem-loop structure, the Rev-responsive element (RRE), located at the 3' end of unspliced and incompletely spliced viral transcripts (Felber et al., 1989; Malim et al., 1989). Multiple domains are required for Rev to function, including an arginine-rich motif (ARM) and a leucine-rich sequence, situated, respectively, at the N-terminal and C-terminal portions of the protein. The ARM domain contains a nuclear/nucleolar localization signal (NLS/NoLS) and binds to the major groove of the RRE. The leucine-rich domain harbours a nuclear export signal (NES) that binds to the cellular protein Crm1, in a Ran-GTP manner[9]. Upon interaction with components of the nuclear pore, the Rev-RRE/Crm1/Ran-GTP complex reaches the cytoplasm, where Ran-GTP is converted into Ran-GDP and the ternary complex disassembles (Askjaer et al., 1998; Fornerod et al., 1997). Subsequently, the NLS interacts with the transport mediator importin β and Rev enters the nucleus again (Fontoura et al., 2005).

Interestingly, recent studies indicate that some retroviruses employ other export pathways to circumvent the nuclear retention of their unspliced transcripts that seem to share common features to both Crm1 and Tap mechanisms. Along this line, it has been recently reported that spumaviruses utilize a Crm1-dependent pathway that does not require any *trans*-acting viral regulatory protein, but putative CTE-like RNA secondary structures that respond to cellular HuR proteins (Bodem et al., 2011). Another study suggests that export of unspliced spumaviruses RNAs may rely on an additional and complementary mechanism based on Gag-RNA interactions (Renault et al., 2011). In MLV infected cells, by contrast, nuclear export of unspliced vRNAs depends on the presence of CTE-like stem-loop structures within the ψ sequence (Basyuk et al., 2005). Finally, MMTV (Byun et al., 2010) and JSRV (Caporale et al., 2009; Hofacre et al., 2009; Nitta et al., 2009) recruit a Crm1-dependent pathway that functions in presence of the signal peptide (SP) of their Env glycoproteins.

1.6.4.1 Retroviral signal peptides and export of viral unspliced RNAs

In eukaryotes, SPs target proteins to the ER membrane, which is the entry site into the secretory pathways, and mediate insertion into the translocon for transport through the lipid bilayer (Paetzel et al., 2002). SPs are generally

[9] Ran (Ras-related nuclear protein) is a member of the Ras GTPases superfamily (Quimby and Dasso, 2003). GTP and GDP stand for guanosine tri- and di-phosphate, respectively.

located at the N-terminal protein extension, but they can also lie within the peptide or at its C-terminal end (Kutay et al., 1995). Despite a great diversity in terms of sequences, the function of SPs has been maintained throughout evolution, as evidenced by the fact that they can be interchanged between different proteins, and even between proteins of different organisms, without being affected (Gierasch, 1989; Izard and Kendall, 1994).

Although SPs are extremely heterogeneous in terms of length and amino acid charge, they share a number of conserved features that are essential for protein export. These include a core region of seven to fifteen hydrophobic residues (h), required for targeting and membrane insertion, a central region (c) of two to nine small and polar residues, which determines the site of SP cleavage, and a polar region (n), usually positively charged. The n-region contributes most to the variation in the overall length of SPs, which can range from fifteen to up to fifty amino acid residues (von Heijne, 1985; von Heijne, 1986; von Heijne, 1990).

Upon insertion into the ER membrane in a loop-like configuration, SPs can be cleaved-off by signal peptidases (SPases) on the membrane lumen. A number of features of the SP sequence influences the outcome of SPase cleavage, such as the steric hindrance and charge of amino acid residues located within the cleavage site, as well as the length of the h-region and properties of the n-region. After being cleaved and released from the rest of the protein, SPs can undergo proteasomal degradation or be further processed by signal peptide peptidases (SPPase) (Paetzel et al., 2002).

New studies revealed that SPs display additional and important functions beyond membrane targeting. Along this line, it has been shown that eukaryotic SPs can exert important roles in cellular processes, such as signal-transduction pathways (e.g., targeting of calmodulin by the pre-prolactin SP) and cellular immune response (e.g., activation of natural killer cells by SP fragments of major histocompatibility complex class I molecules) (Paetzel et al., 2002). However, most examples of post-targeting functions come from viral SPs, including FVs where it has been proposed that it plays a key role during viral budding and particle release and may also determine viral infectivity (Cartellieri et al., 2005; Lindemann et al., 2001; Shaw et al., 2003; Stanke et al., 2005; Wilk et al., 2001). Conversely, in MMTV (Byun et al., 2010) and JSRV (Caporale et al., 2009;

Hofacre et al., 2009; Nitta et al., 2009), SPs are involved in the export of unspliced viral RNAs.

1.6.4.2 MMTV and the export of unspliced viral RNAs

Besides the four basic retroviral genes, MMTV encodes also a regulatory protein (Rem) that derives from a double splicing event of the viral mRNA (Indik et al., 2005; Mertz et al., 2005). At the N-terminus, both Rem and Env contain a SP sequence (MMTV SP) that exhibits HIV Rev-like activity. MMTV SP possesses several Rev-like motifs that are important to exert its function, including NLS, ARM and NES domains and the Rem-responsive element (RmRE) located at the 3' end of the full-length viral RNAs (Byun et al., 2010). Upon targeting Rem and Env to the ER, MMTV SP is cleaved-off by the SPase, retrotranslocates to the cytoplasm and enters the nucleus (Byun et al., 2010; Dultz et al., 2008). Here, it binds to the full-length viral RNAs through the RmRE (Mertz et al., 2009) and export them to the cytoplasm (Byun et al., 2010).

1.6.4.3 JSRV and the export of unspliced viral RNAs

Recent results obtained by Marco Caporale and colleagues in our laboratory shown that the N-terminal portion of JSRV Env glycoprotein contains a SP sequence (JSE-SP) with a HIV Rev-like activity (Fig. 7). Similarly to Rev, JSE-SP function depends on multiple domains, including NLS/NoLS, ARM and NES. Following cleavage from the Env polyprotein at the ER membrane, JSE-SP enters the nucleus, where is targeted to the nucleolus by the NLS/NoLS domain. Evidences suggest that nuclear export is achieved through the binding of the ARM domain of JSE-SP to the signal peptide responsive element (SPRE), a stem-loop RNA secondary structure located at the 3' end of viral unspliced RNAs. JSE-SP acts as a post-transcriptional regulator of viral gene expression, as it enhances Gag synthesis and viral particle release (Caporale et al., 2009). Experiments conducted by Hofacre, Nitta and colleagues confirmed the presence of a *trans* acting factor, which was termed Rej (for regulator of JSRV expression), located at the N-terminus of JSRV Env glycoprotein and that specifically interacts with its responsive element Rej-RE (for Rej-responsive element) (Hofacre et al., 2009; Nitta et al., 2009).

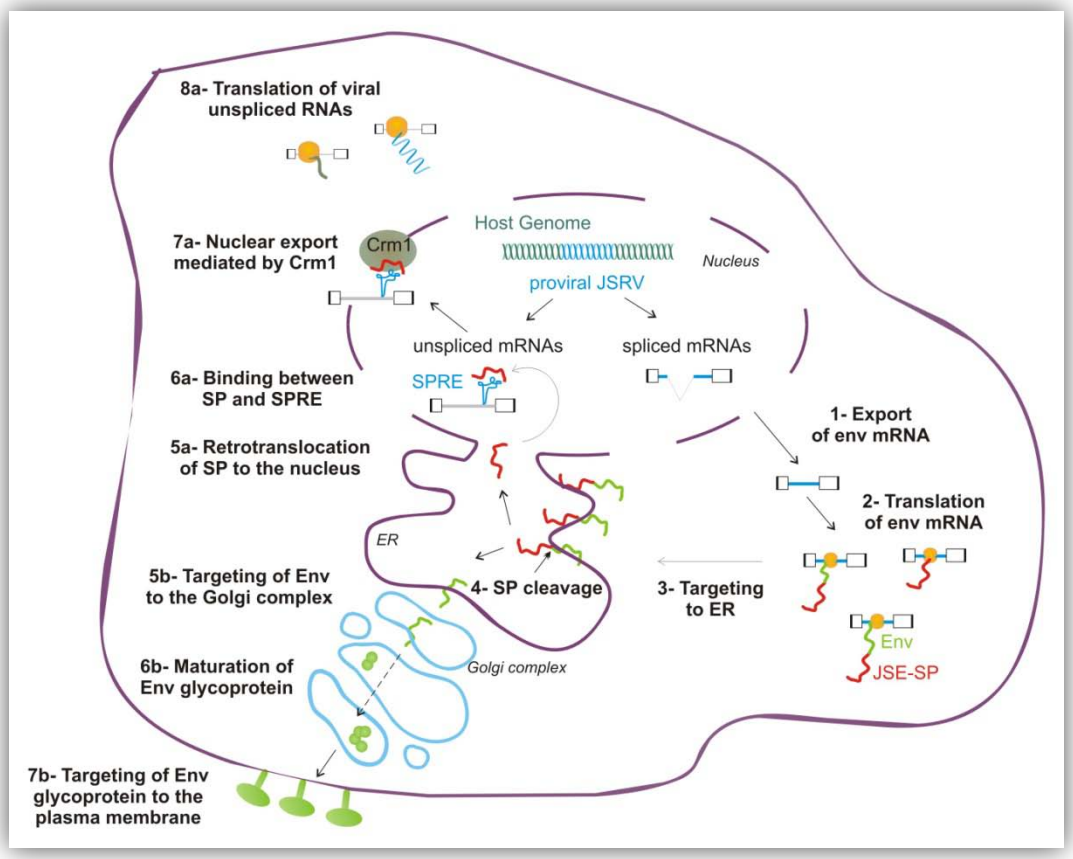


Figure 7| Model of JSE-SP trafficking. See text for details. ER, endoplasmic reticulum; SP, signal peptide; SPRE, signal peptide-responsive element.

1.6.5 Translation of viral proteins

Following nuclear export, retroviral proteins are translated on cellular ribosomes by the host cell translational machinery in a cap-dependent mechanism. However, in order to optimize synthesis of viral mRNAs throughout the replication cycle, retroviruses have evolved alternative initiation strategies (Bolinger and Boris-Lawrie, 2009). Some retroviruses, such as lentiviruses (Berkhout et al., 2011; Herbreteau et al., 2005; Ohlmann et al., 2000), RSV (Deffaud and Darlix, 2000b) and MLV (Berlioz and Darlix, 1995; Deffaud and Darlix, 2000a; Vagner et al., 1995), make use of internal ribosome entry site (IRES) elements (on both 5'UTR and coding regions) to regulate viral expression. IRES are defined as RNA domains where ribosomes are recruited in a cap-independent manner, thus allowing mRNAs translation even during stress conditions that normally inhibit the host cell translation machinery (e.g., hypoxia, apoptosis, angiogenesis). Additional mechanisms include viral proteins that directly bind to secondary structures of the vRNA (e.g., HIV-1 Tat/TAR and HIV Rev/RRE interaction) or that modify the host cell translational apparatus

(e.g., cleavage of eukaryotic initiation translation factor by retroviral proteases), thus overall favouring translation of viral transcripts (Balvay et al., 2007).

A second important strategy adopted by retroviruses to increase efficiency of viral translation is to synthesize different proteins from a single unspliced RNA molecule. For example, in orthoretroviruses, Pro and Pol are produced from the same Gag-Pro and/or Gag-Pro-Pol polyprotein precursors by RNA frameshifting (e.g., HIV-1, RSV and HTLV) or termination codon read-through (e.g., MLV and Feline leukaemia virus, FeLV) (Swanstrom and Wills, 1997).

1.6.6 Genome packaging, assembly and budding

Retroviral assembly is a multistep process finely orchestrated by the Gag polyprotein, whose various domains play distinct roles during viral morphogenesis. Within the context of the unprocessed polyprotein, MA is primarily required for Gag membrane targeting and attachment, CA represents the driving force for the assembly of immature particles, while NC mediates the packaging of vRNAs into forming virions (D'Souza and Summers, 2005). In spumaviruses, by contrast, the Gag polyprotein is not processed in the canonical MA, CA and NC domains, and it seems that both Gag and Pol proteins take part to viral assembly (Lindemann and Rethwilm, 2011).

The process starts with the encapsidation of vRNAs into assembling particles. A unique feature of the vRNA is that it dimerizes with nearby vRNA molecules *via* "kissing" interactions[10] between palindromic stem loops, leading to the formation of a dimer-linkage structure (DLS) (Muriaux and Darlix, 2010). Retroviral dimeric vRNAs are selectively recruited, from the large pool of cellular RNA molecules, by high affinity interactions between the NC domain of Gag and the ψ sequence located within the 5'UTR and the 5' end of *gag* (Johnson and Telesnitsky, 2010) (Fig. 3). Subsequent steps in retroviral assembly include the oligomerization of a few thousand Gag molecules, through their interaction (I) domains around the nucleating core formed by Gag-RNA complex. This process is driven by the CA and NC domains of the Gag polyprotein (D'Souza and Summers, 2005).

[10] "Kissing" interactions between two molecules of vRNAs are non-covalent interactions (D'Souza and Summers, 2005).

Retroviruses have been shown to recruit Gag molecules to two different sites for productive viral assembly:

i. Betaretroviruses (B-type and D-type retroviruses) assemble in the cytoplasm, by forming intracytoplasmic "A-type" particles that are subsequently transported to the plasma membrane, where Env is incorporated during budding (Swanstrom and Wills, 1997). This is also the mechanism adopted by spumaviruses (Yu et al., 2006).
ii. Alpharetroviruses, gammaretroviruses and lentiviruses (C-type retroviruses), on the other hand, recruit Gag polyproteins at the plasma membrane, where formation of immature viral particle occurs (Swanstrom and Wills, 1997).

Interestingly, a single amino acid substitution in the MA domain of Gag transforms M-PMV (i.e., the prototype of D-type retroviruses) to a C-type retrovirus (Rhee and Hunter, 1990), suggesting that these two pathways may share common mechanistic processes.

Plasma membrane targeting is mediated by the MA domain of Gag through the N-terminal myristoyl group and a cluster of basic amino acids (the so-called M domain) that confer specificity to the host cell membrane (Fig. 5) (Hamard-Peron and Muriaux, 2011). During this process, MA also serves as a specific interface between Gag and membrane microdomains, including lipid rafts[11] enriched with phosphatidylinositol 4,5-bisphosphate [PI(4,5)P2] and cholesterol. Two models, not mutually exclusives, have been postulated to explain the association between Gag and lipid rafts. In the first one, the interaction between the M domains of Gag and PI(4,5)P2 may induce exposure of the N-terminal myristoyl group, as well as sequestration of other PI(4,5)P2 chains; these N-terminal myristate groups-PI(4,5)P2 complexes, in turn, may promote Gag multimerization. Alternatively, multiple PI(4,5)P2 chains, associated with a Gag cluster, may induce formation of a stable raft-like domain that, in turn, may target Gag to the cell membrane (Hamard-Peron and Muriaux, 2011; Ono, 2010).

[11] From (Pike, 2006): «*Lipid rafts are small, heterogeneous, highly dynamic, sterol- and sphingolipid-enriched domains that compartmentalize cellular processes. Small rafts can sometimes be stabilized to form larger platforms through protein–protein and protein–lipid interactions*».

Several studies demonstrate that Gag possesses late (L) domains that actively promote separation of the lipid envelope of the nascent virion from the cell surface, by engaging components of cellular budding pathways (Freed, 2004). These pathways are topologically similar to the membrane fission events that occur during cellular cytokinesis (Carlton and Martin-Serrano, 2007). The L domains directly bind to class E vacuolar proteins sorting (Vps), which participate in the biogenesis of three distinct endosomal sorting complexes required for transport (ESCRT) (Göttlinger and Weissenhorn, 2010). To date, three L domains have been characterized: (i) the P(T/S)AP-type, which binds to the Tsg101 protein, (ii) the $LYPx_nL$-type, which mediates the interaction between Tsg101 and its partner ALIX, and (iii) the PPxY-type, which promotes binding of proteins containing tryptophan-tryptophan (WW) domains to Nedd4-related E3 ubiquitin ligases (Göttlinger and Weissenhorn, 2010).

Orthoretroviruses and spumaretroviruses differ also in this step of the viral replication cycle. Indeed, budding and release by spumaviruses strictly depends on Gag-Env interactions, while orthoretroviruses can release viral particles even in absence of Env (Baldwin and Linial, 1998; Pietschmann et al., 1999), although the latter greatly facilitate intracellular Gag trafficking and viral particle release (Arnaud et al., 2007b; Sfakianos and Hunter, 2003). In addition, it has been demonstrated that the late steps of FV replication cycle depend on the presence of the signal peptide (or leader peptide) of the viral Env protein (Elp)[12] (Cartellieri et al., 2005; Lindemann et al., 2001; Wilk et al., 2001). Elp lies at the N-terminal portion of the Env precursor protein (pre-gp130Env), which it is processed by furin or furin-like protease during its intracellular transport towards the cell membrane (Duda et al., 2004; Geiselhart et al., 2004; Lindemann et al., 2001; Netzer et al., 1990; Stange et al., 2008; Wang and Mulligan, 1999). It has been shown that Elp mediates particle egress by specifically interacting with the budding domain in gp130Env and the N-terminus of Gag (Cartellieri et al., 2005; Lindemann et al., 2001; Wilk et al., 2001). Moreover, it has been suggested that ubiquitination of Elp may regulate the

[12] The SP of FV Env has been termed SP_{ENV} (Lindemann et al., 2001), Elp (for *E*nv *l*eader *p*rotein) (Geiselhart et al., 2004; Wilk et al., 2001) and LP (for *l*eader *p*eptide) (Duda et al., 2004). Here, for simplicity, it is used the nomenclature most commonly and recently adopted, i.e., Elp.

egress of viral and subviral particles[13] from infected cells, thus determining viral infectivity (Stanke et al., 2005).

1.6.7 Release and maturation

Finally, when all the viral components are recruited together, the cell membrane bends until the virion pinches-off and is released in the supernatant (Fig. 5). Upon release from the host cell surface, orthoretroviruses undergo an obligatory maturation step that yields infectious virions. This process is mediated by the viral PR enzyme, which promotes proteolytic cleavages of Gag and Gag-Pro-Pol precursors, thus leading to major structural changes (Fig. 5). In spumaretroviruses, by contrast, Gag is not proteolytically processed in NC, CA and MA domains, and thus viral particles retain an "immature" morphology (Göttlinger and Weissenhorn, 2010; Hartl et al., 2011).

1.7 Interplay between retroviruses and their hosts: restriction factors

Host–pathogen interaction is a classic example of genetic conflict, in which both entities try to gain evolutionary advantages over the other. As illustrated above, retroviruses rely completely upon host factors to successfully complete their replication cycle. Hence, over time, they have developed different strategies to hijack the cell machinery and optimize viral replication. Accordingly, hosts have evolved many defence mechanisms to prevent viral transmission, including intrinsic, innate and adaptive immune responses, which are, respectively, the first and second lines of defence against spreading of infectious pathogens. The driving force of host-pathogen interaction is therefore the genetic plasticity that, by continuously demanding functional "innovations", points to different layers of complexity and renders this interplay evolutionary dynamic.

Innate immune response is acquired at birth and changes little throughout the life of an individual. In general, it requires sentinel and cytolytic cells (i.e., macrophages and natural killer cells), as well as the secretion of cytokines (including interferon, IFN) and an array of serum proteins known as complement. If this strategy fails, the adaptive immune response is mobilized. This includes

[13] Subviral particles are defined as non-infectious and capsidless particles that contain only the viral Env glycoprotein (Stanke et al., 2005).

activated lymphocytes (B and T cells) and the synthesis and release of epitope-specific antibodies. Although this mechanism of defence takes longer to develop, it is far more specific and, most importantly, shows memory (Lydyard et al., 2004).

Besides the innate and adaptive immune responses, vertebrates have developed an «*intrinsic immunity*» (Bieniasz, 2004), which consists in a cell-autonomous defence against viral infections. This mechanism of response relies on cellular proteins known as «*restriction factors*», which specifically target different steps of the viral replication cycle, thus potently inhibiting spreading of infection. Restriction factors have been conserved across the speciation of mammals (Luban, 2010), suggesting that they might have co-evolved with retroviruses throughout mammalian orders. They include IFN-induced proteins, such as APOBECs, tripartite motif (TRIM) proteins and bone marrow stromal cell antigen 2 (BST-2), and the zinc finger antiviral protein (ZAP). Retroviruses, in turn, have evolved a number of mechanisms to avoid and/or modify the host immune response and establish successful infections. For example, lentiviruses and spumaviruses encode accessory proteins to overcome cellular restriction (Löchelt et al., 2005; Strebel et al., 2009), whereas MMTV uses cells of the host immune system for viral spread (Golovkina et al., 1998).

1.7.1 TRIM5α

In 2004, Matthew Stremlau and colleagues observed that Old World monkeys'[14] cells were resistant to HIV-1 infections (Stremlau et al., 2004). The factor responsible for this restriction activity was identified as TRIM5α, a protein that is able to block the early steps of retroviral replication by preventing the process of reverse transcription (Fig. 8). The underlying mechanism adopted by TRIM5α to exert its block remains still poorly understood, even though it is now becoming increasingly clear that its restriction activity may occur through multistep and redundant pathways (Anderson et al., 2006; Luban, 2010; Towers, 2007). One possibility is that TRIM5α forms complexes with retroviral cores that are rapidly degraded by the proteasomal machinery in a ubiquitin-mediated mechanism (Wu et al., 2006). Alternatively, TRIM5α may rapidly uncoat incoming

[14] Examples of Old World monkeys (order Primates, parvorder Catarrhini) are macaques, baboons and mandrills.

HIV-1 capsids, thereby causing premature disassembly of viral cores (Stremlau et al., 2006). Finally, it has been proposed that TRIM5α exerts its block by impairing nuclear entry to PICs, thus preventing integration into the host genome (Anderson et al., 2006).

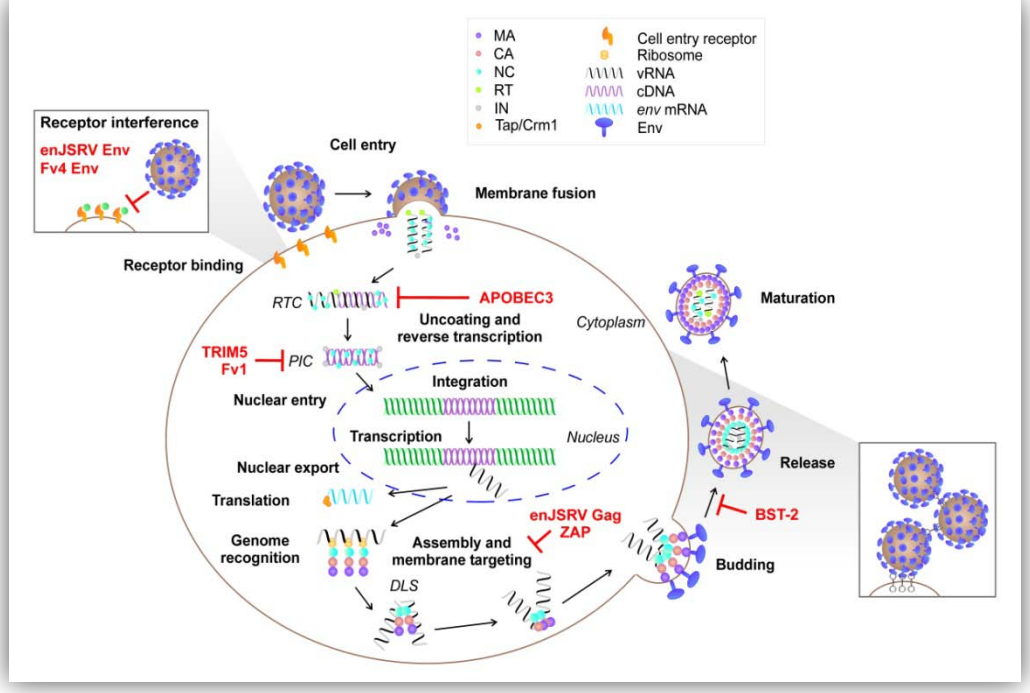


Figure 8| Mechanism of action of host restriction factors. Cellular restriction factors follow different strategies and act at different stages of the retroviral replication cycle. See text for details. APOBEC, apolipoprotein B mRNA-editing enzyme; BST-2, bone marrow stromal cell antigen 2; CA, capsid; cDNA, complementary DNA; DLS, dimer-linkage structure; enJSRV, endogenous Jaagsiekte sheep retrovirus; Fv1, Friend-virus-susceptibility-1; Fv4, Friend-virus-susceptibility-4; IN, integrase; MA, matrix; NC, nucleocapsid; PIC, pre-integration complex; RT, reverse transcriptase; RTC, reverse transcription complex; TRIM, tripartite motif; vRNA, viral genomic RNA (unspliced); ZAP, Zinc-finger antiviral protein. Figure adapted from (D'Souza and Summers, 2005) and (Jern and Coffin, 2008).

TRIM5α is one of the multiple protein isoforms encoded by *TRIM5* family genes, defined by the presence of three zinc-binding domains: an amino-terminal ring finger that functions as E3 ubiquitin ligase, one or two B-box domains and a coiled-coil region known as B30.2/SPRY domain. The two first domains are required for efficient restriction activity, whereas the latter enhances binding avidity for the viral capsid by promoting TRIM5α oligomerization. In particular, the B30.2/SPRY domain exhibits species-specific restriction activity to HIV-1. Replacement of the amino-terminal residues in the B30.2/SPRY domain of human TRIM5α (TRIM5α_{hu}) with the corresponding amino acids of rhesus TRIM5α (TRIM5α_{rh}) confers to the resulting TRIM5α_{hu} mutant protein the ability to block HIV-1, with an efficiency comparable to that of the wild-type TRIM5α_{rh} (Stremlau

et al., 2005). The importance of this domain is further emphasized by the fact that these residues have been positively selected during primate evolution (Sawyer et al., 2005; Song et al., 2005), suggesting that TRIM5α has been co-evolving with retroviruses for millions of years.

The *TRIM* family arose through gene duplication events. The human genome, for instance, contains over seventy *TRIM* family genes, including an active *TRIM5* gene and a *TRIM5* pseudogene. Multiple copies of *TRIM5* genes are also found in mouse, rat and cow genomes, whereas in the dog genome there is only one *TRIM5* gene present and this has been interrupted by gene insertion and has therefore decayed into a pseudogene (Johnson and Sawyer, 2009). This pattern of gene duplication and/or loss indicates highly dynamicity of the *TRIM5* loci, which might have contributed to the "endogenous" restriction of retroviruses throughout mammals evolution. This notion is further supported by recent findings that *TRIM5* is involved in promoting innate immune signalling pathways (Pertel et al., 2011). Moreover, it is noteworthy that a number of TRIM proteins have been found to block different steps of the retroviral replication cycle (Uchil et al., 2008).

1.7.2 CypA and TRIM5-CypA

CypA is a cellular peptidyl-prolyl isomerase that is incorporated in about 10% of HIV-1 virions through interactions with the CA domain of the Gag polyprotein (Franke et al., 1994; Thali et al., 1994). Most likely, CypA has been selectively co-opted by HIV-1 as it promotes viral infectivity by preventing TRIM5α_{hu}-CA interaction (Sokolskaja et al., 2006). CypA is required in target but not in producer cells, since mutational inactivation of CypA binding site, or inhibition of the CA-CypA interaction by cyclosporin A (CsA), inhibits or attenuates HIV-1 replication (Braaten and Luban, 2001). A similar block to HIV-1 infections is exerted by TRIM5α_{rh} in Old World monkey cells (Berthoux et al., 2005), suggesting that this step may be a point of vulnerability in the retroviral cycle.

Some species belonging to *Aotus* and *Macaca* genera (New World monkey[15]) express a TRIM5-CypA fusion protein that derives from the replacement of the TRIM5α B30.2/SPRY domain with a CypA pseudogene (Johnson and Sawyer, 2009;

[15] Examples of New World monkeys (order Primates, parvorder Platyrrhini) are marmosets, squirrel monkeys and owl monkeys.

Nisole et al., 2004; Sayah et al., 2004). Some studies speculate that TRIM5-CypA arose after the divergence of New and Old World monkeys, when a LINE1 catalysed the insertion of a CypA complementary DNA into the *TRIM5* locus (Patthy, 1999; Towers, 2007). TRIM-CypA confers resistance to HIV-1 infection (Hofmann et al., 1999) and might have provided significant selective advantages to these animal species. It is worth noting that a human TRIM5-CypA-like gene would be a useful antiviral against HIV infections.

1.7.3 APOBEC

APOBEC proteins belong to a large family of cytidine deaminases, unique to mammals, involved in a variety of processes, including lipid metabolism, antibody diversification and inhibition of the replication of retroviruses and some DNA viruses. These enzymes replace the cytidine bases on single-stranded DNA or RNA uracil bases, thereby acting as DNA- and RNA-editing enzymes. In mammals, such substitutions may alter the amino acid sequence of important proteins, thus causing a variety of diseases, including cancers and immune disorders (Ross, 2009). In retroviruses, APOBEC3 proteins are packaged into newly forming virions and induce mutations in the minus strand of viral cDNA, thus greatly reducing viral replication efficiency (Fig. 8). It has been shown that APOBEC3 proteins can inhibit viral replication in either a cytidine deamination-dependent or -independent manner, but the precise mechanism underlying these processes remains still to be defined (Holmes et al., 2007).

Retroviruses have evolved a variety of strategies in order to counteract the effects of cellular APOBEC3 proteins and prevent their packaging into newly forming virions. These comprise (i) proteasome-mediated degradation of APOBEC3G and APOBEC3F by the viral Vif protein in HIV-1 (Mehle et al., 2004; Yu et al., 2003; Sheehy et al., 2002), and (ii) inhibition of the interaction between APOBEC3G and vRNA–NC complexes in HTLV-1 (Derse et al., 2007). Spumaviruses evade APOBEC3 antiviral activity by a distinct strategy that is mediated by the accessory protein Bet. Although the biochemical details of this mechanism have not been completely elucidated, several lines of evidence suggest that Bet forms complexes with APOBEC3C proteins that are sequestered in subcellular compartments and undergo proteosomal degradation (Löchelt et al., 2005; Perkovic et al., 2009).

APOBEC3 genes arose through gene duplication of a single-copy primordial gene, and their copy number varies from one in mice, two to three in artiodactyls species and up to seven in humans (Conticello et al., 2005; LaRue et al., 2008). A growing body of evidence indicates that the primary function of these genes in mammals is to provide an innate immune defence to retrovirus and retrotransposon mobilization (Bogerd et al., 2006; Chen et al., 2006; Esnault et al., 2006; Esnault et al., 2008; Ikeda et al., 2011). This hypothesis is supported by the fact that primate mammals exhibit a larger repertoire of *APOBEC3* genes than non-primate mammals, probably due to the higher exposure of the first to potentially invasive retroelements (Conticello et al., 2005). It is noteworthy, however, that the entire human *APOBEC3* locus is flanked by repetitive elements (mainly LTRs from ERVs class I), suggesting that such amplification and recombination of *APOBEC3* genes in primates might have been facilitated by retroviral elements (Conticello et al., 2005). Thus, it appears that host genomes have co-opted retroviral elements that fulfil key functions for the host biology, while host/virus interplay might have provided the driving force for the rapid expansion of the entire *APOBEC3* locus and other restriction factors in primates.

1.7.4 ZAP

ZAP is a host antiviral protein that inhibits the late stages of replication of MLV (Gao et al., 2002), alphaviruses (Bick et al., 2003) and filoviruses (Müller et al., 2007). Putative orthologs of *ZAP* gene can be found in the genomes of a variety of mammals, including rat (where it was firstly identified) (Gao et al., 2002), as well as chicken and fish (Amé et al., 2004), suggesting that this gene family dates back to the origin of vertebrates. Viral recognition by ZAP is mediated by four CCCH-type zinc-finger motifs, which specifically bind to target viral mRNAs and mediate exosome-dependent RNA degradation (Gao et al., 2002; Guo et al., 2007) (Fig. 8). Surprisingly, phylogenetic analyses found no evidence of positive selection of this domain throughout primates evolution, arguing that the antiviral activity of ZAP was not shaped by rapid alterations in the viral RNA binding sites (Kerns et al., 2008). In primates and rat, *ZAP* encodes two isoforms that derive from alternative RNA splicing events. In primates, the longer isoform ZAP(L) contains a poly(ADP-ribose) polymerase (PARP)-like domain, which is absent in the shorter protein ZAP(S). In rat, by contrast, neither isoforms possess a PARP-like domain (Gao et al., 2002; Kerns et al., 2008). Recent studies

revealed that the PARP-like domain has been positively selected throughout primates evolution, even though the mechanisms by which this region mediates ZAP antiviral activity have not yet been characterized (Kerns et al., 2008). However, it is possible to speculate that the PARP-like domain may increase the affinity of ZAP for viral RNAs or, alternatively, counteract putative viral factors involved in ZAP degradation (Kerns et al., 2008).

1.7.5 BST-2

BST-2, also known as CD317, HM1.24 or tetherin, is a transmembrane protein that restricts many enveloped viruses, including retroviruses, filoviruses and herpesviruses. BST-2 is expressed constitutively in several cell types, such as B cells, T cells, macrophages and cancer cell lines and it can be induced by type I IFN in response to viral infections (Evans et al., 2010; Neil et al., 2008; Van Damme et al., 2008).

BST-2 localizes within several endosomal compartments, including *trans*-Golgi network (TGN) and recycling endosomes. In addition, it is present at the plasma membrane where it tethers fully formed virions to the cell surface, thereby preventing their release (Fig. 8). Virtually, any lipid-enveloped virus could be restricted by BST-2. However, because the antiviral activity of BST-2 most likely requires co-localization with budding virions, its restriction is probably limited to viruses that bud from cellular membranes enriched in BST-2 (Habermann et al., 2010).

It has been proposed that BST-2 exerts its antiviral activity without cellular cofactors. This “direct tethering” model is strongly supported by an elegant experiment that demonstrated that an artificial protein, lacking significant sequence homology to the native BST-2 but retaining its topology, was still able to mimic its antiviral activity (Perez-Caballero et al., 2009). Key domains for BST-2 restriction are the N-terminal transmembrane (TM) portion, which faces the cytoplasm, and the glycosylphosphatidylinositol (GPI) membrane anchor at the C-terminal portion. This unusual topology enables BST-2 to interact with the nascent virions and tether them to the cell surface (Perez-Caballero et al., 2009). In addition, recent studies revealed that BST-2 is incorporated into nascent infectious virions (Fitzpatrick et al., 2010). The exact mechanism underlying BST-2 incorporation and virion tethering has not yet been elucidated.

Several lines of evidences favour the "membrane spanning" model, whereby BST-2 is incorporated into virions as a dimer, whose TM domains are inserted into the virion envelope, while the GPI anchors remain embedded in the host-cell membrane (Perez-Caballero et al., 2009). Retained virions can be internalized and subsequently degraded by late endosomes, due to the interaction between the C-terminal domain of BST-2 and the cellular endocytic machinery (Miyakawa et al., 2009). Alternatively, virions can be trapped at the cell surface by BST-2 and form large aggregates (Casartelli et al., 2010). In both cases, their spread as cell-free virions is restricted.

Over time, retroviruses have evolved different strategies to counteract BST-2 restriction. For example, HIV-1 encodes the Vpu protein that enhances viral particle release from infected cells by preventing the incorporation of BST-2 into budding virions. It has been proposed that Vpu may down-regulate BST-2 expression at the plasma membrane by (i) promoting proteosomal degradation of the proteins already present on the cell surface, and/or (ii) sequestering newly synthesized proteins into the TGN, thus impairing them to reach the cell membrane (Neil et al., 2008; Van Damme et al., 2008). Similarly, it has been suggested that HIV-2 Env may remove BST-2 proteins from the cell surface and subsequently divert them from the normal recycling pathway, thus depleting their global level (Hauser et al., 2010). Finally, it has been recently demonstrated that the SIV Nef protein reduces the expression of cell surface BST-2 by sequestering the protein away from sites of particle release (Zhang et al., 2011).

1.8 Endogenous retroviruses (ERVs)

ERVs represent remnants of past rounds of germ line infections by exogenous retroviruses. They have been found in all vertebrates studied to date, where they represent a significant percentage of the total genome. Up to 8-10% of human and mouse genomes, for instance, is thought to be of retroviral origin (Jern and Coffin, 2008).

Over time, in some species, including human and non-human primates, most ERVs have accumulated mutations and/or deletions, which prevent them to express viral proteins and release potential infectious viruses (Jern and Coffin,

2008). Conversely, in other species, such as koala and sheep, some ERVs have maintained intact open reading frames for most of their genes and, therefore, are potentially able to produce infectious particles. These, in turn, can re-infect the host germ line and give rise to genomic amplification of some proviral loci. Interestingly, both koala and sheep genomes are thought to be invaded by exogenous retroviruses at the present time, suggesting that, in these animal species, the process of endogenization is still ongoing (Arnaud et al., 2007a; Tarlinton et al., 2006).

For many years, ERVs have been considered as merely molecular "junk" or parasites. It is now clear that host genomes have coevolved with ERVs, preventing or minimizing the deleterious consequences of their unrestrained integrations while capitalizing their adaptive potential: in other words, turning some "junk" into treasure (Goodier and Kazazian, 2008). Indeed, over time, some ERVs have been positively and repeatedly selected by host genomes, and fulfil now useful functions in diverse aspects of host biology, including antiviral activity and placental morphogenesis (Blaise et al., 2003; Dupressoir et al., 2005; Varela et al., 2009).

1.8.1 ERVs and host defence

Some ERVs have been co-opted by host species for their role in protecting them against related exogenous and pathogenic retroviruses. For instance, in mouse, Fv1 blocks the early steps of MLV replication by binding to the viral CA (Hilditch et al. 2011) (Fig. 8), with a mechanism similar to that adopted by TRIM5α. *Fv1* is a relic of a *gag*-like gene from an endogenous retrovirus not related to MLV (Best et al., 1996; Friend, 1957; Lilly, 1967). This locus exists as two major alleles, $Fv1^n$ and $Fv1^b$, exhibiting co-dominance and interference with different strains of MLV. $Fv1^n$ is present in NIH-Swiss mice and confers resistance to B-tropic[16] but not to N-tropic MLV strains. $Fv1^b$, on the other hand, is expressed in BALB/c mice and displays the opposite phenotype (Lilly and Steeves, 1973). A less common third allele, $Fv1^{nr}$, inhibits B-tropic viruses and some N-tropic viruses (Steeves and Lilly, 1977). The mouse genome possesses also the *Fv4* locus, a remnant of a

[16] MLVs are classified into four different subgroups, depending on host susceptibility to viral infections. Ecotropic MLVs can only infect murine cells, although different ecotropic viruses do not necessarily use the same cell surface receptor. Xenotropic MLVs, on the other hand, can infect only non-murine cells. Finally, poly and amphotropic MLVs have the ability to infect both murine and non murine cells through different cellular receptors (Hunter, 1997).

defective ERV that encodes an ecotropic MLV-like Env (Kozak et al., 1984). Expression of Fv4 can saturate cell surface receptors sufficiently to block entry by exogenous ecotropic MLV strains (Fig. 8) (Ikeda and Odaka, 1983; Ikeda et al., 1985).

Similarly to Fv4, endogenous sheep betaretroviruses (enJSRVs) prevent JSRV infections by receptor interference in *in vitro* experiments (Spencer et al., 2003) (Fig. 8). Moreover, some of them can block JSRV *in vitro*, by a unique mechanism of interference that occurs at a post-integration step of the viral replication cycle (see paragraph 1.10.3) (Arnaud et al., 2007b; Mura et al., 2004; Murcia et al., 2007) (Fig. 8).

Endogenous MMTV proviruses (Mtvs) exert a completely different restriction activity, as they reduce, or completely abolish, viral spread of their exogenous counterparts by altering host immune response (Fig. 8). MMTV infects newborn pups during the first weeks of their life, passing through the milk of viremic female mice. The primary target of MMTV are B lymphocytes and dendritic cells (Baribaud et al., 1999) that, following infection, express Sag (for *sup*er*ant*i*g*en) at their cell surface (Korman et al., 1992) and initiate T cell responses (Baribaud et al., 1999). Activation of T cells, in turn, promotes division of infected cells and, thus, viral spread (Golovkina et al., 1998). By contrast, expression of different subtypes of Sag by Mtvs leads to progressive deletion of T cells, thereby reducing spread of exogenous MMTV (Golovkina et al., 1992; Golovkina et al., 1998; Mustafa et al., 2000). Recently, it was shown that *Mtv*-null BALB/c mice are resistant to MMTV infection, suggesting that they may possess other mechanisms of viral defence besides Mtvs. Interestingly, these mice were found to be resistant to *Vibrio cholerae* and susceptible to *Salmonella typhimurium* infections, arguing that Mtv-encoded Sag may also provide a unique genetic susceptibility towards specific bacteria (Bhadra et al., 2006).

1.8.2 ERVs and placental development

Several lines of evidence suggest that retroviruses have contributed to the evolution of placental mammals. Intact *env* genes derived from full-length or defective ERVs are highly expressed in the genital tract and placental tissues of many mammals, including mice, sheep and humans (Blaise et al., 2003; Blond et al., 2000; Dunlap et al., 2006b; Dupressoir et al., 2005; Mangeney et al., 2007;

Mi et al., 2000). It has been proposed that ERVs might have infected some primitive aplacental mammal-like species at an early intrauterine stage, giving rise to cellular proliferation and formation of a primitive placenta (Harris, 1991). Alternatively, ERVs might have been independently acquired by mammals, during evolution, for a convergent biological role in placental morphogenesis (Stoye, 2009; Villareal, 1997).

Env proteins promote membrane fusion, a process that highly resembles the formation of the syncytiotrophoblast. A systematic screening of the human genome led to the identification of *syncytin-1* and *syncytin-2* genes, which derive from HERV *env* genes and are specifically expressed at the cytotrophoblast–syncytiotrophoblast interface of placenta[17] (Blaise et al., 2003). Syncytin-1 was shown to be directly involved in cell fusion (Blond et al., 2000; Mi et al., 2000), whereas Syncytin-2 displays immunosuppressive activity, most likely associated with maternal-fetal tolerance (Mangeney et al., 2007). The identification of sequences in mouse (*syncytin-A* and *syncytin-B*) (Dupressoir et al., 2005) and sheep (enJSRV *env*) (Dunlap et al., 2006b), which exhibit similar physiological properties to those of human *syncytin* genes (Dunlap et al., 2006b; Dupressoir et al., 2009), strongly supports the hypothesis that ERVs have been positively selected for their critical roles in the evolution of placenta and viviparity in mammals.

1.8.3 ERVs and effects on host transcriptome

Curiously, despite the copious number of ERVs and related elements in host genomes, most tissues do not express high levels of ERV RNA or replicating viruses. Hosts control ERVs expression in many ways, including ERVs diversion outside of genes and/or in antisense orientation, cellular restriction factors, DNA methylation and small interfering RNAs (siRNAs) (Maksakova et al., 2008). Moreover, ERVs expression is restricted in most differentiated tissues in order to prevent insertional mutagenesis events that may have potentially deleterious effects for the host (Rowe and Trono, 2011). Indeed, ERVs usually "tend" to integrate and become fixed into non essential regions of the genome, where their presence is less harmful for the host (Jern and Coffin, 2008). Nevertheless, they can influence the transcriptional activity of nearby genes, due to the

[17] Refer to figure 13.

potential binding sites for transcription factors in their proviral LTRs. For example, the transcriptional activity of human beta-like globin genes is regulated by a solo LTR sequence, related to HERVs and integrated upstream of the locus control region (Long et al., 1998; Pi et al., 2010).

There has been considerable debate on whether ERVs may be associated with diseases or not. Some studies have shown that HERVs may play roles in human neuronal syndromes, including multiple sclerosis (Antony et al., 2007). In contrast, some retrovirologists have dismissed them as merely "human rumour viruses", because they produce high levels of "background noise" that may alter the signal from infectious retroviruses or, alternatively, produce false-positive results (Jern and Coffin, 2008; Voisset et al., 2008). Further work is needed to evaluate carefully any potential correlation between ERVs and diseases.

1.9 Jaagsiekte sheep retrovirus (JSRV)

JSRV is an exogenous pathogenic retrovirus and the etiological agent of ovine pulmonary adenocarcinoma (OPA), a contagious lung cancer of sheep (Palmarini et al., 1999a). OPA was firstly reported in 19th century in South Africa, when farmers coined the term *jaagsiekte* to describe the respiratory distress observed in affected animals and that they associated to sheep being chased (in Afrikaans *jag* means chase and *ziekte* sickness) (Tustin, 1969). Since then, OPA has spread in many geographical areas around the world, where it affects both domestic and wild sheep (Griffiths et al., 2010).

The earliest evidence that OPA is induced by a retrovirus dates back to 1970s, when retrovirus-like particles were observed by electron microscopy in lung cancer cells (Malmquist et al., 1972; Perk et al., 1971). These results were strengthened by the identification of RT activity (Verwoerd et al., 1980) and retroviral RNA (Perk et al., 1974; Verwoerd et al., 1980) in tumour extracts, and the experimental reproducibility of clinical and histological signs of the disease in sheep injected with tumour extracts from OPA-affected animals (Martin et al., 1976). In 1991, Denis York and colleagues isolated the genomic RNA of what we now know as JSRV-SA (JSRV-South Africa), from washes of OPA-affected lungs (York et al., 1991). One year later, the same group provided the first evidence of enJSRVs in the genome of sheep and goats (York et al., 1992). This discovery

raised suspicions about a potential role of enJSRVs in the development of OPA. These doubts were soon ruled out when it was demonstrated that only JSRV could be detected from lung tumour samples (Bai et al., 1996; Palmarini et al., 1996a). Conclusive evidence that JSRV was the causative agent of OPA came from the isolation of the infectious molecular clone $JSRV_{21}$. The inoculation of new-born lambs with viral particles collected from the supernatants of 293T cells expressing $JSRV_{21}$ proved that this virus is necessary and sufficient to induce OPA (Palmarini et al., 1999a).

1.9.1 Ovine pulmonary adenocarcinoma (OPA)

Sheep affected by OPA show different symptoms, including progressive dyspnoea associated with loss of weight, and usually die for respiratory failure after a protracted incubation period (Sharp and DeMartini, 2003). One of the characteristic clinical signs of OPA is the production of copious amount of fluid in the lungs, which drains from the nostrils of affected sheep once their hind limbs are raised above their head. However, in many cases, no lung fluid can be observed and, therefore, definitive diagnosis of OPA can be made only after histopathological examination (De las Heras et al., 2003). At the post-mortem exam, naturally infected animals at advanced stages of OPA present usually a thin carcass with enlarged lungs infiltrated with tumour and airways filled with fluid. Extrathoracic metastases have also been reported, but are generally rare (DeMartini et al., 1988). Remarkably, JSRV infected animals do not develop antibodies against the virus (Ortín et al., 1998), probably because of the expression of the highly related enJSRVs, during ontogeny, that may lead to tolerance (Spencer et al., 2003).

In natural conditions, JSRV can infect both adult sheep and lambs, perhaps through aerosolized particles and maternal colostrum/milk (Caporale et al., 2005; Grego et al., 2008). Curiously, the virus is primarily detected in peripheral blood leucocytes and lymphoid organs rather than in lungs, and only a minority of naturally JSRV infected animals develop OPA (Caporale et al., 2005). These findings led to the hypothesis that JSRV may use lymphoreticular cells as a viral reservoir, and spreads to the lungs only when the latter are more susceptible to infection. In a recent study, Claudio Murgia and colleagues found that, in the

lungs, JSRV infects and transforms proliferating type 2 pneumocytes (also termed lung alveolar proliferating cells, LAPCs) (Murgia et al., 2011) (Fig. 9).

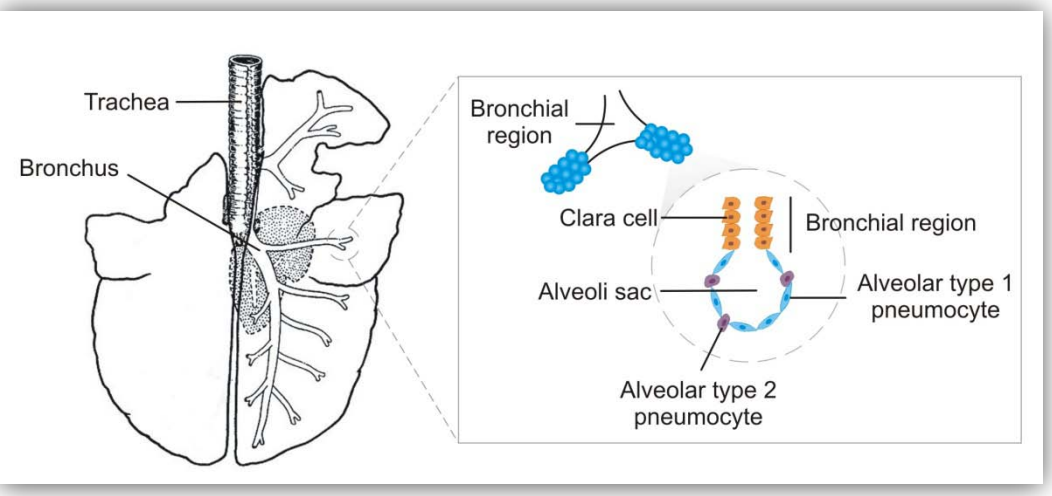


Figure 9| Schematic representation of cell types present in the terminal bronchioli and alveoli of sheep lungs. Figure adapted from (Murgia et al., 2011).

These cells are normally abundant in young lambs during post-natal development, and in adult sheep, following injury to the bronchioalveolar epithelium. Sheep naturally developing OPA present often a wide variety of other respiratory pathogens, including lungworms (Snyder et al., 1983), Maedi-Visna virus (MVV, lentivirus) (Dawson et al., 1990) and other bacterial infections. It has been proposed that these pathogens may act synergistically to render more susceptible sheep lungs to JSRV transformation (Caporale et al., 2005; Murgia et al., 2011). Similarly to most retroviruses [see paragraph 1.6.3 and (Suzuki and Craigie, 2007)], JSRV infects more efficiently cells in division, when the nuclear membrane dissolves. Lung inflammation induced by primary pathogens may indeed promote active divisions of LAPCs to repair injured tissues, thus providing optimal conditions for JSRV transformation. This hypothesis is supported by the observation that lung lesions increase the number of LAPCs, and render infected animals more susceptible to JSRV infection and OPA (Murgia et al., 2011).

OPA represents an excellent animal model to study human bronchioalveolar carcinoma, as both diseases share many clinical and histological traits (Palmarini and Fan, 2001; Murgia et al., 2011). Indeed, recent results obtained in our group suggest that human lung cancer may originate from cell types that are more closely related to those present in sheep than in mice (Murgia et al., 2011).

1.9.2 JSRV: genomic organization and mechanism of oncogenesis

JSRV is an exogenous sheep betaretrovirus phylogenetically related to ENTV, M-PMV and MMTV. Its genome is approximately 7.5 Kb in length and exhibits a simple organization, typical of replication competent retroviruses. Besides encoding the classical retroviral genes *gag*, *pro*, *pol* and *env*, JSRV harbours an additional open reading frame of unknown function (hence termed *orf-x*), which overlaps the 3' end of *pol* (Palmarini et al., 1999a; York et al., 1992) (Fig. 10).

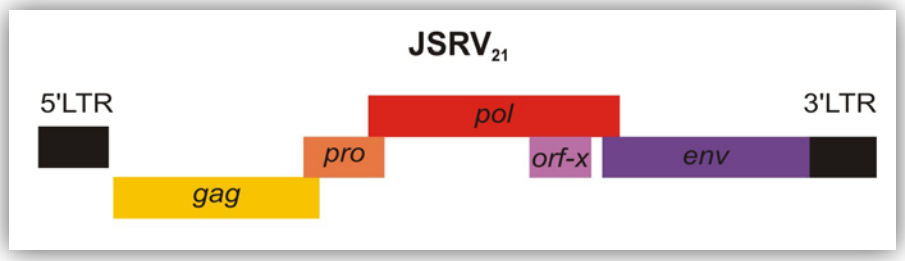


Figure 10 | Genetic organization of the $JSRV_{21}$ molecular clone.

This reading frame displays no homology to any known viral or cellular genes, and is highly conserved among different JSRV isolates (Rosati et al., 2000), although it does not appear to be required for viral pathogenesis (Cousens et al., 2007). To date, three highly related JSRV isolates (Palmarini and Fan, 2003), have been cloned from OPA affected sheep: JSRV-SA (York et al., 1991), $JSRV_{21}$ (Palmarini et al., 1999a) and $JSRV_{JS7}$ (DeMartini et al., 2001). $JSRV_{21}$ is probably the most extensively studied among these three viral strains.

The lung tropism of JSRV is determined by its LTR sequences. Indeed, besides containing the viral promoter, these regions include enhancer elements activated by lung-specific transcription factors, including the hepatocyte nuclear factor-3β (HNF-3 β), nuclear factor I (NFI), and CCAAT/enhancer binding protein (C/EBP) (McGee-Estrada and Fan, 2006; McGee-Estrada et al., 2002; Palmarini et al., 2000).

The *gag* gene encodes a single polyprotein that is cleaved into at least five products: matrix (MA), p15, capsid (CA), nucleocapsid (NC) and p4 (Murcia et al., 2007; Palmarini et al., 1999b). Similarly to other retroviruses, JSRV contains a putative M domain at the N-terminal portion of Gag consisting of a myristate group and a stretch of basic residues, both required for viral assembly and

trafficking (Mura et al., 2004; Murcia et al, 2007; Swanstrom and Wills 1997). Mutations in this region alter the ability of JSRV to reach the cell membrane and release viral particles (Mura et al., 2004; Murcia et al., 2007). The L domain lays within the p15 protein and contains core amino acid motifs of PSAP and PPAY (Murcia et al., 2007) analogous to those present in M-PMV (Gottwein et al., 2003). In accordance to what observed in other retroviruses, mutations in the JSRV L domain result in a defect of the late steps of viral replication cycle that impairs viral budding (Freed, 2002; Murcia et al., 2007).

As described for other betaretroviruses, *pro* and *pol* are encoded in different open reading frames (Palmarini et al., 1999a). The *pro* gene expresses (i) a deoxyuridine triphosphatase (dUTPase) that prevents incorporation of deoxyuridine triphosphate (dUTP) by the reverse transcriptase, and (ii) a protease that cleaves viral polyprotein precursors. The *pol* gene encodes the reverse transcriptase and the integrase, involved, respectively, in the reverse transcription and integration processes (Leroux et al., 2007).

The *env* mRNA is approximately 2.4 Kb in length and derives from a single-splicing event (Palmarini et al., 2002). After maturation, it gives rise to the transmembrane (TM) and surface (SU) domains. SU mediates viral entry into the cells by interacting with Hyal2, the cellular receptor of JSRV (Spencer et al., 2003). Hyal2 is a member of the hyaluronglucosaminidase family, involved in the enzymatic degradation of hyaluronic acids present in vertebrates' extracellular matrix. It is ubiquitously expressed in sheep, in accordance with the ability of JSRV to infect different cell types both *in vitro* (Palmarini et al., 1999b) and *in vivo* (Palmarini et al., 1996b). However, viral expression is restricted to specific bronchioalveolar epithelial cells due to the tropism conferred to JSRV by the LTR sequences. The TM domain anchors the virus to the cell lipid bilayer and confers to JSRV the ability to induce cell transformation (Palmarini et al., 2001b). In particular, the cytoplasmic tail (CT) of the Env glycoprotein bears a YXXM motif[18] that, most likely, activates Ras/MEK/MAPK and PI-3K/Akt-dependent pathways (Chow et al., 2003; Maeda et al., 2005; Palmarini et al., 2001b; Varela et al., 2006). Remarkably, expression of JSRV Env glycoprotein alone is sufficient to induce cell transformation both *in vitro* (Maeda et al., 2001) and *in vivo* (Murgia

[18] Y for tyrosine, X for any amino acid and M for methionine.

et al., 2011). This renders JSRV the only virus to harbour a dominant oncoprotein that is necessary and sufficient to trigger tumor development (Alberti et al., 2002; Murgia et al., 2011).

As already mentioned (paragraph 1.6.4.3), recent studies demonstrated that JSRV Env possesses a signal peptide (JSE-SP) located at the N-terminal portion of the protein (Caporale et al., 2009; Hofacre et al., 2009). Similarly to MMTV SP (Byun et al., 2010), JSE-SP plays a critical role in viral replication cycle, as it acts as a post-transcriptional regulator of viral gene expression (Caporale et al., 2009; Hofacre et al., 2009; Nitta et al., 2009) (see paragraph 4.1).

1.10 Endogenous sheep betaretroviruses: enJSRVs

In a recent study, Frédérick Arnaud and colleagues screened a bacterial artificial chromosome (BAC) library derived from the genomic DNA of a single Texel ram. Sequence analyses revealed that sheep genome harbours, at least, twenty-seven copies of endogenous betaretroviruses, highly related to the exogenous and pathogenic JSRV and hence termed enJSRVs (Arnaud et al., 2007a) (Fig. 11). Most of these loci possess defective genomes, due to the presence of premature termination codons, large deletions and/or recombinations. However, five enJSRV loci (enJSRV-7, enJSRV-15, enJSRV-16, enJSRV-18 and enJSRV-26) display intact genomic organization and uninterrupted open reading frames for all of the retroviral genes (*gag*, *pro*, *pol*, *orf-x*, and *env*), resembling replication competent retroviruses. Four of the five intact enJSRVs (enJSRV-15, enJSRV-16, enJSRV-18 and enJSRV-26) exhibit identical 5' and 3' LTRs, which is indicative of relatively recent integration into the host germ line. This hypothesis is further reinforced by the presence of two loci (enJSRV-16 and enJSRV-18) that are 100% identical at the nucleotide level along their entire genomes (Arnaud et al., 2007a).

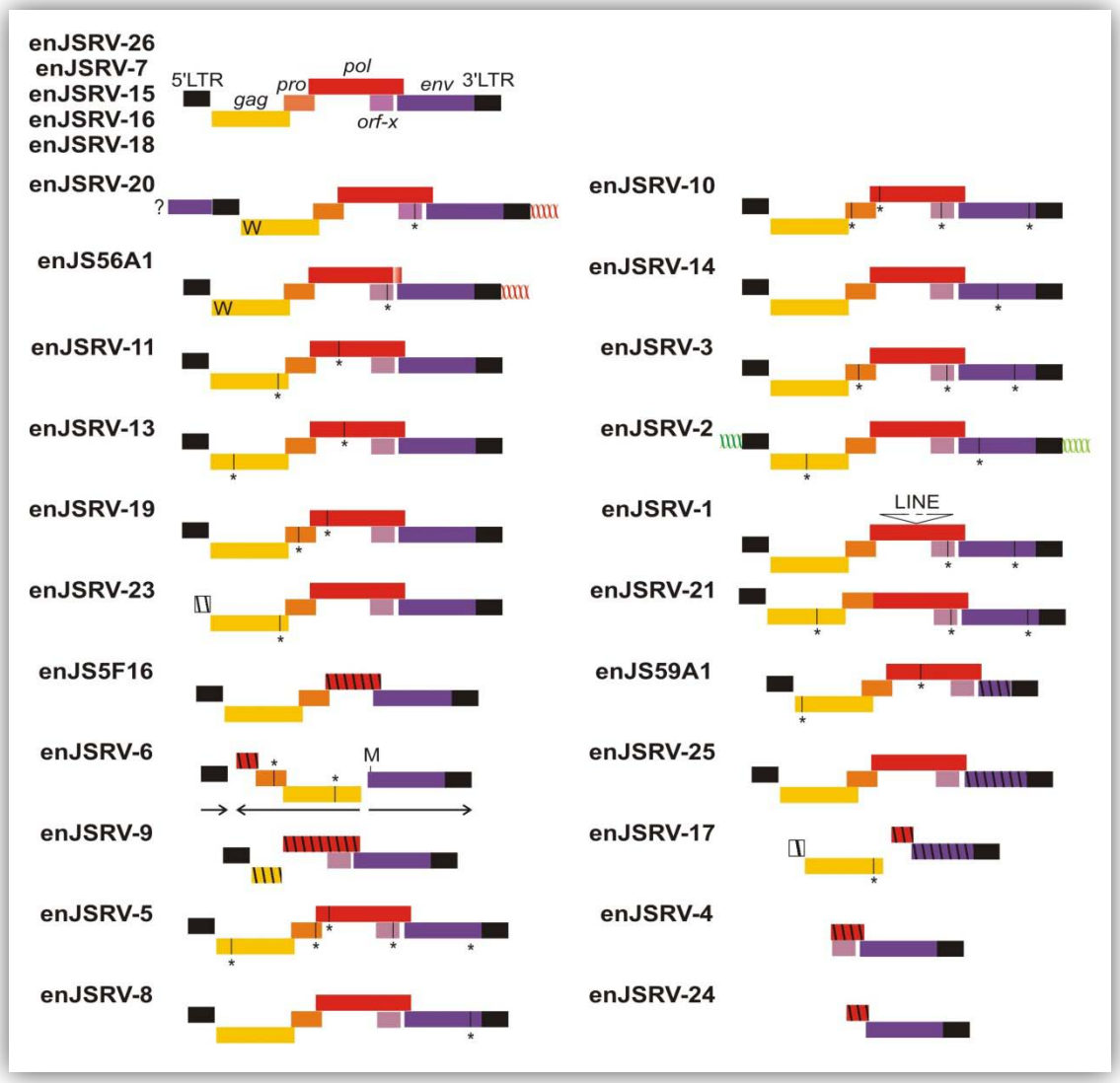


Figure 11 | Genetic organization of enJSRV proviruses. All the genomic sequences flanking enJSRV proviruses contain a six base pairs duplication that is the hallmark of retroviral integration. The only exceptions are represented by enJSRV-20, which contains a portion of an *env* gene (indicated by a purple box and a question mark) before the 5′LTR, and enJSRV-2, which does not contain the same six base pairs sequences flanking the LTRs. Five of the twenty-seven enJSRVs possess an intact genomic organization, typical of replication competent exogenous retroviruses (top). The two transdominant proviruses enJS56A1 and enJSRV-20 possess a trypthophan residue (W) at position 21 of Gag and identical 3' genomic flanking regions. The enJSRV-6 locus possesses an additional methionine (M) in Env besides the canonical start codon present in JSRV and other enJSRV loci. Moreover, in enJSRV-6, *gag* and *pro* are in opposite direction compared to the 5′ and 3′ LTRs and *env* (indicated by horizontal arrows). enJSRV-1 presents a LINE element within the *pol* coding region. Premature termination codons are represented by a vertical line and an asterisk (*). Large deletions in proviral genomes are indicated by hatched boxes. Figure adapted from (Arnaud et al., 2007a).

The enJSRVs loci share 85-89% identity with the *gag* and *env* sequences of the infectious molecular clone $JSRV_{21}$ (Arnaud et al., 2007a; Palmarini et al., 1999a). The major differences are found in the U3 sequences (Bai et al., 1996) and three regions spanning *gag* and *env*, referred to as variable regions (VR1, VR2 and VR3). VR1 and VR2 lie within the MA domain of Gag, while VR3 encompasses the last sixty-seven amino acids of the TM domain of Env. Sixteen of the twenty-

seven enJSRVs possess intact open reading frames for *env* (Fig. 1.11), even though two of them (enJSRV-4 and enJSRV-24) lack most of the other genes (Arnaud et al., 2007a; Palmarini et al., 2000). Standard entry assays revealed that enJSRV Env mediates viral entry *via* Hyal2, which serves also as a cellular receptor for the exogenous JSRV and ENTV sheep betaretroviruses (Spencer et al., 2003). However, enJSRVs Env glycoproteins lack the YXXM motif critical for JSRV cell transformation and, therefore, are unable to induce foci in classical transformation assays of rodent and chicken cell lines (Arnaud et al., 2007a; Palmarini et al., 2001b).

1.10.1 Evolutionary history of enJSRVs

Once retroviruses infect germ cells, they leave a permanent footprint in the genome of their hosts, which can reveal the dynamics of host-pathogen interplay across long evolutionary periods, spanning even millions of years. ERVs hold therefore great potential as informative markers for evolutionary studies of both viral and host genomes. Along this line, in a recent work conducted by Bernardo Chessa and colleagues in our laboratory, the characterization of the enJSRVs distribution in various breeds of domestic sheep (*Ovis aries*) has provided valuable insights into the history of sheep domestication (Chessa et al., 2009).

The "age" of a provirus can be directly inferred by knowing the time of speciation of species phylogenetically related. Alternatively, retroviral integration time can be estimated assessing the sequence divergence of 5' and 3'LTRs, assuming that both regions were identical at the time of infection but evolved separately afterwards, at the same evolutionary rate of non-coding regions. According to that, "young" ERVs most likely possess identical or nearly-identical LTRs, while "old" proviruses have LTRs significantly divergent. However, it is important to note that LTRs may have different evolutionary rates, and homologous ERV loci may undergo different selective pressures in different species (Martins and Villesen, 2011).

Sequence analyses and phylogenetic data suggest that enJSRVs entered the host genome before the speciation of *Ovis* and *Capra* genera, approximately 5-7 million years ago (MYA) (Fig. 12). Some enJSRV loci were found in all of the species of *Ovis* genus, such as *O. aries, O. nivicola, O. canadensis* and *O. dalli*, whereas others were restricted to *O. aries*, including eight insertionally

polymorphic loci[19]. These findings and the current knowledge on ruminant evolution suggest that the insertionally polymorphic enJSRVs entered the sheep genome less than 9,000 years ago, after sheep domestication. In particular, enJSRV-26 was found only in a single Texel ram. Considering the history of this breed, it is possible to speculate that enJSRV-26 integrated very recently in the sheep genome, probably less than 200 years ago, and may even be a unique integration event (Arnaud et al., 2007a).

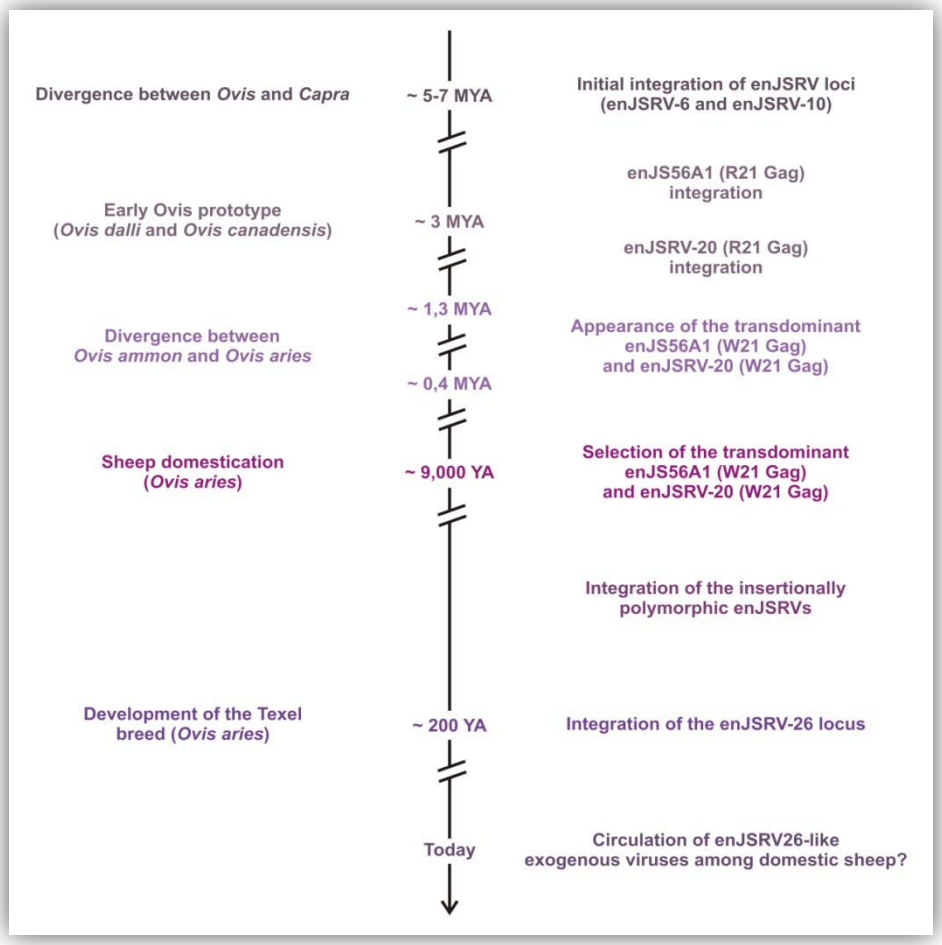


Figure 12 | Evolutionary history of enJSRVs. Figure adapted from (Arnaud et al., 2007a).

1.10.2 Role of enJSRVs in sheep reproductive biology

enJSRVs are abundantly expressed in sheep reproductive organs, including endometrial luminal and glandular epithelia of uterus and epithelia of oviducts and cervix (Palmarini et al., 1996b; Palmarini et al., 2000; Palmarini et al., 2001a; Spencer et al., 1999). enJSRV mRNAs can also be detected in the lymphoid cells of the lamina propria of the gut, in the bronchial epithelial cells

[19] These enJSRVs are present only in a proportion but not all domestic sheep.

of the lungs and in the cortico-medullary junction of the thymus, where T lymphocytes undergo the process of maturation. Expression of enJSRVs in these organs may render sheep tolerant towards related exogenous betaretroviruses, and explain why JSRV infected animals do not develop antibodies against it (Ortín et al., 1998; Sharp and Herring, 1983).

In the conceptus (embryo/foetus and associated extraembryonic membranes), enJSRV *env* mRNA is mainly detected in trophoblast giant binucleate cells (BNCs) and multinucleated syncytia, both required for implantation and nutrition of the conceptus (Palmarini et al., 1996b; Palmarini et al., 2000; Palmarini et al., 2001a; Spencer et al., 1999) (Fig. 13).

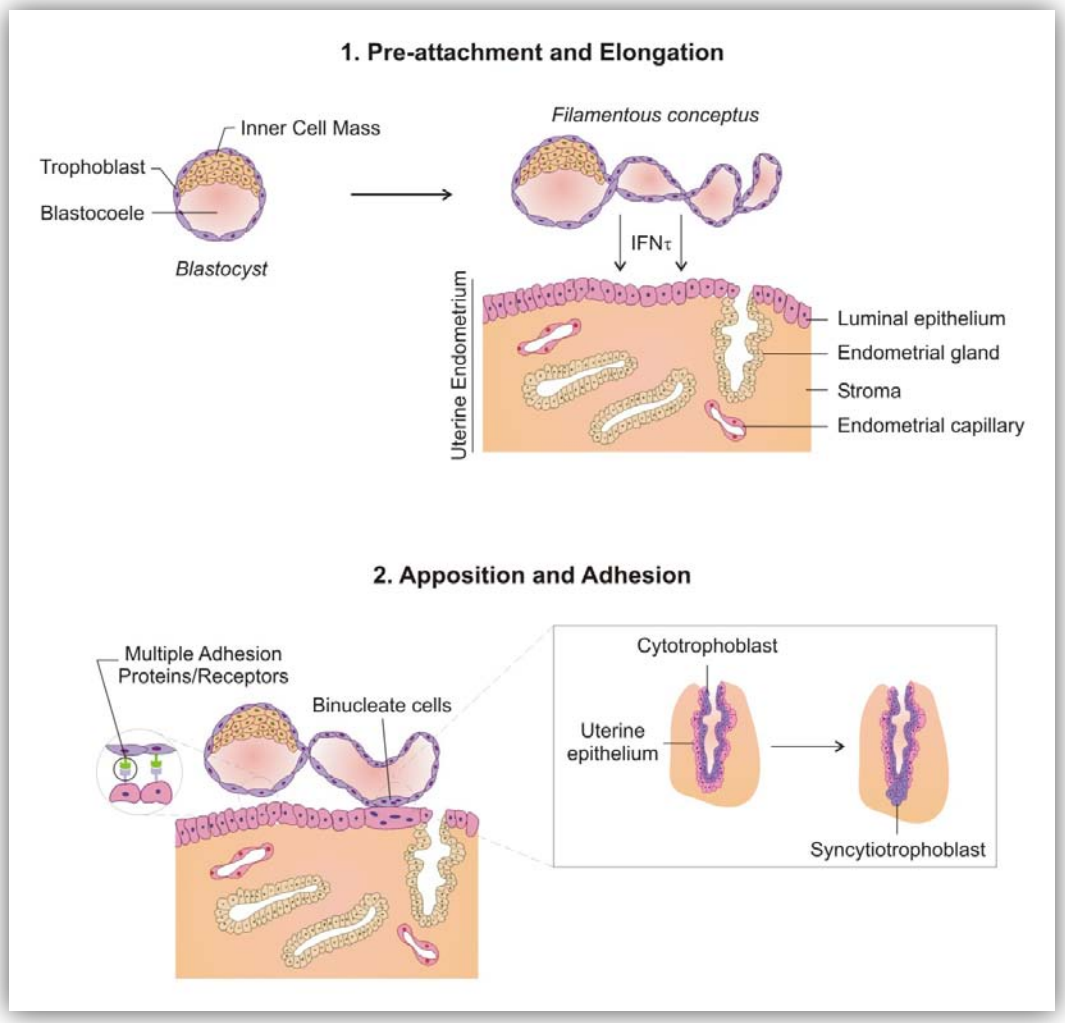


Figure 13 | Schematic representation of blastocyst implantation in sheep. Figure adapted from (Spencer et al., 2004).

In vivo experiments demonstrate that inhibition of enJSRV Env expression retards blastocyst growth and elongation, and inhibits differentiation of trophoblast giant BNC, resulting in loss of pregnancy (Dunlap et al., 2006a).

These results indicate that enJSRV *env* mRNA is required for conceptus elongation and trophectoderm growth in sheep. During early pregnancy, the developing conceptus secretes IFN tau (IFNτ), which functions as a pregnancy recognition signal in ruminants. Like other type I IFNs, IFNτ activates signalling pathways involved in maintaining maternal tolerance of foetal allograft and protecting the conceptus from viral infections (Bazer et al., 2008). In the ovine endometrium, IFNτ and progesterone upregulate expression of BST-2 (Arnaud et al., 2010) and enJSRV Env (Palmarini et al., 1996b; Palmarini et al., 2000; Palmarini et al., 2001a; Spencer et al., 1999), respectively. In particular, IFNτ enhances BST2 mRNA only in the stroma and not in the luminal epithelium endometrium, where enJSRVs are expressed. The *BST-2* gene is duplicated in ruminants in the A and B isoforms (oBST-2A and oBST-2B). Phylogenetic analyses indicate that this duplication occurred approximately 25 MYA (Arnaud et al., 2010), before the speciation of the Bovinae subfamily (Hassanin and Douzery, 2003) and, thus, before the initial integration of enJSRVs in the host genome (Arnaud et al., 2007a). Both isoforms are able to block enJSRV viral particles *in vitro*, even though with different efficiencies (Arnaud et al., 2010). Preliminary data obtained by Lita Murphy, a PhD candidate in our group, indicate that the differences in the antiviral restriction by oBST-2A and oBST-2B are attributable to differences in their amino acid sequences (Murphy and Palmarini, unpublished data).

1.10.3 Role of enJSRVs in host defence

Besides their role in conceptus development and placenta morphogenesis, enJSRVs could protect their host by interfering with related exogenous retroviruses at early and late stages of the retroviral cycle (Fig. 8). *In vitro* experiments demonstrated that JSRV cannot enter cell lines derived from the ovine genital tract that express enJSRV RNAs. Thus, it is possible that enJSRVs block JSRV entry by receptor interference, as both endogenous and exogenous betaretroviruses utilise Hyal2 as a cellular receptor (Spencer et al., 2003). This mechanism could also be adopted by the host to control unrestrained viral infections by newly emergent enJSRVs. Indeed, a recent study conducted by Sarah Black and colleagues demonstrated that the enJSRVs expressed in the uterine endometrial epithelia of sheep can release viral particles into the uterine lumen. By using a *trans* species embryo model, in which bovine embryos were

transferred into ovine uteri, these authors found that enJSRV viral particles can potentially infect the trophoblast of the developing conceptus. Interestingly, sequence analyses revealed that only the most recently integrated enJSRV loci (including enJSRV26-like sequences) were found to be consistently expressed in the uterine endometrium and form infectious viral particles (Black et al., 2010a).

enJS56A1, one of the enJSRV loci, can block JSRV release by a mechanism known as JLR (for *J*SRV *l*ate *r*estriction) (Mura et al., 2004). In addition, it has been recently demonstrated that enJS56A1 is able to restrict viral release by intact enJSRV proviruses as efficiently as the exogenous JSRV (Arnaud et al., 2007a) (see paragraph 3.1). Interestingly, enJSRV-26, which is the "youngest" enJSRV isolated to date, possesses the unique ability to escape JLR (Arnaud et al., 2007a). The presence of such escape provirus strongly supports the hypothesis that transdominant proviruses play indeed a critical role as restriction factors against related exogenous retroviruses. The fact that (i) enJSRV-26 was detected only in a single Texel ram (Arnaud et al., 2007a) and (ii) enJSRV26-like sequences were found in *trans* species conceptuses of two Texel animals (Black et al., 2010a) strongly argues that an exogenous retrovirus related to enJSRV-26 is still circulating within sheep population. Moreover, it is noteworthy that five of the twenty-seven enJSRV loci isolated to date (enJSRV-7, enJSRV-15, enJSRV-16, enJSRV-18 and enJSRV-26) are insertionally polymorphic and display an intact genomic organization (Arnaud et al., 2007a). Overall these findings suggest that, in sheep, the process of endogenization is still ongoing and, therefore, the evolutionary interplay between endogenous and exogenous sheep betaretroviruses and their hosts has not reached an equilibrium yet.

Chapter II

Materials and methods

2.1 Plasmids

Plasmids pCMV4JS21, pCMV5-enJS26, pCMV5-enJS18, and pCMV2en56A1 have been described previously (Arnaud et al., 2007a; Palmarini et al., 1999a; Palmarini et al., 2000). Briefly, pCMV4JS21 is an expression plasmid for the infectious and oncogenic molecular clone $JSRV_{21}$, while pCMV5-enJS26, pCMV5-enJS18 and pCMV2en56A1 express, respectively, the endogenous proviruses enJSRV-26, enJSRV-18 and enJS56A1. Viral particles can be recovered from the supernatant of 293T cells transiently transfected with either pCMV4JS21, pCMV5-enJS26 or pCMV5-enJS18 (Arnaud et al., 2007a; Palmarini et al., 1999a). Conversely, pCMV2en56A1 is unable to release viral particles *in vitro* due to the presence of a defective Gag protein (Mura et al., 2004). In all the expression plasmids described above, the U3 region in the proximal LTR was replaced by the cytomegalovirus (CMV) immediate early promoter, as already described (Arnaud et al., 2007a; Palmarini et al., 1999a; Palmarini et al., 2000). The complete sequences of the exogenous $JSRV_{21}$ and enJSRV-26, enJSRV-18 and enJS56A1 proviruses have been already published (Arnaud et al., 2007a; Palmarini et al., 1999a; Palmarini et al., 2000) and are deposited in GenBank (http://www.ncbi.nlm.nih.gov/), under accession numbers AF105220.1, EF680297.1, EF680301.1 and AF153615.1, respectively.

All of the chimeras/mutants employed in this study were derived from the plasmids listed above and are schematically represented in figure 14. Restriction maps of pCMV5-enJS26 and pCMV5-enJS18 expression plasmids are reported in figure 15.

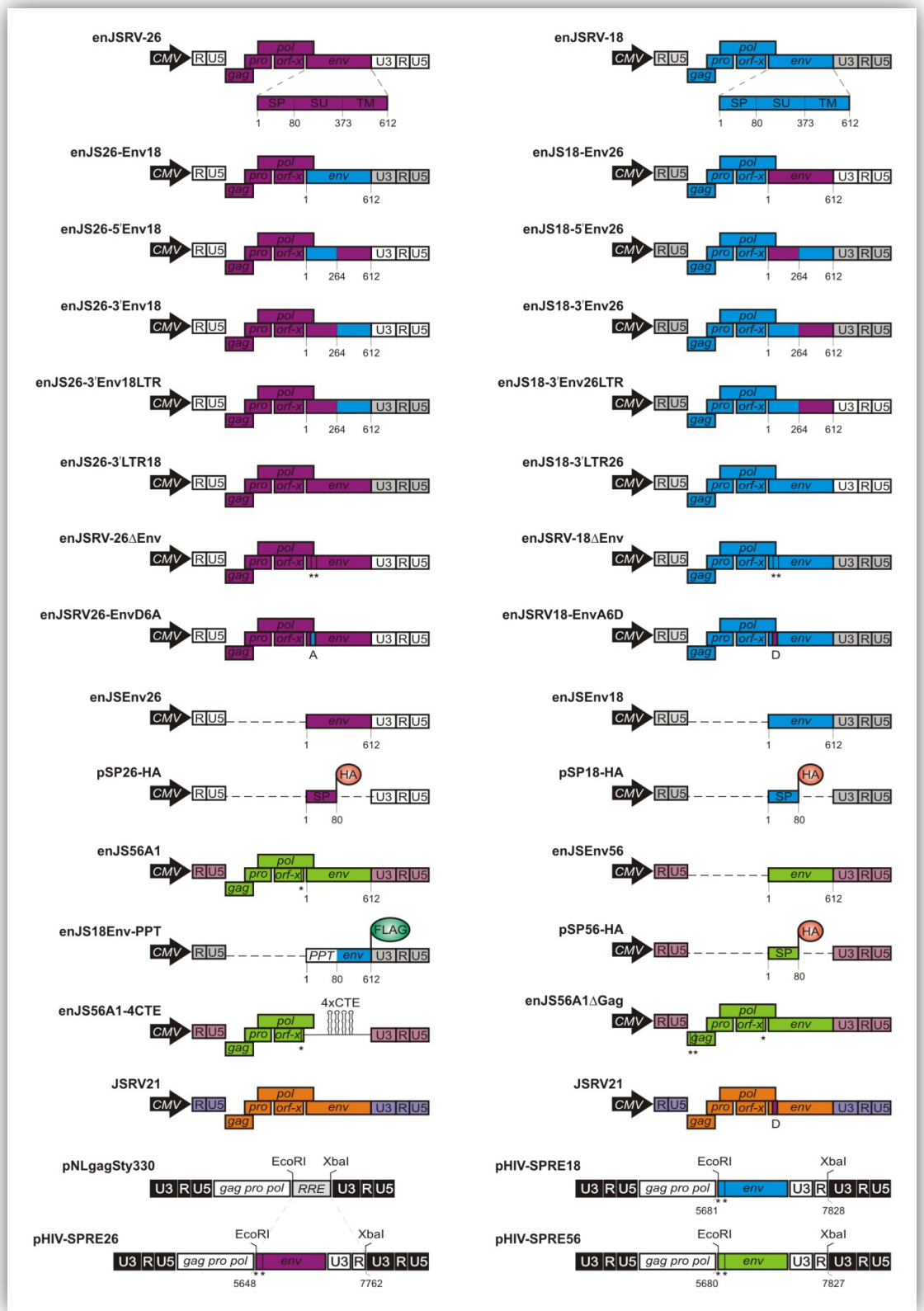


Figure 14 | Schematic representation of the expression plasmids employed in this study. Numbers indicate amino acid residues in the Env glycoprotein. The alanine and aspartic acid are indicated with the one-letter code as A and D, respectively. Premature termination codons are represented with vertical lines and asterisks. pNLgagSty330 is a HIV Gag-Pol expression plasmid that encodes the HIV Rev responsive element (RRE). pHIV-SPRE26, pHIV-SPRE 18 and pHIV-SPRE56 were derived from pNLgagSty330 by replacing the HIV RRE with the correspondent signal peptide responsive element (SPRE) after enzymatic digestion with *EcoRI* and *XbaI*. CMV, cytomegalovirus; CTE, constitutive transport element; R, terminally redundant sequence; U5, unique 5' sequence; U3, unique 3' sequence; SP, signal peptide; PPT, preprotrypsin.

Chimeras enJS26-Env18 and enJS18-Env26 were derived by digesting the correspondent full-length expression plasmids with *BamHI* and *KpnI*, and swapping the *env* gene between enJSRV-26 and enJSRV-18.

In enJS26-5'Env18 and enJS18-5'Env26 chimeras, the 5' end of the *env* gene was exchanged between enJSRV-26 and enJSRV-18, following restriction digestion with *BamHI* and *SwaI* of the corresponding full-length expression plasmids.

enJS26-3'Env18 and enJS18-3'Env26 chimeras were derived by site-directed mutagenesis, by replacing the 3' end of enJSRV-26 *env* with the corresponding sequence in enJSRV-18, and *vice versa* (enJS26-3'Env18: CTA to TTA at position 1696 and TGT to TGC at position 1707, with respect to enJSRV-26 *env*; enJS18-3'Env26: TTA to CTA at position 1696 and TGC to TGT at position 1707, with respect to enJSRV-18 *env*).

enJS26-3'Env18LTR is an expression plasmid of a chimeric provirus containing most of the wild-type enJSRV-26 except for the 3' end of the sequence (i.e., the C-terminus Env and the 3'LTR) that derives, instead, from the corresponding sequence of the enJSRV-18 provirus. Accordingly, enJS18-3'Env26LTR contains the genomic sequence of the enJSRV-18 provirus apart from the 3' end of the Env protein and the 3'LTR, which were replaced by those of the enJSRV-26 provirus. enJS26-3'Env18LTR and enJS18-3'Env26LTR were obtained by digesting, respectively, enJS26-3'Env18 and enJS18-3'Env26 with *PmlI* and *KpnI* and then swapping the corresponding LTR regions. Chimeras enJS26-3'LTR18 and enJS18-3'LTR26 express, respectively, the majority of enJSRV-26 and enJSRV-18 genomic sequences, except for the 3'LTR that were exchanged between the two proviruses. These chimeras were derived by digesting the corresponding wild-type expression plasmids with *PmlI* and *KpnI* and then swapping the LTR regions between them.

All of the enJSRV-26 and enJSRV-18 chimeras were a courtesy of Frédérick Arnaud.

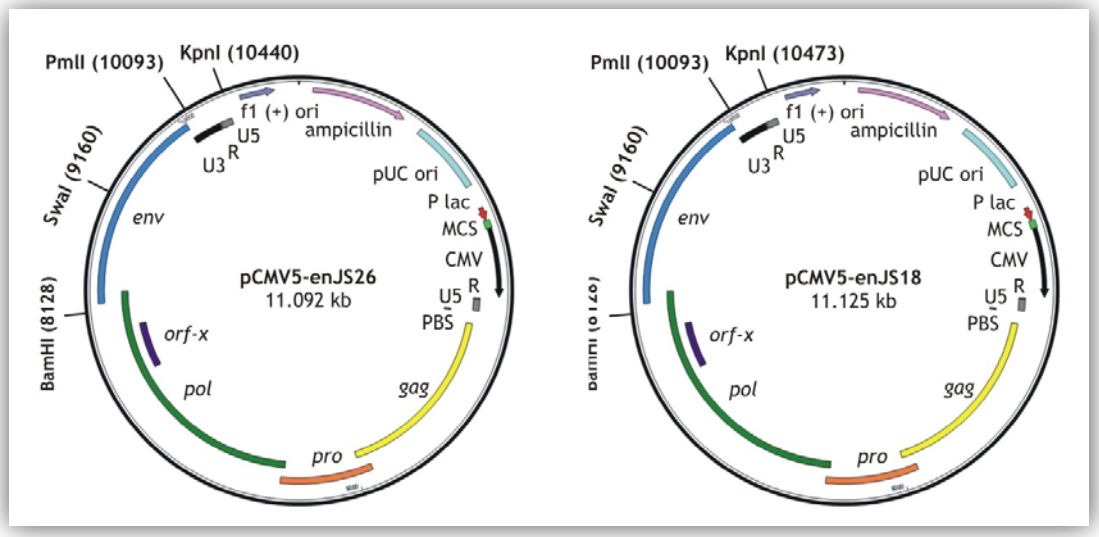


Figure 15| Restriction maps of the expression plasmids for enJSRV-26 and enJSRV-18. Maps were generated using the SeqBuilder program (Lasergene 9 Core suite). Relevant restriction sites are indicated. f1 (+) ori, phage origin of replication; ampicillin, antibiotic resistance; pUC ori, prokaryote origin of replication; P lac, lactose promoter; MCS, multiple cloning site; CMV, cytomegalovirus promoter; PBS, primer binding site.

Mutants enJSRV-26ΔEnv and enJSRV-18ΔEnv contain two premature termination codons at the first and third methionine of *env*, derived by site-directed mutagenesis. As a result, enJSRV-26ΔEnv and enJSRV-18ΔEnv encode the genomic sequence of the correspondent wild-type provirus except the Env glycoprotein.

Mutants enJSRV26-EnvD6A, enJSRV18-EnvA6D and JSRV-EnvA6D express the full-length viruses containing a single point mutation in their Env glycoproteins, and were derived by site-directed mutagenesis. The nomenclature of the mutants indicates the virus from which they were derived (enJSRV-26, enJSRV-18 or $JSRV_{21}$, respectively), followed by the protein mutated (the Env glycoprotein), a single letter indicating the amino acid mutated (A is for alanine and D for aspartic acid), a number representing its position in Env, and a letter indicating the amino acid residue in the resulting mutant.

Plasmids enJSEnv26, enJSEnv18 and enJSEnv56 express, respectively, the Env of enJSRV-26, enJSRV-18 and enJS56A1, and have already been described (Arnaud et al., 2007a). These mutants were obtained from the pCMV3JS21ΔGP expression plasmid (Maeda et al., 2001) by replacing the *env* and 5' LTR of $JSRV_{21}$ with those of each provirus.

pSP26-HA, pSP18-HA and pSP56-HA encode, respectively, the signal peptide of the Env glycoprotein (SP) of enJSRV-26, enJSRV-18 and enJS56A1 proviruses. These mutants were obtained by polymerase chain reaction (PCR) from the corresponding Env expression plasmids encompassing the signal peptide region, and were tagged with the hemagglutinin (HA) epitope (YPYDVPDYA).

In enJS18Env-PPT, the signal peptide of enJSRV-18 Env was replaced by the signal peptide of the human preprotrypsin protein (PPT), followed by the FLAG epitope fused at the N-terminus of enJSRV-18 surface domain (SU) of Env. enJS18Env-PPT has been described elsewhere (Caporale et al., 2009) and was provided by Claudio Murgia.

In enJS56A1-CTE, enJS56A1 *env* was replaced by four copies of the constitutive transport element (CTE) of Mason-Pfizer Monkey Virus (M-PMV) by restriction digestion with *BamHI*. The M-PMV CTE was derived from pSarm4 as already described, and was a gift of Eric Hunter (Rhee et al., 1990). This mutant was a courtesy of Marco Caporale.

Mutant enJS56A1ΔGag contains two nonsense mutations replacing the first and third methionine of enJS56A1 Gag, and was obtained by site-directed mutagenesis.

pNLgagSty330 and pRev were kindly provided by Barbara Felber and have already been described (Felber et al., 1989; Mermer et al., 1990). pNLgagSty330 is a Rev and Tat-dependent HIV-1 Gag-Pol expression plasmid and, for simplicity, is termed pHIV1-RRE in this study. pRev is an expression plasmid for HIV-1 Rev. The expression plasmid for HIV-1 Tat was a gift from Mauro Giacca, and has been described elsewhere (Vardabasso et al., 2008). pHIV-SPRE26 and pHIV-SPRE18 were obtained by PCR by replacing the Rev responsive element (RRE) in pHIV1-RRE with the signal peptide responsive elements (SPRE), encompassing *env* and 3'LTR, of enJSRV-26 and enJSRV-18, respectively. In order to avoid spurious expression of enJSRV-26 and enJSRV-18 SPs, PCR amplifications were carried out on enJSRV-26ΔEnv and enJSRV-18ΔEnv mutants, respectively. The expression plasmid for SPRE56, which was termed pHIV-SPRE56 in this study, has been already described elsewhere (Caporale et al., 2009) and was a courtesy of Marco Caporale. This plasmid contains two premature termination codons in enJS56A1

env. After PCR amplification, fragments were digested with *EcoRI* and *XbaI* and ligated into the pHIV-1 RRE plasmid.

For cloning purposes, PCR was performed using PfuUltra II Fusion HS DNA Polymerase (Stratagene). For site-directed mutagenesis experiments, the QuickChange site-directed mutagenesis kit (Stratagene) was employed according to manufacturer's instructions. Restriction enzymes were purchased from Roche and New England Biolabs. Ligation reactions were carried out using T4 DNA Ligase (Roche). Plasmid DNA was produced in DH5α strain of *E. coli* (Invitrogen) or XLI-Blue Supercompetent Cells (Stratagene), using the DNA Maxiprep kit (Invitrogen). All of the constructs described above were sequenced to ensure the presence of the introduced mutations. A list of all the primers used in this study is reported in the appendix (table 3).

2.2 Cell cultures, transfections and viral preparations

293T and COS cells were cultured in Dulbecco's modified Eagle medium (GIBCO), supplemented with 10% foetal bovine serum at 37°C, 5% CO_2 and 95% humidity. For western blotting purposes, cells were transiently transfected with 2 μg of the appropriate plasmids in Petri dishes of 10 cm in diameter, unless otherwise indicated, using the Calphos mammalian transfection kit (Clontech), as recommended by the manufacturer. The empty vector pcDNA3.1 (Invitrogen) was used to calibrate the amount of DNA used in each experiment. Cells supernatants were collected at 48 hours post-transfection, and viral particles were concentred by ultracentrifugation (35000 rpm, SW41 Beckman centrifuge rotor, for 1 hour at 4°C on 29% sucrose cushion) and then resuspended in 1x TNE buffer, as already described (Arnaud et al., 2007a; Palmarini et al., 1999a). For analysis of intracellular proteins, cells were lysed by standard techniques as described previously (Arnaud et al., 2007a). Briefly, at 48 hours post-transfection, cells were washed once with cold phosphate buffer saline (PBS), and lysed on ice for 10 minutes with modified RIPA buffer (150 mM Tris-HCl, pH7.4; 1% NP-40; 1 mM EDTA; 150 mM NaCl; 1 μM PMSF; 1 mM NaF) supplemented with a cocktail of protease inhibitors (Complete™, Roche), according to the manufacturer's recommendations. Cells were then snap frozen in liquid nitrogen, thawed on ice, and sonicated. Lysates were further centrifuged at 10000 rpm (SW41 Beckman centrifuge rotor) for 30 minutes at

4°C, supernatants were collected, and protein concentration was determined by the method of Lowry (Lowry et al., 1951).

2.3 Western blotting

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting were performed on concentrated viral particles and cell lysates (15 μl of 200x concentrated virus and 200 mg of protein extracts), as previously described (Arnaud et al., 2007a; Palmarini et al., 1999a). Briefly, after SDS-PAGE, proteins were transferred to nitrocellulose membranes (Hybond, Amersham), and blocked for 1 hour at room temperature with blocking buffer (5% skimmed milk in TBS/T [0.1% Tween 20 in TBS]). Membranes were then rinsed with TBS/T three times for 5 minutes, and incubated with the selected primary antibody. This step was performed either for 1 hour at room temperature or overnight at 4°C. For western blotting, all primary and secondary antibodies employed in this study were diluted in blocking buffer. enJSRVs/JSRV Gag protein was detected using a rabbit polyclonal serum against the JSRV major capsid protein (CA) (working dilution: 1:5000) (Mura et al., 2004). A rabbit polyclonal serum towards the JSRV transmembrane protein (TM) was employed to detect Env proteins (working dilution: 1:100). Signal peptides tagged with the HA epitope were detected with a mouse monoclonal anti-HA antibody (Abcam) (working dilution: 1:3000), while γ-tubulin was detected with a rabbit polyclonal antibody (Sigma) (working dilution: 1:1000). After incubation with the primary antibody, membranes were rinsed three times in TBS/T for 5 minutes, and further exposed to the appropriate peroxidase-conjugated secondary antibodies for 1 hour at room temperature. Donkey anti-rabbit (F[ab']2 fragment) (Amersham, GE) (working dilution: 1:50000) and donkey anti-mouse (IgG F[ab']2) (Fitzgerald) (working dilution: 1:5000) were used as secondary antibodies. If membranes were exposed again to a different primary antibody, they were stripped with Restore Western Blot Stripping Buffer (Thermo Scientific) for 1 hour at 37°C, and washed three times with TBS/T for 5 minutes. Membranes were subsequently developed by using ECL Plus (Amersham) and signals were further quantified by measuring chemiluminescence in a Molecular Dynamics Storm 840 imaging system, employing the ImageQuant TL software (Molecular Dynamics). Each experiment was repeated independently at least three times, and results are presented as the mean value for each sample (±

standard error). P values were calculated using the 2-tailed Student's T-test, assuming that the two groups had equal variance.

2.4 Immunoprecipitation

Immunoprecipitation was performed essentially as already described (Murcia et al., 2007). Briefly, 293T cells were transfected with 2 μg of the appropriate plasmids, as indicated in Results. 48 hours after transfection, cells were lysed as described above. Protein extracts were then sonicated and quantified. 2 mg of whole-cell extract were rocked with 20 μl of anti-FLAG M2 Affinity Gel (Sigma) and incubated at 4°C for 1 hour. The beads were then washed three times with the IP buffer (20 mM Hepes buffer, pH 7.4; 150 mM NaCl; 0.1% Triton), supplemented with the protease inhibitor cocktail and resuspended in 50 μl of IP buffer or 3X FLAG epitope (Sigma). Samples were then boiled for 5 minutes and subjected to SDS-PAGE gradient gels (NuPAGE 4-12% Bis-Tris Gel; Invitrogen) and western blot analysis, as described above. The FLAG epitope was detected with a mouse monoclonal anti-FLAG (Sigma) antibody (working dilution: 1:5000).

2.5 Confocal microscopy

293T and COS cells were plated onto two well-chambered glass slides (Lab-Tek; Nalge Nunc International) and transfected with 2 μg of the appropriate plasmids, employing Lipofectamine supplemented with PLUS Reagent (Invitrogen), according to the manufacturer's instructions. 24 to 48 hours after transfection, cells were washed with PBS and fixed with 3% v/v formaldehyde for 15 minutes. After fixation, cells were processed essentially as already described (Mura et al., 2004; Sfakianos and Hunter, 2003). Briefly, cells were permeabilized with PBS containing 0.1% Triton X-100 for 10 minutes at room temperature. Cells were then blocked for 5 minutes at room temperature, firstly with PBS containing 0.4% fish skin gelatine and 0.2% Tween 20, and secondly with PBS containing 2.5% normal goat serum and 0.2% Tween 20. Next, cells were incubated with the appropriate primary antibody overnight at 4°C. SP proteins tagged with the HA epitope were detected with mouse monoclonal anti-HA (Abcam) antibody (working dilution: 1:800). Rabbit polyclonal antibody to fibrillarin (Abcam) was used as a marker for the nucleolus (working dilution: 1:300). Cells were then washed three times with PBS containing 0.2% Tween 20 and blocked as described

above. Subsequently, cells were incubated for 45 minutes at 37°C with the appropriate secondary antibody. Goat anti-mouse and anti-rabbit immunoglobulin G conjugated respectively with Alexa Fluor 488 and Alexa Fluor 594 (Molecular Probes) were used as secondary antibodies (working dilutions: 1:3000). For confocal analysis, all primary and secondary antibodies were diluted in PBS containing 2.5% normal goat serum and 0.2% Tween 20. Finally, cells were washed three times, firstly with PBS containing 0.2% Tween 20, and secondly with PBS. Slides were mounted with medium containing DAPI (4',6'-diamidino-2-phenylindole; Vectashield, Vector Laboratories) and analyzed with a Leica TCS SP2 confocal microscope. Single sections from confocal optical sections along the z axis were analyzed.

2.6 Enzyme-linked immunosorbent assay (ELISA)

ELISA assay was performed essentially as already described (Caporale et al., 2009). Briefly, 293T cells were transfected with 1 μg of the appropriate plasmid (or 0.2 μg for HIV-1 Tat), as illustrated in Results. 48 hours after transfection, 100 μl of cell supernatants were collected, filtered through a 0.45 μm microfilter (Millipore) and assessed for the presence of HIV-1 Gag proteins using a Murex HIV Antigen mAb kit (Abbot Murex), according to the manufacturer's instructions. Absorbance was read at 450 nm with a microplate reader. Each experiment was repeated independently at least three times, and results are presented as the mean value for each sample (± standard error).

2.7 Quantitative polymerase chain reaction (qPCR)

2.7.1 Overview of the assay

Quantitative PCR (qPCR) is a technique that allows accurate amplification and quantification of DNA. The assay is based on the PCR procedure, with the advantage that the amplified products are detected in "real time", and not at the end of the reaction as for the standard technique. This permits a more sensitive quantification of target molecules (Higuchi et al., 1992). The amplified products can be detected employing fluorescence-labelled probes, such as TaqMan or molecular beacons probes, or DNA binding dyes, like SYBR Green. Both methods rely on the emission of fluorescence that is directly proportional to the amount of target product present in the reaction at any given cycle

(Wilhelm and Pingoud, 2003). As the qPCR reaction proceeds, the quantity of amplified products, and thereby of fluorescence signals, increases in the exponential phase, proportional to the initial amount of target DNA. The point at which fluorescence crosses the threshold value (arbitrarily chosen within the exponential phase of the reaction) is termed "threshold cycle" (Ct). This value is inversely proportional to the starting quantity of the amplified target DNA, i.e., the higher is the value of the Ct, the smaller is the initial amount of DNA present in the reaction mixture (Kubista et al., 2006). qPCR products can be quantified by (i) relative quantifications, which determine the levels of a target gene in test samples by expressing it as relative to the levels in reference samples (e.g., untreated *versus* treated cells, normal *versus* cancer cells, etc.), or (ii) absolute quantifications, which employ an external standard curve. The standard curve is generated by amplifying multiple dilutions containing known amounts of target DNA and, subsequently, relating them to their respective Ct values. Such a standard curve is then used to extrapolate the quantity of DNA (or copy number) in unknown samples. In both relative and absolute quantifications, it is always recommended to perform an additional normalisation of the target gene with an endogenous standard gene. For this purpose, housekeeping genes, such as the glyceraldehyde-3-phosphate dehydrogenase (G3PDH or GAPDH), albumin, actins, tubulins, cyclophilin, 18S rRNA or 28S rRNA, are often used as reference genes (Mar et al., 2009; Popovici et al., 2009). In this study, the *β-actin* and *SOX9* were used as reference genes. They were selected among different functional classes, in order to reduce the chance that they might be co-regulated, thus affecting the accuracy of the results.

β-actin belongs to the actin family, a diverse group of proteins that includes α, *β* and γ isoforms (Garrels and Gibson, 1976). Actins are the most abundant proteins in eukaryotic cells, highly conserved from yeast to humans (Schmidt and Hall, 1998). They have also been found in plants as actomyosin complex, where they are involved in intracellular movements as well as in plant cell cycle (Kost et al., 2002; Takagi, 2003; Vorobev and Ia, 1963; Wick, 1991). Actins play important roles in different cellular processes, including cell motility and cytoskeleton organization (Hunter and Garrels, 1977). In particular, the β-actin, together with the γ-actin, is one of the major components of the microfilamentous structures in mammals and avian non-muscle cells (Clarke and

Spudich, 1977), and is involved in cytokinesis and cell motility (Weaver and Weissmann, 1979). Sequential mutations in the β-actin protein have been shown to be associated with incremental increase in human fibroblasts tumorigenicity (Leavitt et al., 1982).

SOX9 is a member of the *SOX* family, which consists of a large number of genes well conserved in vertebrates and invertebrates, closely related to the SRY (Sex Determining Region of Y chromosome) DNA-binding HMG (High Mobility Group) box (Denny et al., 1992; Gubbay et al., 1990; Wright et al., 1993). *SOX* genes family encodes transcription factors involved in several aspects of developmental processes, such as central nervous system differentiation, sex determination, and heart and kidney development (Bowles et al., 2000). In particular, *SOX9* appears to be a key regulator of skeletal development and sex determination, as mutations in a single allele cause a severe skeletal malformation syndrome (referred to as campomelic dysplasia in humans) and autosomal sex reversal (Bi et al., 2001; Foster et al., 1994). Moreover, it has been shown that heterozygous *SOX9* mutant mice die perinatally due to haploinsufficiency of the gene (Bi et al., 2001). Finally, individuals with trisomy for chromosome 17, including the region containing *SOX9*, show skeletal anomalies and central nervous system deficiencies (Lenzini et al., 1988).

2.7.2 Determination of the copy number variation of the transdominant enJS56A1 provirus in sheep genomic DNA

qPCR assays were designed in order to assess the copy number variation of target genes (enJS56A1, enJSRV-6 and enJSRV-18) relative to reference genes (*β-actin* and *SOX9*) in sheep genomic DNA. The copy number variation of each locus was determined by employing an absolute quantification method. To this end, external standard curves were generated. PCR products for enJS56A1 (532 bp), enJSRV-6 (577 bp), enJSRV-18 (497 bp), *β-actin* (438 bp) and *SOX9* (438 bp) were amplified from sheep genomic DNA, employing antisense primers complementary to the genomic 3' flanking region of each provirus[20], and sense primers designed in a conserved region in *env*[21]. Primers for *β-actin*[22] and *SOX9*[23] were designed on

[20] The antisense primers used to amplify the enJSRV-6 and enJSRV-18 proviruses have been already described (Arnaud et al., 2007a).
[21] The *env* sequences of enJSRV-6, enJSRV-18 and enJS56A1 proviruses have been already reported and deposited in GenBank, under accession numbers EF680319.1, EF680301.1 and AF153615.1, respectively (Arnaud et al., 2007a; Palmarini et al., 2000).

highly conserved regions of both genes. A list of all the primers used in this study is reported in the appendix (table 4). qPCR assays were carried out in triplicate in a total volume of 25 μl, and performed in a Mx30005P (Stratagene) thermocycler, using the Brilliant II SYBR Green QPCR Low ROX Master Mix (Stratagene) and the Brilliant SYBR Green QPCR Core Reagent Kit (Stratagene), according to the manufacturer's instructions. The reaction mixture contained 20 ng of sheep genomic DNA. A negative control, containing all reagents but no genomic DNA, was also used to check for contamination by exogenous DNA. The reaction was subjected to a denaturation step at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 30 seconds, primer annealing at the temperature appropriate for each primer (55°C to 59°C for 30 seconds), and elongation at 72°C for 35 seconds, ending with a melting curve analysis to validate the specificity of the PCR products. The amplified fragments were then cloned into a PCR-4 TOPO vector (Invitrogen) and sequenced. qPCR reactions were performed on serial dilutions of plasmid DNA encoding for each locus (from 10^2 to 10^7 copies) and containing a known number of molecules, and run in parallel with the samples of sheep genomic DNA. Reactions ended with a dissociation curve to validate the specificity of the PCR products. The presence of a single peak in the dissociation curve is indicative of specific amplification of target genes, whereas multiple extra peaks are usually due to contaminating DNA or primer dimers.

Below are reported the amplification plots (Fig. 16), and the dissociation curve (Fig. 17) of serial dilutions of the *β-actin* expression plasmid, used to generate the standard curve for the reference gene.

[22] Sense and antisense primers were designed using the *β-actin* sequence of *Ovis aries* deposited in GenBank under accession number NM_001009784.1.
[23] Oligonucleotide primers were designed in the 3'UTR of the ovine *SOX9* gene (Payen and Cotinot, 1994).

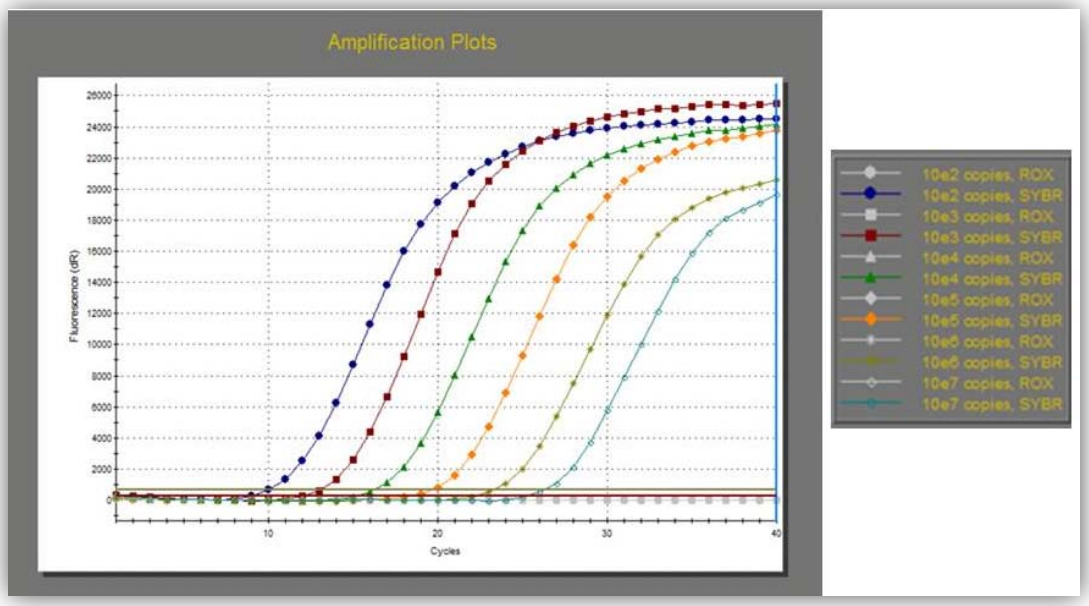


Figure 16 | Amplification plots of the *β-actin* gene standard curve. The curve was generated on serial dilutions (from 10^2 to 10^7 copies) of plasmid DNA containing the *β-actin* gene, as described in the text. The fluorescence data generated in the amplification plots are expressed as number of cycles (X-axis) *versus* fluorescence (Y-axis), and are baseline-corrected raw fluorescence (dR). ROX dye was used as a negative control for the qPCR assay.

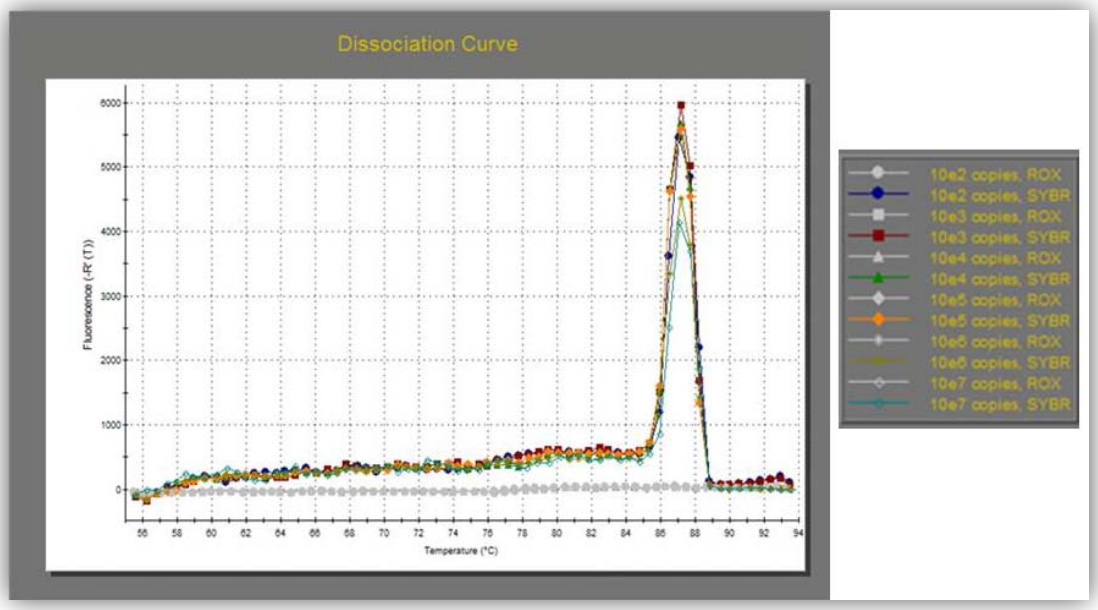


Figure 17 | Dissociation curve of serial dilutions (from 10^2 to 10^7 copies) of plasmid DNA encoding the *β-actin* gene. Dissociation curve analysis was performed at the end of PCR cycles. Data were obtained by increasing the temperature of reaction solutions from 55°C to 95°C, while continuously collecting fluorescence data. The increase in temperature causes PCR products to undergo denaturation, thereby reducing the fluorescence of specific products containing the SYBR Green dye. The fluorescence data generated in the dissociation curve are expressed as a function of fluorescence (Y-axis), and are baseline-corrected raw fluorescence (dR). ROX dye was used as a negative control for the qPCR assay.

Copy number of standard DNA molecules was calculated by the MxPro QPCR Software (Stratagene), using the following formula (Godornes et al., 2007):

$$Number\ of\ copies_{/\mu l} = \frac{6.022 \times 10^{23}\left(molecules_{/mole}\right) \times DNA\ concentration\left(g_{/\mu l}\right)}{Number\ of\ bases\ pairs \times 660\ daltons}$$

The resulting Ct values were plotted as a function of the $\log_{10}$ concentration of the input plasmid amount. The slope of the line was used to determine PCR efficiency and the unknown fold change of each locus in sheep genomic DNA. The equation that relates the slope to the amplification efficiency is:

$$PCR\ Efficiency = 10^{\left(-1_{/slope}\right)} - 1$$

where PCR Efficiency corresponds to the proportion of template molecules that are doubled every cycle. From this equation it follows that a slope of –3.322 will result in a reaction with 100% efficiency, as for the *β-actin* gene standard curve (Fig. 18).

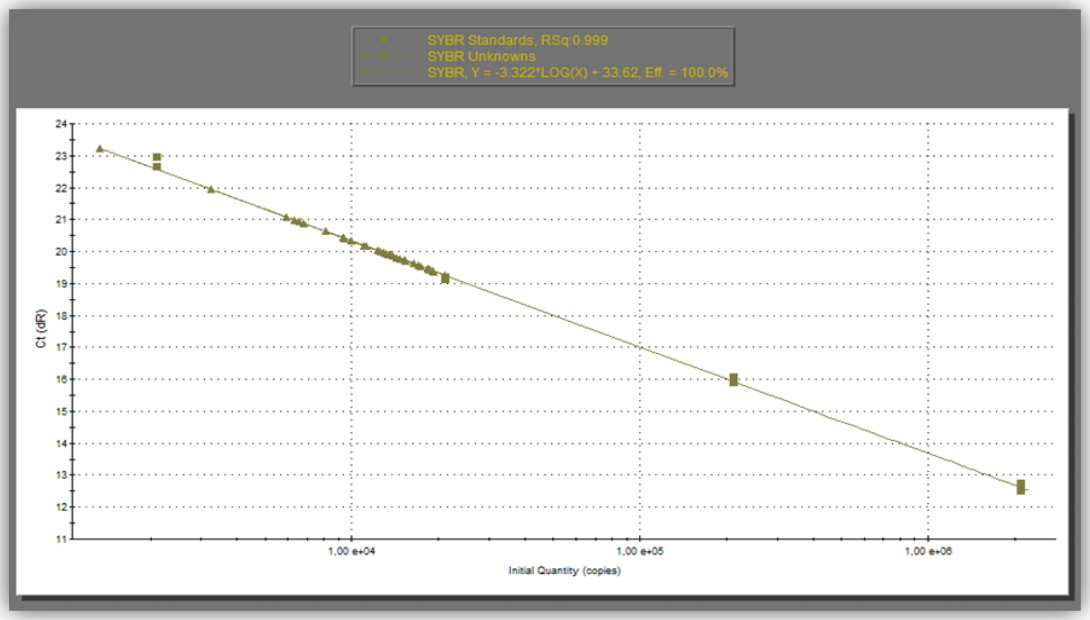


Figure 18 | Standard curve of the *β-actin* gene. Ct values are plotted on the standard curve as a function of the $\log_{10}$ concentration of the input amount of the expression plasmid containing the *β-actin* gene (SYBR Standards), and genomic DNA of the samples analyzed (SYBR Unknowns). The R Squared (RSq) value is an indicator of the quality of the fit of the standard curve to the standard data points plotted. The value is between 0 and 1: the closer the value is to 1, the better the fit of the line. The equation for the line is: $Y = m \times \log x + b$, where m is the slope of the line.

Results were expressed as the ratio between the estimated number of molecules in the target and reference genes in each sample, using the following formula (Bieche et al., 1998):

$$N = \frac{copy\ number\ of\ target\ gene}{copy\ number\ of\ reference\ gene}$$

2.7.3 Analysis of genomic DNA of wild and domestic sheep

qPCR assays were designed to estimate the dosage of the transdominant enJS56A1 provirus (target gene) relative to the *β-actin* and *SOX9* genes (reference genes) in genomic DNA of domestic and wild sheep. Gene dosage of enJSRV-6 and enJSRV-18 proviruses was also determined as additional control.

The genomic DNA tested was collected from various breeds of domestic sheep (*Ovis aries*) (Dorset, Suffolk, Texel, Jacob, Red Maasai, Merino, Xalda, Rambouillet, Soay, Norway and Finsheep), wild sheep (*O. dalli*, *O. canadensis*, *O. ammon* and *O. vignei*), Mediterranean mouflon (*O. orientalis musimon*), and members of the genera *Budorcas* (*B. taxicolor*) and *Pseudois* (*P. nayaur*). All DNA samples were obtained and used in a previous study (Arnaud et al., 2007a). To determine the fold change of each locus, qPCR assays were carried out as already described. Standard curve efficiency was 99.2% for enJS56A1, 96.6% for enJSRV-6, 99% for enJSRV-18, 100% for *β-actin,* and 100% for *SOX9*.

2.7.4 Analysis of genomic DNA of healthy and OPA affected domestic sheep

qPCR assays were designed to estimate the copy number variation of enJS56A1 in the genomic DNA of healthy and OPA-positive sheep, belonging to different flocks where OPA was endemic. The genomic DNA tested was collected from a total of fourteen sheep (seven healthy and seven OPA affected animals) belonging to the Blackface breed. To determine the copy number of the enJS56A1 locus in each animal, samples were subjected to an absolute qPCR analysis, as described above, employing *β-actin* as a reference gene. Standard curve efficiency was 92.9% for enJS56A1 and 99% for *β-actin*. All the samples used in this study were a courtesy of Dr. Chris Cousens at the Moredun Research Institute in Edinburgh. Note that the term “healthy sheep” refers to animals that did not show any clinical sign of OPA at the time of blood collection. By contrast, “OPA affected sheep” stands for animals that developed OPA and were found positive for JSRV by PCR on lung tumour samples.

2.8 Genotyping enJS56A1-like proviruses

The presence of a codon encoding an arginine or tryptophan residue at position 21 of enJS56A1 (and enJS56A1-like proviruses) Gag protein was assessed by PCR. Genomic DNA samples collected from various breeds of domestic (Texel, n=2; Merino, n=1) and wild sheep (*O. dalli*, n=1; *O. canadensis*, n=1; *O. ammon*, n=1), and the Mediterranean Mouflon (*Ovis orientalis musimon*, n=1) were amplified by PCR using a forward primer complementary to the genomic 5' flanking region of enJS56A1 or enJSRV-20 proviruses, and a reverse primer complementary to their *gag*, as previously described (Arnaud et al., 2007a). PCR reactions were carried out in a total volume of 50 μl, and contained 100 ng of sheep genomic DNA and the HotStar Taq DNA polymerase system. Standard amplification cycles and annealing temperature at 55°C were used as recommended by the manufacturer (Qiagen). PCR products were then cloned into a pCR4-TOPO vector (Invitrogen), and forty individual clones for each PCR product were completely sequenced.

2.9 Fluorescence *in situ* hybridization (FISH)

FISH analysis was performed by Giulia Pia di Meo in the Leopoldo Iannuzzi laboratory in Naples (Italy). The assay was carried out essentially as already described (Chessa et al., 2009). Briefly, sheep peripheral blood cells were cultured at 37°C in RPMI 1640 medium (Invitrogen), supplemented with 15% foetal bovine serum, 1.5% concanavalin A (Sigma) and penicillin-streptomycin (Invitrogen). Cells were then synchronized with 300 μg/ml thymidine (Sigma) and, after 18 hours, washed and resuspended in medium containing 15 μg/ml bromodeoxyuridine (Sigma) and 30 μg/ml Hoechst 33258 (Invitrogen). Cells were then incubated for 6 hours at 37°C (the last hour in presence of 0.1 μg/ml colcemid, Sigma), treated with a hypotonic solution, and washed three times with methanol-acetic anhydride. Cell suspensions were then plated onto slides, incubated overnight at 50°C, and stained for 10 minutes with 25 μg/ml Hoechst 33258. Slides were further probed with biotin-labelled BAC clones containing the appropriate provirus. Hybridization, chromosome staining, signal detection and image processing were performed as already described (Chessa et al., 2009) in at least thirty methaphases for each probe. Chromosome identification was carried out using the R-banding karyotype, by adding Fluorescein Avidin DCS and Biotinylated Anti-Avidin (Vectors Laboratories), as recommended by the

manufacturer. Chromosome identification and band nomenclature followed the International System for Chromosome Nomenclature of Domestic Bovids (ISCNDB2000, 2001).

Chapter III

Molecular determinant of JLR escape

3.1 Introduction

As already discussed in paragraph 1.10.3, enJS56A1 can block JSRV release by a mechanism known as JLR (for *J*SRV *l*ate *r*estriction) (Mura et al., 2004). This restriction activity is unique compared to those exerted by other retroviral and cellular restriction factors (e.g., Fv1, TRIM5α and APOBEC) (Luban, 2010), as it occurs at late steps of the viral replication cycle. In a previous work, Manuela Mura and colleagues demonstrated that enJS56A1 cannot release viral particles in the supernatant of transfected cells, despite abundant levels of Gag polyprotein in cell lysates. Electron microscopy experiments revealed that enJS56A1 can form virus-like particles that assemble in the cytoplasm, but they cannot reach the plasma membrane and bud (Mura et al., 2004). The defect in enJS56A1 exit was mapped to a trypthophan residue (W) in position 21 of Gag polyprotein, which substitutes an arginine (R) that is well conserved in all betaretroviruses (Mura et al., 2004). It has been suggested that the presence of W21 affects the general conformation of enJS56A1 Gag that, as a result, behaves like an unfolded or misfolded protein and is degraded by the proteasome. The W21R mutation confers to enJS56A1 a transdominant phenotype over JSRV: when co-expressed in the same cells, enJS56A1 Gag associates with JSRV Gag early after its synthesis, resulting in the formation of aggregates that are unable to traffic to the cell membrane, and are ultimately degraded by the proteasome (Arnaud et al., 2007b; Mura et al., 2004; Murcia et al., 2007).

A recent study conducted by Frédérick Arnaud and colleagues revealed that, during domestication, the sheep genome has been invaded by endogenous proviruses highly related to the exogenous and pathogenic JSRV, and hence termed enJSRVs. Five of the twenty-seven enJSRV loci (enJSRV-7, enJSRV-15, enJSRV-16, enJSRV-18 and enJSRV-26) display an intact genomic organization. Interference assays demonstrated that the transdominant enJS56A1 is able to block viral release from intact enJSRVs as efficiently as the exogenous JSRV. Interestingly, enJSRV-26, which is the "youngest" enJSRV locus isolated to date, possesses the unique ability to escape JLR (Arnaud et al., 2007a).

In this chapter are presented the experimental evidence that led to the characterization of the molecular determinant that allows enJSRV-26 to escape the restriction mechanism exerted by enJS56A1.

3.2 Results

3.2.1 The 5' end of enJSRV-26 env contains the determinant of JLR escape

The first aim of this study was to identify the molecular determinant that allows enJSRV-26 to escape JLR. Sequence analysis revealed that, at the nucleotide level, enJSRV-26 is 98% identical to another insertionally polymorphic provirus in the sheep genome, enJSRV-18. In particular, enJSRV-26 and enJSRV-18 *gag* are 100% identical. However, despite their high degree of similarity, these viruses display different phenotypes in the presence of the transdominant enJS56A1: enJSRV-18 is restricted like the exogenous JSRV and other enJSRVs, while enJSRV-26 is able to escape JLR (Arnaud et al., 2007a) (Fig. 19).

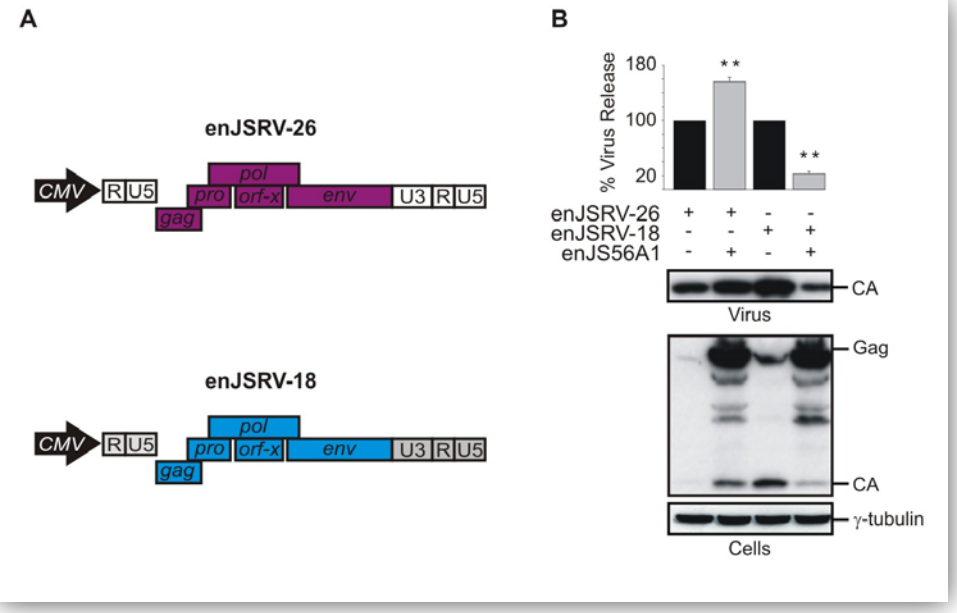


Figure 19| enJSRV-26 is able to escape JLR. (A) Schematic representation of expression plasmids encoding enJSRV-26 and enJSRV-18. CMV, cytomegalovirus immediate early promoter; R, terminally repeated sequences; U5, unique 5' sequence; U3, unique 3' sequence. (B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids (2 μg) indicated in the panel. Membranes were incubated with antibodies against the major capsid protein of JSRV (CA) or γ-tubulin as loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by the wild-type enJSRV-26 or enJSRV-18 expressed in the absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences ($P<0.01$) are indicated with two asterisks.

As JLR depends on the interaction between transdominant and functional Gag proteins, one would expect that the lower the levels of functional Gag, the better the efficiency of JLR. However, it is noteworthy that enJSRV-26, which

exhibits lower levels of Gag compared to enJSRV-18, is able to elude the restriction mechanism induced by enJS56A1 (Fig. 19). In addition, even by normalizing the levels of enJSRV-26 and enJSRV-18 Gag, interference assays revealed that the latter is still restricted by enJS56A1 (Fig. 20).

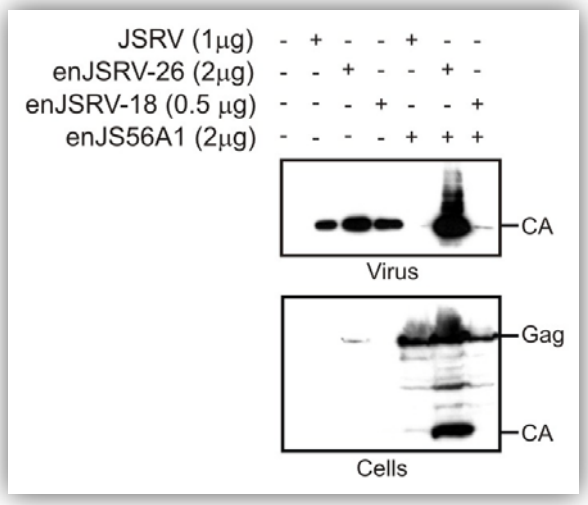


Figure 20| The amount of functional Gag protein does not determine JLR escape. Western blot analysis of concentrated supernatants (virus) and cell extracts (cell) of 293T cells co-transfected with fixed amount of enJS56A1 expression plasmid (2 μg), and JSRV (1 μg), enJSRV-26 (2 μg), or enJSRV-18 (0.5 μg) expression plasmids. Membranes were incubated with an antibody towards the major capsid protein of JSRV (CA). Note that JSRV and enJSRV-18 are restricted by enJS56A1 albeit they express the same levels of Gag as enJSRV-26.

These results indicate that the relative amount of functional Gag does not necessarily influence the ability to escape JLR. Indeed, similar studies showed that JSRV is still restricted by the transdominant enJS56A1 even when the ratio between the two expression plasmid is 10 (JSRV) to 1 (enJS56A1) (Mura et al., 2004).

Sequences alignments revealed that the major differences between enJSRV-26 and enJSRV-18 are found in the 3'LTR (Fig. 21) and *env* (Fig. 22).

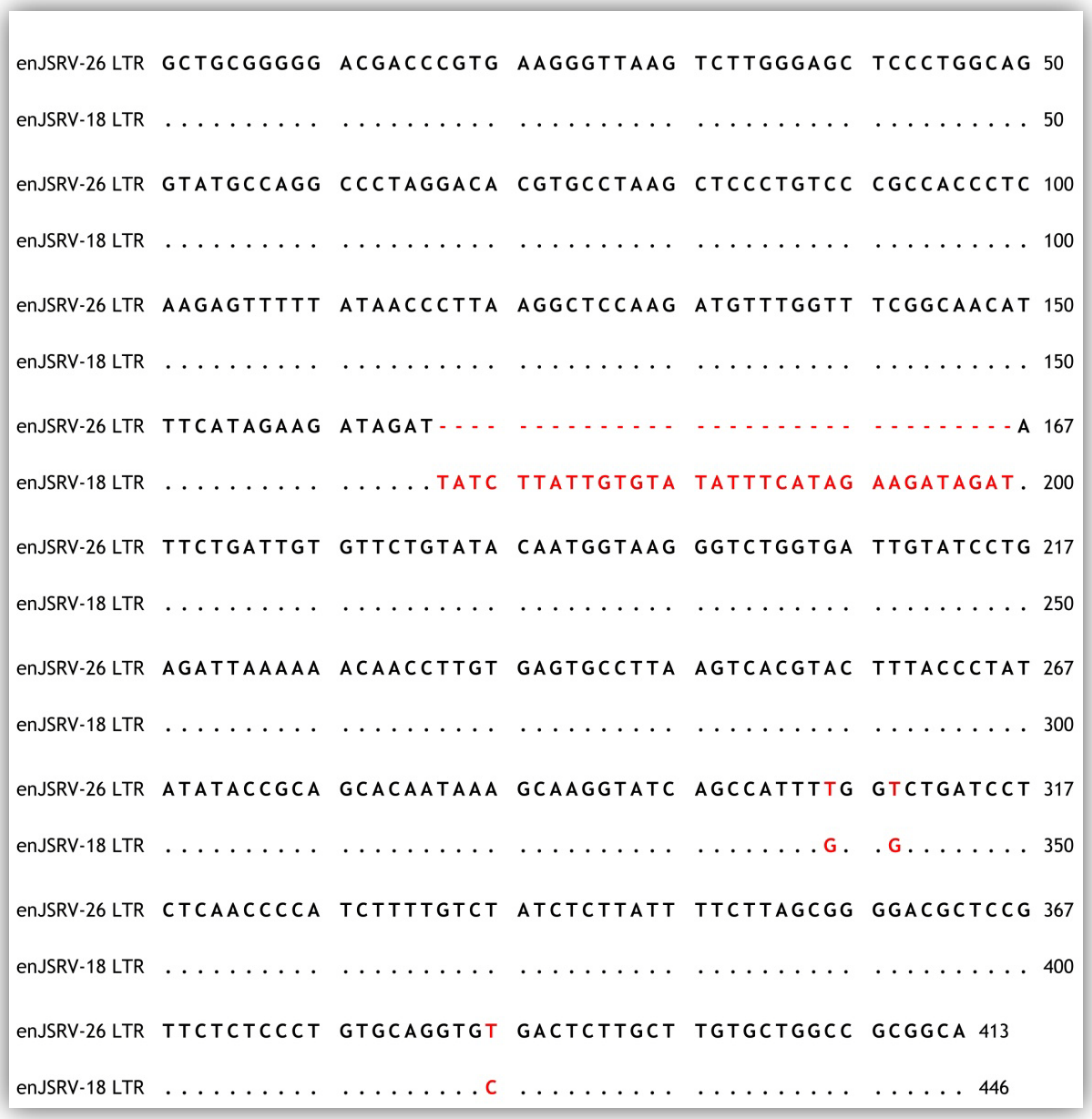


Figure 21 | Alignment of enJSRV-26 and enJSRV-18 LTR sequence. Sequence alignment was generated by CLC Sequence Viewer program version 6.5.1. Identical nucleotides are represented as dots, while different nucleotides are reported in red. The deletion of thirty-three base pairs exhibited by enJSRV-26 LTR is indicated as a red dashed line. Numbering starts from the U3 sequence of the LTR. In enJSRV-26, U3, R and U5 are located respectively between nucleotide positions 1 and 287, 288 and 300, 301 and 413. In enJSRV-18, the corresponding sequences span from nucleotide positions 1 and 320, 321 and 333, and 334 and 446 respectively. The enJSRV-26 and enJSRV-18 LTR sequences reported in the figure have been deposited in GenBank, under accession numbers EF680297.1 and EF680301.1, respectively.

As shown in figure 21, enJSRV-26 and enJSRV-18 LTRs differ for three nucleotides located within the U5 sequence. In addition, enJSRV-26 LTR exhibits a deletion of thirty-three base pairs in the U3 region, with respect to the corresponding sequence in enJSRV-18 LTR.

The *env* genes of enJSRV-26 and enJSRV-18 differ for ten nucleotides, resulting in six synonymous and four nonsynonymous mutations (Fig. 22).

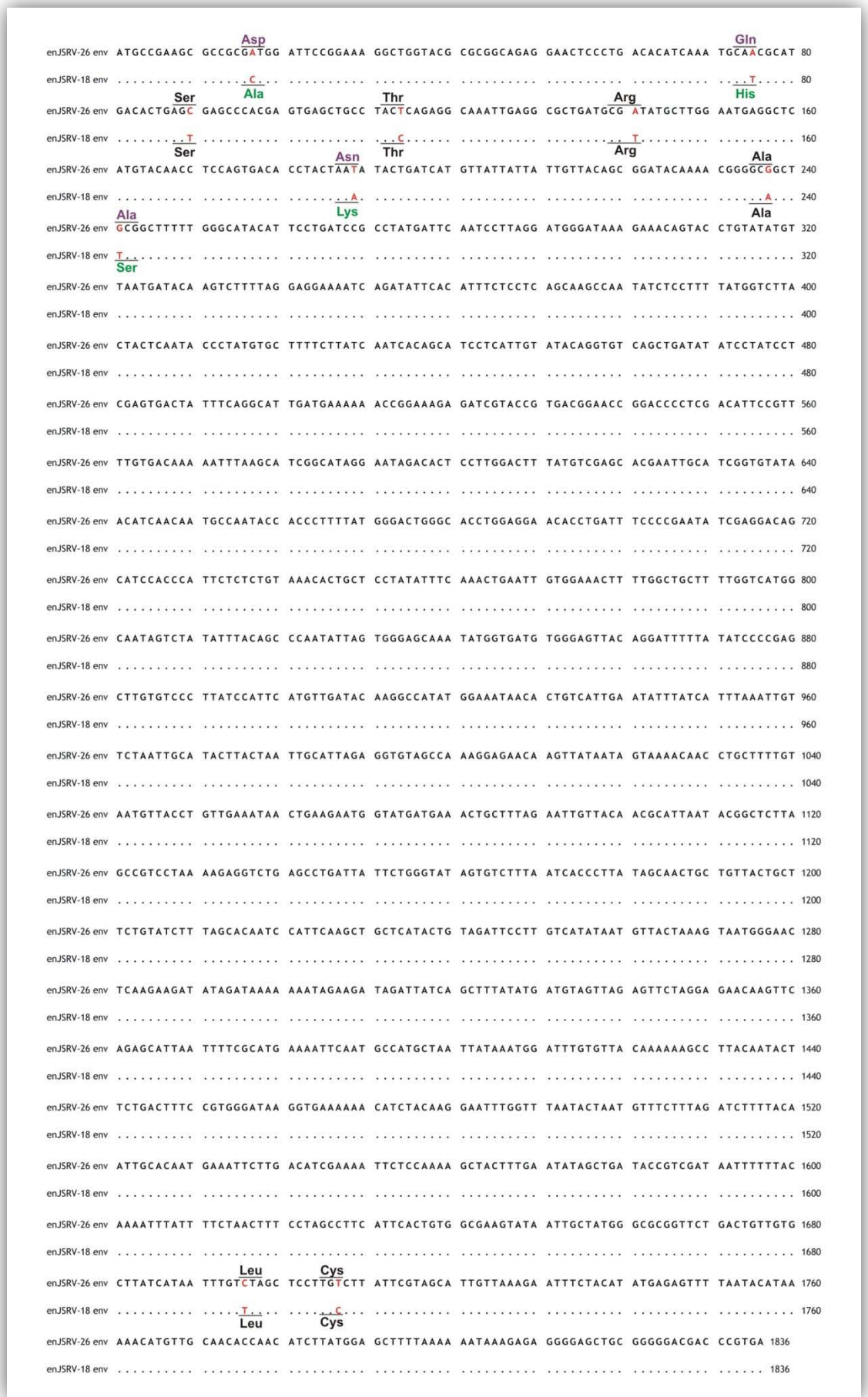


Figure 22 | Alignment of enJSRV-26 and enJSRV-18 *env* sequence. Legend on next page.

(From previous page) Sequence alignment was generated by CLC Sequence Viewer program version 6.5.1. Nucleotide numbering is from transcription start site. Identical nucleotides are represented as dots, while different nucleotides are indicated in red. The amino acid residues resulting from the four nonsynonymous (indicated respectively in violet and green in the resulting enJSRV-26 and enJSRV-18 Env glycoproteins) and six synonymous mutations (in black) are reported above each relevant codon. The enJSRV-26 and enJSRV-18 *env* sequences reported in figure have been deposited in GenBank, under accession numbers EF680297.1 and EF680301.1, respectively. Ala.

In particular, four of the six synonymous mutations are located at the 5' end of *env*, while the two remaining ones lie at the 3' end of the sequence. Conversely, all of the four nonsynonymous mutations are found in the 5' end of *env* (Fig. 22).

A synonymous mutation, also known as a silent mutation, is a base change in the DNA coding sequence that does not affect the amino acid sequence of the resulting protein. Indeed, these changes usually occur at the third position of the codon and, due to the redundancy of the genetic code[24], do not alter the polypeptide sequence. For instance, both AGC (enJSRV-26 *env*) and AGT (enJSRV-18 *env*) codons, located between nucleotide positions 98 and 100, encode a serine (Ser), while the threonine (Thr) is encoded by the ACC (enJSRV-26 *env*) and ACT (enJSRV-18 *env*) triplets spanning from nucleotide positions 122 to 125. The CGA (enJSRV-26 *env*) and CGT (enJSRV-18 *env*) codons, located between nucleotide positions 149 and 151, encode an arginine (Arg), whereas the alanine is encoded by the GCG (enJSRV-26 *env*) and GCA (enJSRV-18 *env*) codons between nucleotide positions 235 and 237. Both CTA (enJSRV-26 *env*) and TTA (enJSRV-18 *env*) triplets, spanning from nucleotide positions 1696 to 1698, encode a leucine (Leu), while the TGT (enJSRV-26 *env*) and TGC (enJSRV-18 *env*) codons, at nucleotide positions 1707-1709, encode a cysteine (Cys) (Fig. 22).

A nonsynonymous mutation, on the other hand, is a point mutation, which causes insertion of a different amino acid into the growing polypeptide chain, giving rise to an altered protein. For instance, the codon GAC spanning from nucleotide positions 16 and 18 in enJSRV-26 *env*, which encodes an aspartic acid (Asp), in enJSRV-18 *env* is mutated to GCT, which codes for an alanine (Ala). The CAA codon in enJSRV-26 *env* (nucleotide positions 73-75) encodes a glutamine (Gln) and is replaced with a CAT nucleotide triplet in enJSRV-18 *env*, which codes for a histidine (His). Between nucleotide positions 197 and 199, enJSRV-26 *env* possesses the AAT codon (asparagine, Asn) which in enJSRV-18 *env* is

[24] Several codons can encode for the same amino acid residue.

mutated to AAA (lysine, Lys). Finally, the GCG codon in enJSRV-26 *env*, at nucleotide positions 251-253, encodes an alanine (Ala) and is replaced with the TCG nucleotide triplet (serine, Ser) in enJSRV-18 *env* (Fig. 22).

As illustrated above, the major differences between enJSRV-26 and enJSRV-18 are found in their *env* genes and 3'LTR sequences, therefore it could be possible that these regions might contain the main determinant of enJSRV-26 JLR escape. In order to experimentally prove this hypothesis, a series of chimeras between enJSRV-18 and enJSRV-26 *env* genes and/or 3'LTR sequences were derived, and interference assays were carried out in presence of the transdominant enJS56A1 (Fig. 23). The release of viral particles from enJS26-Env18 was restricted in presence of enJS56A1, while enJS18-Env26 escaped JLR (Fig. 23A). Furthermore, enJS18-5'Env26 and enJS26-3'Env18 chimeras were able to elude enJS56A1-induced restriction, whereas enJS26-5'Env18 and enJS18-3'Env26 virus exit was impaired by enJS56A1 (Fig. 23, panels B and C). Finally, enJS26-3'LTR18 and enJS26-3'Env18LTR chimeras were able to escape JLR, whereas the presence of enJS56A1 restricted viral particles release from enJS18-3'LTR26 and enJS18-3'Env26LTR (Fig. 23, panels D and E). Collectively, these results demonstrate that the 5' end of enJSRV-26 *env* contains the main determinant of JLR escape.

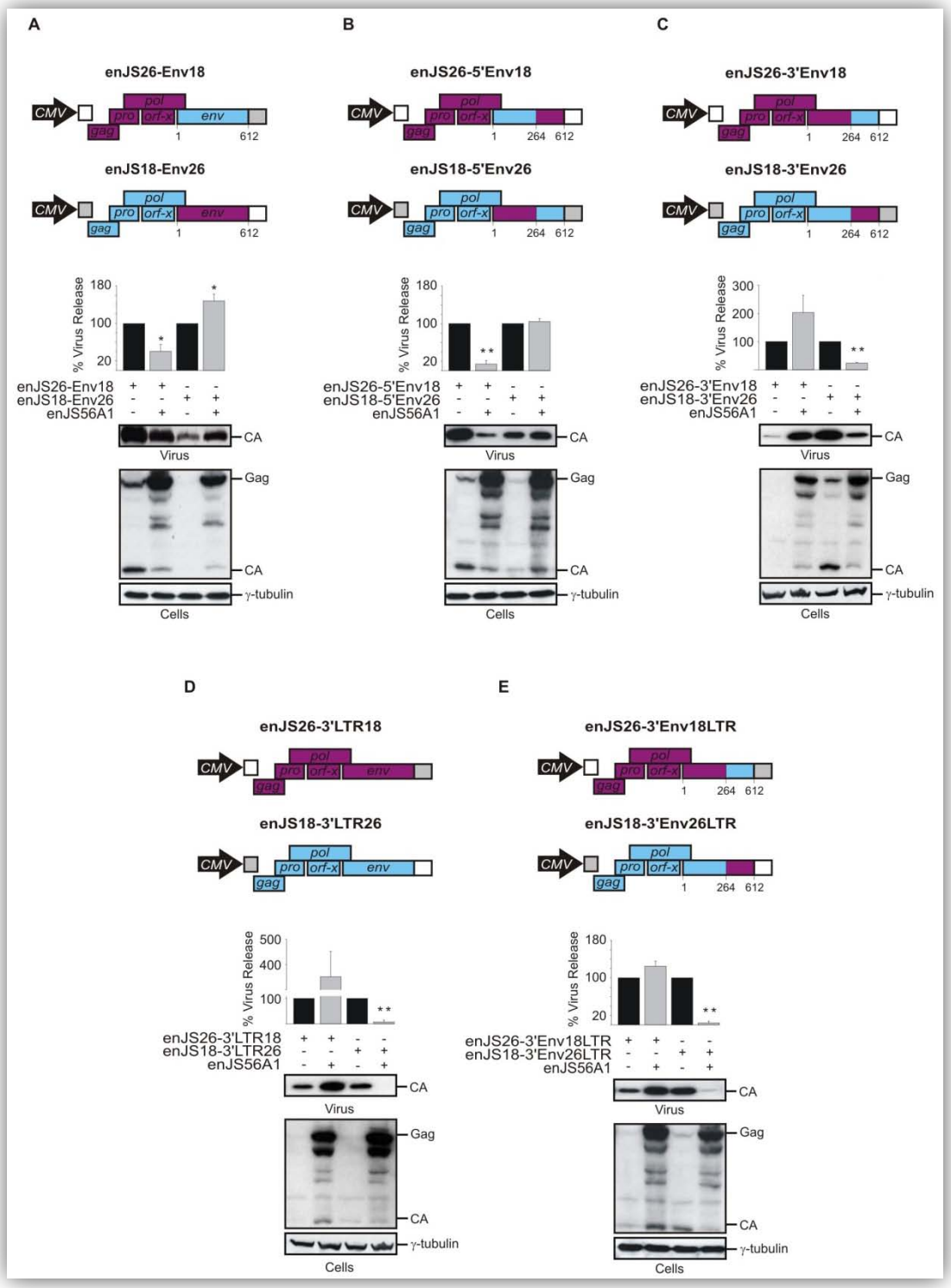


Figure 23 | The 5' end of enJSRV-26 *env* contains the main determinant of JLR escape. (A-E) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated and schematically represented above each panel. Numbers indicate amino acid residues in Env. For simplicity, the U5, R and U3 regions of enJSRV-26 and enJSRV-18 are indicated as white and grey boxes respectively. Membranes were incubated with antibodies towards the major capsid protein of JSRV (CA) or γ-tubulin as loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. Note that for the graphs of panels C and D different scales were adopted. The values obtained by each chimera expressed in absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P<0.05$) or two ($P<0.01$) asterisks. Only those chimeras containing the 5' end of enJSRV-26 *env* are able to escape enJS56A1 restriction. CMV, cytomegalovirus immediate early promoter.

3.2.2 The enJSRV-26 Env protein per se induces JLR escape

The data presented above demonstrated that the main determinant of JLR escape is located at the 5' end of enJSRV-26 *env*. However, it was not possible to discriminate whether the enJSRV-26 Env protein *per se*, or *cis*-acting regions within the env mRNA, determined JLR escape. To this end, enJSRV-26ΔEnv and enJSRV-18ΔEnv mutants were generated by inserting two premature termination codons in the *env* coding region, in order to prevent Env glycoprotein expression while maintaining the full-length viral genome. Interference assays were then carried out with enJS56A1 and the Env mutants (enJSRV-26ΔEnv or enJSRV-18ΔEnv) in presence or absence of their corresponding Env proteins (enJSEnv26 and enJSEnv18, respectively) (Fig. 24).

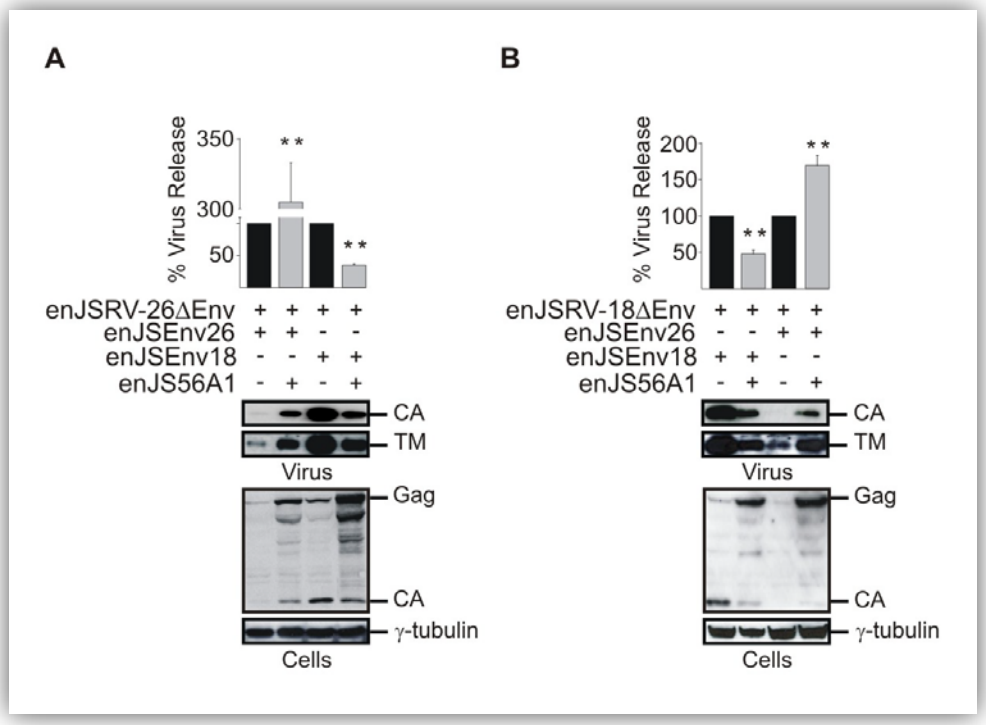


Figure 24 | The enJSRV-26 Env glycoprotein *per se* is involved in JLR escape. (A and B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the major capsid protein (CA) of JSRV or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by each mutant expressed in absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with two ($P<0.01$) asterisks. Env expression was controlled by incubating membranes with antibodies towards the transmembrane (TM) domain of JSRV.

enJSRV-26ΔEnv was able to escape JLR when its Env (enJSEnv26) was provided *in trans*, whereas it was restricted in presence of enJSRV-18 Env (enJSEnv18) (Fig. 24A). On the other hand, enJSRV18ΔEnv was restricted by enJS56A1 in presence

of its Env (enJSEnv18), while escaped JLR when co-expressed with enJSRV-26 Env (enJSEnv26) (Fig. 24B). These results indicate that the enJSRV-26 Env protein *per se*, and not *cis*-acting regions in its mRNA, is sufficient to determine JLR escape.

3.2.3 Aspartic acid at position 6 (D6) in enJSRV-26 Env is the main determinant of JLR escape

The results collected thus far shown that the 5' end of enJSRV-26 Env glycoprotein contains the determinant of JLR escape. As already discussed in chapter I (paragraphs 1.6.4.3 and 1.9.2), previous studies demonstrated that the N-terminal region of JSRV Env contains a signal peptide sequence (JSE-SP), which favours full-length viral RNA nuclear export and enhances Gag proteins synthesis and viral particle release (Caporale et al., 2009; Hofacre et al., 2009). Interestingly, enJSRVs were also shown to possess signal peptides located at the 5' end of their Env glycoproteins (residues 1 to 80) with the same domains and functions as JSE-SP (Caporale et al., 2009).

Analysis of the signal peptide sequences of enJSRV-26 and enJSRV-18 Env glycoproteins (SP26 and SP18, respectively) revealed that they possess three different amino acid residues (aspartic acid at position 6, D6; histidine at position 25, H25; lysine at position 63, K63) (Fig. 25). Interestingly, two other enJSRVs loci, enJSRV-7 and enJSRV-15, that are restricted like enJSRV-18 by enJS56A1, contain only one different residue (D6) from SP26. The alanine (A) at position 6 is well conserved in the signal peptides of JSRV and all the insertionally polymorphic enJSRVs, with the exception of enJSRV-26, where is replaced by an aspartic acid (D) residue (Fig. 25). Note that the A6D substitution in the signal peptide of the enJSRV-26 Env does not alter the hydrophobic profile of the glycoprotein [cfr. Fig. 26 and (Varela et al., 2009)].

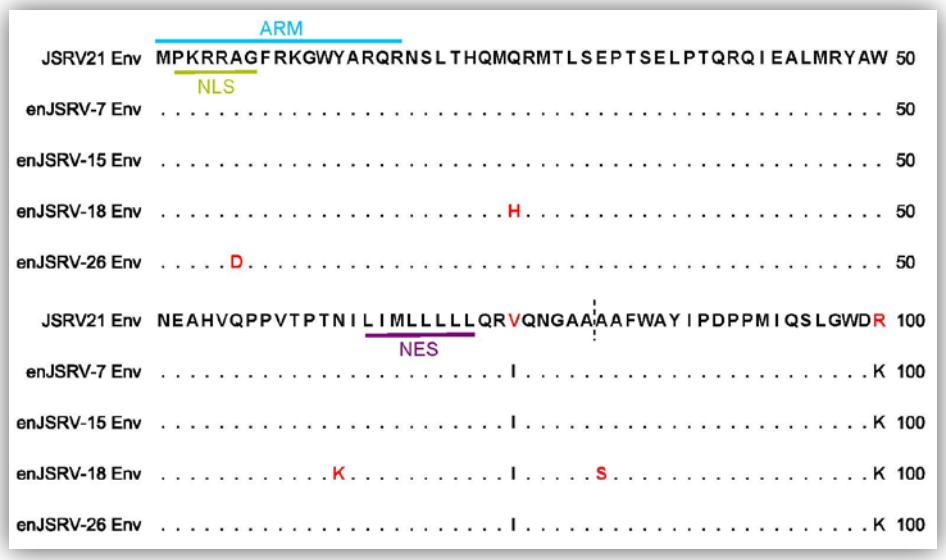


Figure 25| Amino acid sequence alignments of the signal peptide of four insertionally polymorphic enJSRV proviruses and the exogenous JSRV. Sequence alignments were generated by CLC Sequence Viewer, version 6.5.1 program. Amino acid sequences are reported according to the one-letter code. Nucleotide numbering is from transcription start site. Dots represent identical residues, while in red are indicated different amino acids. The dashed line represent the cleavage site for the signal peptidase. Complete sequences of Env glycoproteins have been deposited in GenBank, under accession numbers AAD45228.2 (JSRV Env), ABV71076.1 (enJSRV-7 Env), ABV71081.1 (enJSRV-15 Env), ABV71091.1 (enJSRV-18 Env) and ABV71071.1 (enJSRV-26 Env). ARM, arginine-rich motif; NLS, nuclear localization signal; NES, nuclear export signal.

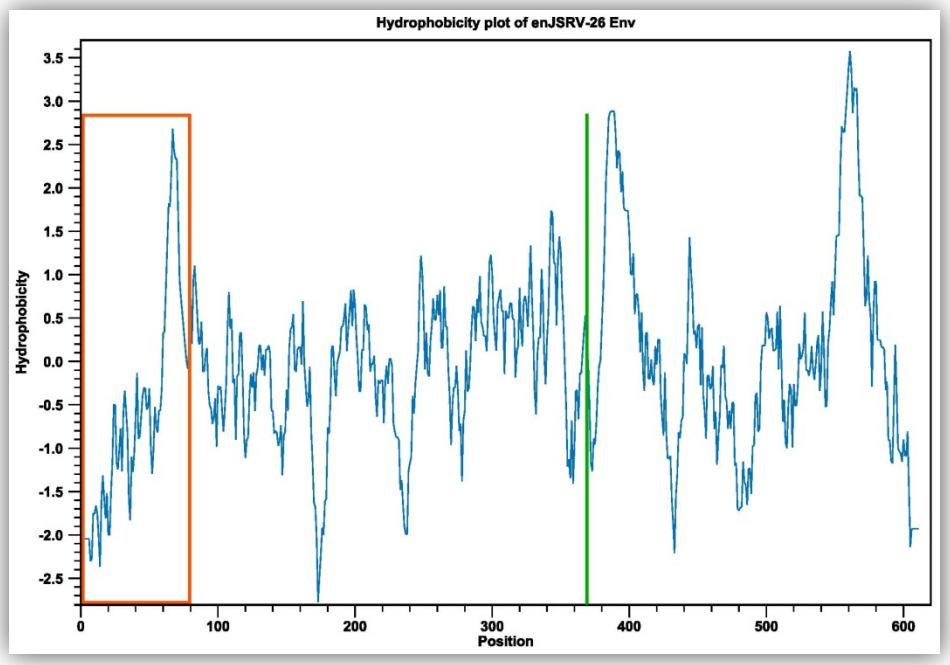


Figure 26| Hydrophobic profile of the enJSRV-26 Env glycoprotein. The hydrophobic profile of the enJSRV-26 was generated by CLC Sequence Viewer, version 6.5.1 program and calculated according to the Kyte and Doolittle method (Kyte and Doolittle, 1982). The orange box indicates the signal peptide, while the green line represents the consensus proteolytic cleavage site separating SU from TM.

Thus, it was possible that the D6 residue might play a critical role in enJSRV-26 JLR escape. In order to test this hypothesis, full-length single mutants were

derived by replacing the alanine at position 6 in enJSRV-18 Env with an aspartic acid (enJSRV18-EnvA6D), and mutating the aspartic acid at position 6 in enJSRV-26 Env to an alanine (enJSRV26-EnvD6A) (Fig. 27).

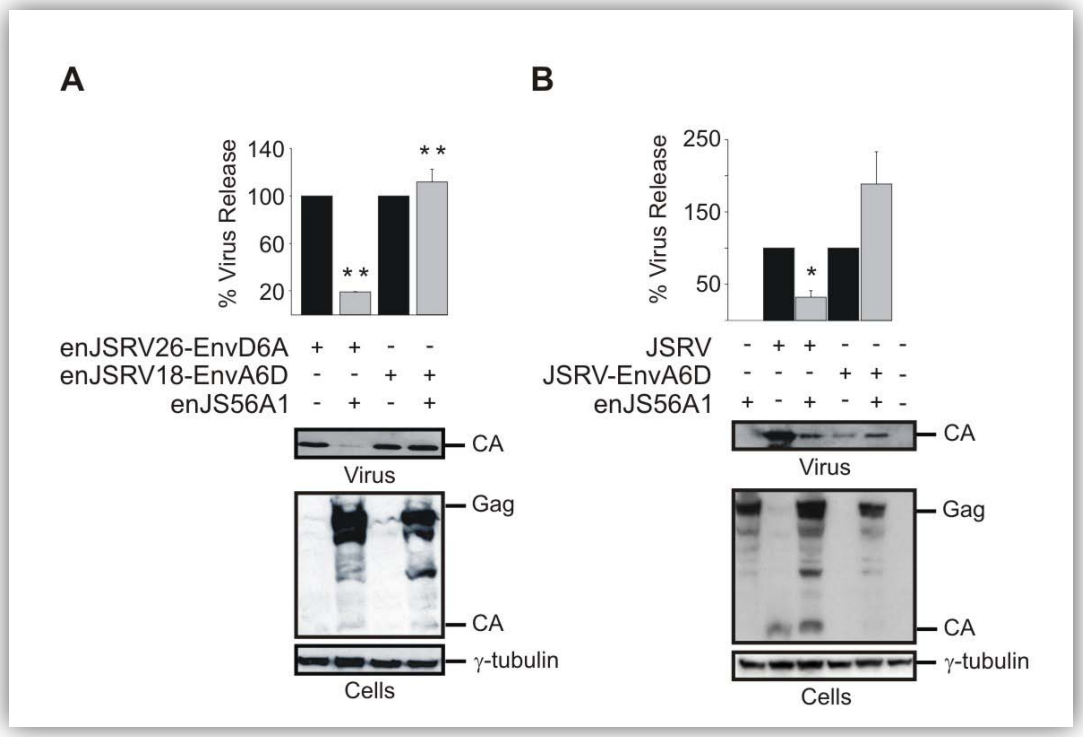


Figure 27 | Aspartic acid residue at position 6 (D6) in enJSRV-26 Env is responsible for JLR escape. (A and B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the major capsid protein (CA) of JSRV or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by each mutant expressed in absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P<0.05$) or two ($P<0.01$) asterisks.

The D6A mutation was found to confer to enJSRV-26 susceptibility to JLR, while the opposite mutation (A6D) allowed enJSRV-18 to escape enJS56A1-induced restriction (Fig. 27A). Interestingly, similar results were obtained for JSRV, normally restricted by enJS56A1 (Fig. 27B). Overall the data collected thus far demonstrate that the A6D mutation in SP26 is the main determinant of JLR escape.

3.3 Discussion

The data presented in this chapter demonstrate that the molecular determinant of JLR escape maps to the aspartic acid residue (D) at position 6 of the signal peptide of enJSRV-26 Env glycoprotein (SP26). The D6 residue substitutes an alanine (A6), well conserved in JSRV and all the other enJSRVs. The A6D mutation allows enJSRV-18 and JSRV to escape the restriction induced by enJS56A1, while the opposite mutation (D6A) renders enJSRV-26 susceptible to JLR.

The D6 residue lies within the N-terminus of enJSRV-26 Env, and indeed only those mutants containing this protein portion in their chimeric Env (enJS18-Env26, enJS18-5'Env26, enJS26-3'Env18, enJS26-3'LTR18 and enJS26-3'Env18LTR) are able to escape JLR. Accordingly, chimeras expressing the 5' end of enJSRV-18 Env (enJS26-Env18, enJS26-5'Env18, enJS18-3'Env26, enJS18-3'LTR26 and enJS18-3'Env26LTR) are restricted by enJS56A1 (Fig. 23).

Interestingly, although some chimeras share the same portions of Env, they are inhibited by enJS56A1 at different levels. For instance, even if both enJS26-5'Env18 and enJS26-Env18 contain the 5' end of enJSRV-18 Env, the first chimera is more restricted than the second by the transdominant enJS56A1 (Fig. 23, panels A and B). Similarly, in presence of enJS56A1, enJS18-Env26 is able to release greater levels of viral particles than enJS18-5'Env26, despite both chimeras contain the determinant to escape JLR (i.e., the A6D substitution in Env) (Fig. 23, panels A and B). Finally, enJS18-3'LTR26 is restricted by enJS56A1 at greater levels than the correspondent wild-type enJSRV-18 virus (cfr. Fig. 19 and Fig. 23D) while, in presence of enJS56A1, enJS26-3'LTR18 releases higher amount of viral particle compared to the wild-type enJSRV-26 (cfr. Fig. 19 and Fig. 23D). These differences may be due to modifications in the genomic sequence of the chimeras (i.e., introduction or removal of portions of *env* or LTR) that may alter the RNA stability, thereby leading to decreased viral protein expression and particle release. It is noteworthy that enJSRV-26 LTR contains a deletion in the U3 sequence, which may render its RNA less stable.

In presence of the transdominant enJS56A1, the co-expression of enJSRV-26ΔEnv and enJSRV-18ΔEnv with their respective Env proteins (enJSEnv26 and

enJSEnv18) does not recapitulate the phenotypes of the corresponding wild-type viruses. Indeed, the levels of viral particles released by both enJSRV-26 and enJSRV-18 are considerably lower than those of the mutants in presence of the corresponding Env (cfr. Fig. 19 and Fig. 24). These differences may be explained assuming that enJSEnv26 and enJSEnv18 expression plasmids produce higher levels of Env than those encoding the wild-type viruses. Indeed, in a previous study, enJSRVs Env glycoprotein was found to enhance Gag expression and viral particle production (Caporale et al., 2009). Thus, in our system, one could speculate that higher levels of Env glycoprotein may promote Gag synthesis and viral particle release.

It was recently demonstrated that that the signal peptide of JSRV and enJSRV Env possesses a predicted nuclear localization signal (NLS) and an arginine-rich RNA-binding motif (ARM) that determine its intracellular localization as well as function. Indeed, a signal peptide deleted of NLS or ARM sequence displays altered intracellular localization and/or inability to enhance Gag expression and viral particle release (Caporale et al., 2009). Interestingly, the A6D mutation lies within the NLS and ARM motifs of SP26, and could therefore affect the intracellular localization and/or the biological function of the latter. In the next chapter these hypothesis will be investigated.

Chapter IV

How enJSRV-26 escapes enJS56A1-induced restriction

4.1 Introduction

Recent studies have shown that the signal peptide of JSRV Env glycoprotein (JSE-SP) functions as a post-transcriptional regulator of viral gene expression, as it is able to enhance nuclear export of unspliced viral RNAs and increase viral particle release (Caporale et al., 2009; Hofacre et al., 2009; Nitta et al., 2009). Similarly to HIV Rev (Stauber et al., 1995), JSE-SP contains a nuclear localization signal (NLS), an arginine-rich RNA-binding motif (ARM) and a nuclear export sequence (NES) (Caporale et al., 2009; Hofacre et al., 2009). These motifs map to the N-terminal end of the protein and determine its intracellular localization (Caporale et al., 2009; Hofacre et al., 2009). In particular, the NLS domain targets JSE-SP to the nucleoli of transfected cells, while the NES motif is involved in the nucleocytoplasmic export of the aforementioned protein (Caporale et al., 2009; Hofacre et al., 2009). NLS, ARM and NES motifs have also been shown to be important for the functional activity of JSE-SP (Caporale et al., 2009; Hofacre et al., 2009). Indeed, mutations in these domains alter JSE-SP function, resulting in a decrease of Gag synthesis and viral production (Caporale et al., 2009).

Upon targeting Env to the ER, JSE-SP is cleaved-off from the protein and enters the nucleus. Once there, JSE-SP is targeted by NLS to the nucleolus, where it forms ring-like structures (Caporale et al., 2009; Hofacre et al., 2009). The nucleolus is a dynamic subnuclear compartment involved in various processes, including ribosome biogenesis and regulation of cell cycle (Emmott and Hiscox, 2009). These functions are frequently mediated by the sequestration or release of nucleolar proteins. Viruses usually interact with the nucleolus to take over host cell functions and recruit nucleolar proteins to help with virus replication (Wang et al., 2010). It is possible to hypothesize that JSE-SP localizes to the nucleolus to hijack the host cellular machinery, thereby facilitating JSRV replication. There, most likely, JSE-SP binds to its signal peptide-responsive element (SPRE), an RNA secondary structure located at the 3' end of the viral RNA (spanning the last fifty nucleotides of *env* and one hundred fourteen nucleotides of U3) (Caporale et al., 2009). A recent study conducted by Marco Caporale and colleagues in our laboratory demonstrated that the function of JSE-SP depends on the presence of an intact SPRE. Indeed, 293T cells co-transfected with expression plasmids for JSE-SP and SPRE mutants were barely

able to release viral particles (Caporale et al., 2009). The binding between JSE-SP and SPRE seems to be mediated by ARM (Nitta et al., 2009), and results in the nucleocytoplasmic export of full-length viral transcripts *via* the cellular Crm1 protein (Caporale et al., 2009; Nitta et al., 2009).

The results presented in the previous chapter demonstrate that the main determinant of JLR escape is the residue D6 of SP26. Interestingly, this amino acid residue maps to predicted NLS and ARM motifs of the protein. Thus, it could be speculated that the A6D mutation might affect the intracellular localization and/or the activity of SP26. In this chapter are illustrated all the experiments performed to unveil the mechanism adopted by enJSRV-26 in order to escape JLR.

4.2 Results

4.2.1 SP26 does not localize to the nucleolus

The first aim of this study was to assess whether the A6D mutation could affect the intracellular localization of SP26. To this end, the hemagglutinin (HA) epitope was fused to the C-terminus of SP26 and SP18, in order to generate pSP26-HA and pSP18-HA expression plasmids, respectively. Confocal microscopy of cells transfected with pSP26-HA or pSP18-HA showed that both SP18 and SP26 localized in the cytoplasm and in the nucleus of transfected cells. However, while SP18 co-localized with nucleolar markers, such as fibrillarin, SP26 displayed a diffuse nuclear staining pattern with no accumulation in the nucleoli (Fig. 28). The relative number of cells expressing SP18 with nucleolar localization was about 80-fold higher than those expressing SP26 (Fig. 28C). Collectively, these data suggest that the A6D mutation most likely affects the intracellular localization of SP26.

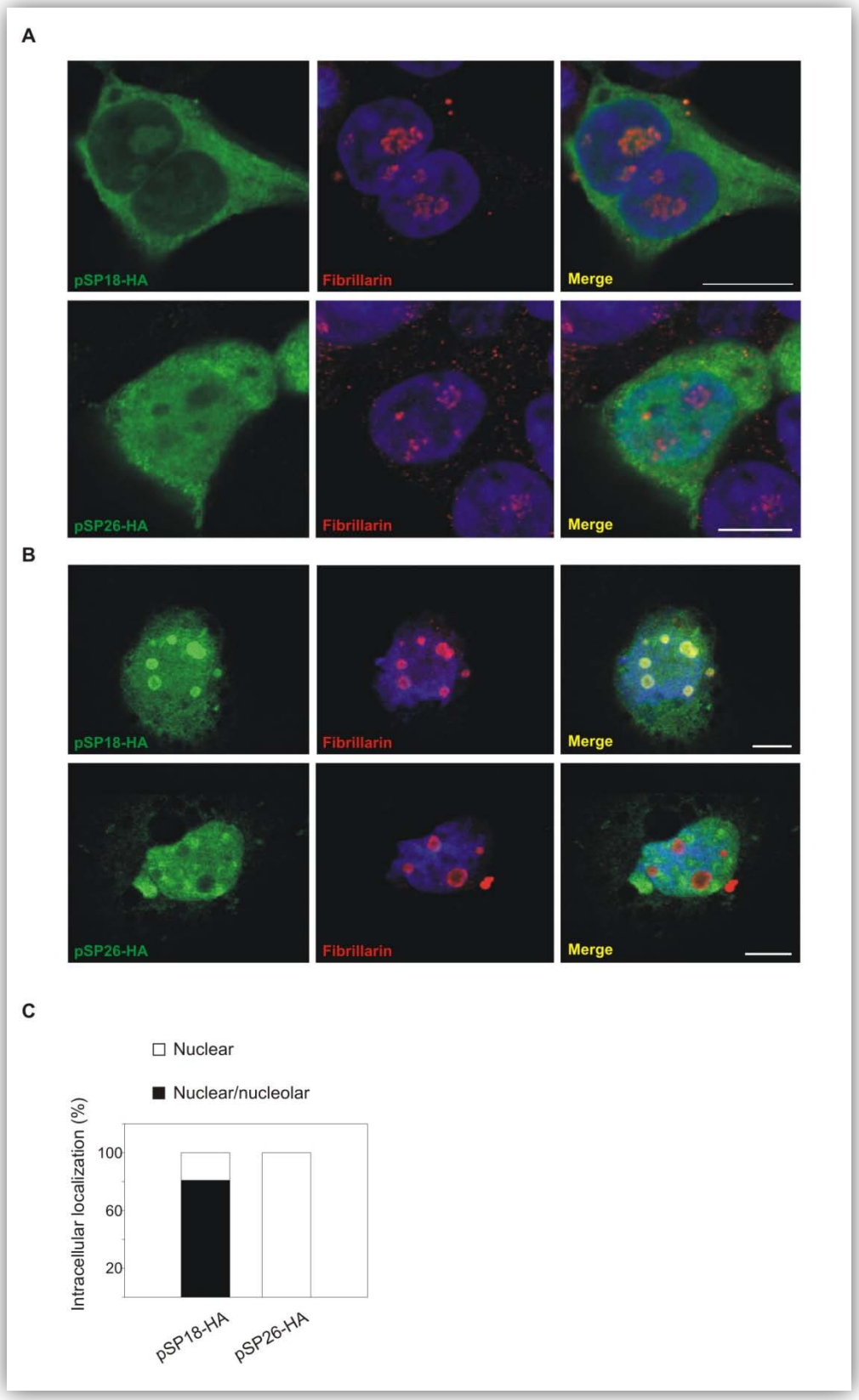


Figure 28| SP26 does not localize to the nucleolus. 293T (A) and COS (B) cells were transfected with pSP26-HA or pSP18-HA expression plasmids, fixed at 24 hours post-transfection, and incubated with anti-HA (green) and fibrillarin (red) antibodies. Nuclei are shown in blue. Both SPs display nuclear localization, but only SP18 shows a strong co-localization with the nucleolar marker fibrillarin. Bars correspond to 10 μm. (C) Percentage of 293T cells expressing pSP26-HA or pSP18-HA and displaying nuclear or nuclear/nucleolar localization. Approximately fifty cells were counted for each experiment.

4.2.2 The presence of SPRE does not alter the intracellular localization of SP26

Both JSRV and enJSRVs contain a signal peptide responsive element (SPRE) located at the 3' end of the viral genome, spanning the *env* and U3 regions. This sequence was found to affect the intracellular localization of JSE-SP *in cis* (Caporale et al., 2009; Nitta et al., 2009). Indeed, in presence of SPRE, JSE-SP relocalized from the nucleus to the cytoplasm of transfected cells, facilitating the export of unspliced viral RNAs (Caporale et al., 2009; Nitta et al., 2009). Hence, the next step of this study was to assess whether the SPREs of enJSRV-26 and enJSRV-18 were also able to affect the intracellular localization of their corresponding signal peptides.

To this end, confocal microscopy experiments were performed on 293T cells co-transfected with pSP26-HA or pSP18-HA expression plasmids and their corresponding Env mutants (enJSRV-26ΔEnv or enJSRV-18ΔEnv respectively), containing intact SPREs. In presence of enJSRV-26ΔEnv, SP26 was found to display a diffuse nuclear staining in the cytoplasm and nucleus of transfected cells with no accumulation in the nucleoli (Fig. 29A), thereby maintaining the same localization pattern as when expressed in absence of SPRE (cfr. Fig. 29, panels A and C and Fig. 28). Conversely, in presence of enJSRV-18ΔEnv, SP18 accumulated preferentially in the nucleoli (Fig. 29B), but the relative number of cells displaying nucleolar localization was lower compared to those expressing SP18 by itself (Fig. 29C *versus* 28C). These results suggest that only the SPRE of enJSRV-18 is able to affect the intracellular localization of its corresponding signal peptide.

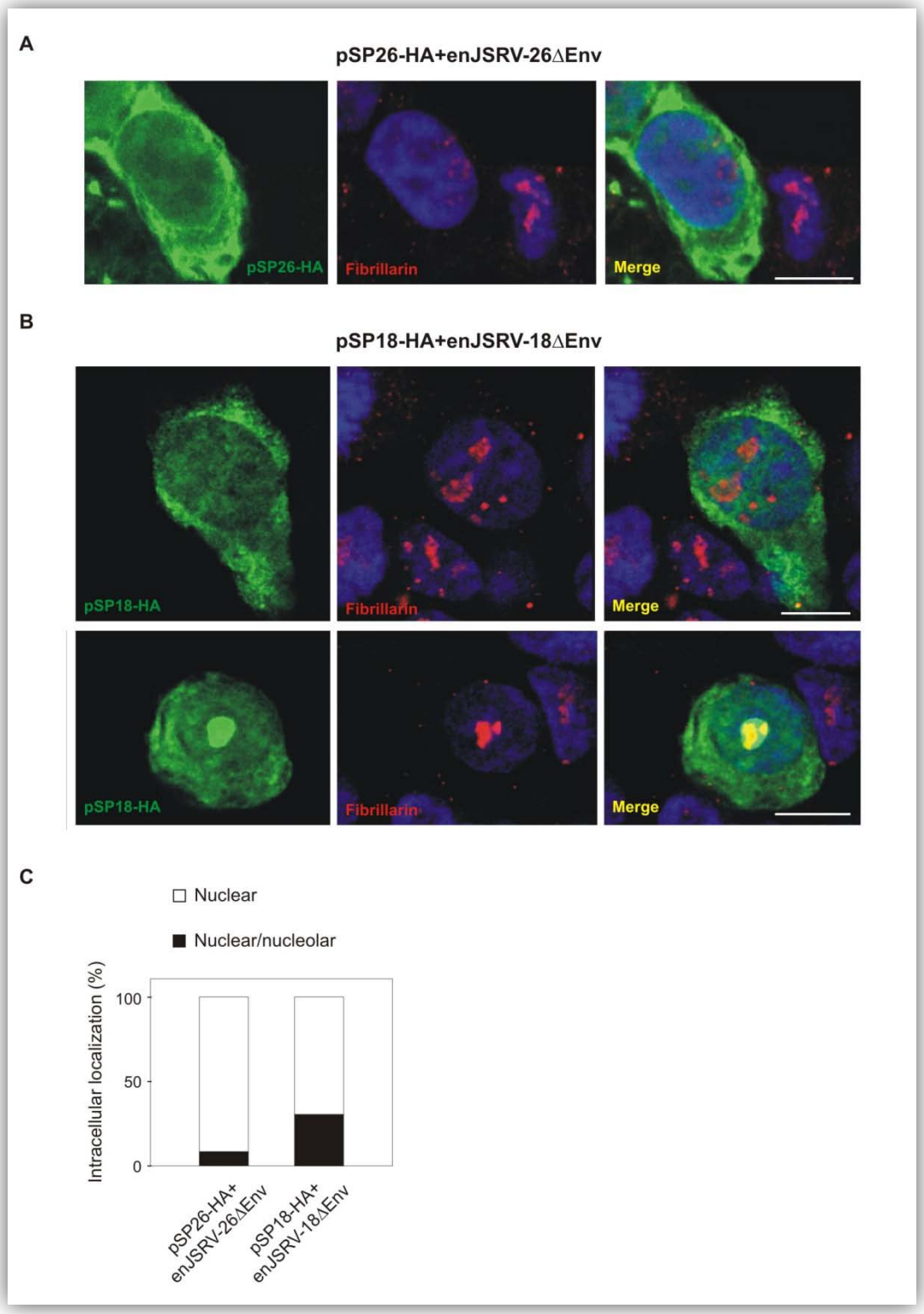


Figure 29 | The presence of the enJSRV-26 signal peptide responsive element (SPRE) does not relocalize SP26. (A and B) 293T cells transfected with the indicated plasmids were fixed at 24 hours post-transfection and stained with anti-HA (green) and fibrillarin (red) antibodies. For confocal experiments of 293T cells expressing pSP18-HA and enJSRV-18ΔEnv expression plasmids, both nuclear (top) and nuclear/nucleolar (bottom) staining are shown. Scale bars correspond to 10 μm. (C) Quantification of nuclear or nuclear/nucleolar staining pattern of SP26 and SP18 expressed in presence of the respective Env mutants in 293T cells. Approximately fifty cells were counted for each experiment.

4.2.3 The signal peptide of enJS56A1 Env does not relocalize SP26 to the nucleolus

Previous studies showed that a HIV-1 Rev mutant with an altered RNA-binding site (RevM5) localized predominantly in the cytoplasm of mammalian cells and was unable to bind the RRE (Malim et al., 1989). However, in presence of HIV-1 Rev, RevM5 accumulated in the nucleoli, indicating that the wild-type protein was able to relocalize this mutant (Daelemans et al., 2004). Therefore, in our experimental system, it was possible that, when co-expressed in the same cell type, the functional SP56 could rescue the defect of SP26 by relocalizing it to the nucleolus.

To test this hypothesis, confocal microscopy experiments were performed on 293T cells transiently transfected with expression plasmids encoding SP26 or SP18 (pSP26-HA or pSP18-HA), their correspondent Env mutants (enJSRV-26ΔEnv or enJSRV-18ΔEnv respectively), and the expression plasmid for the full-length enJS56A1. In presence of enJS56A1, SP26 displayed its typical diffuse nucleocytoplasmic pattern with no accumulation in the nucleoli, indicating that it was not relocalized by the functional SP56 (cfr. Fig. 30, panels A and C and Fig. 29). Similarly, the presence of SP56 did not affect the nucleolar localization of SP18 (cfr. Fig. 30, panels B and C and Fig. 29).

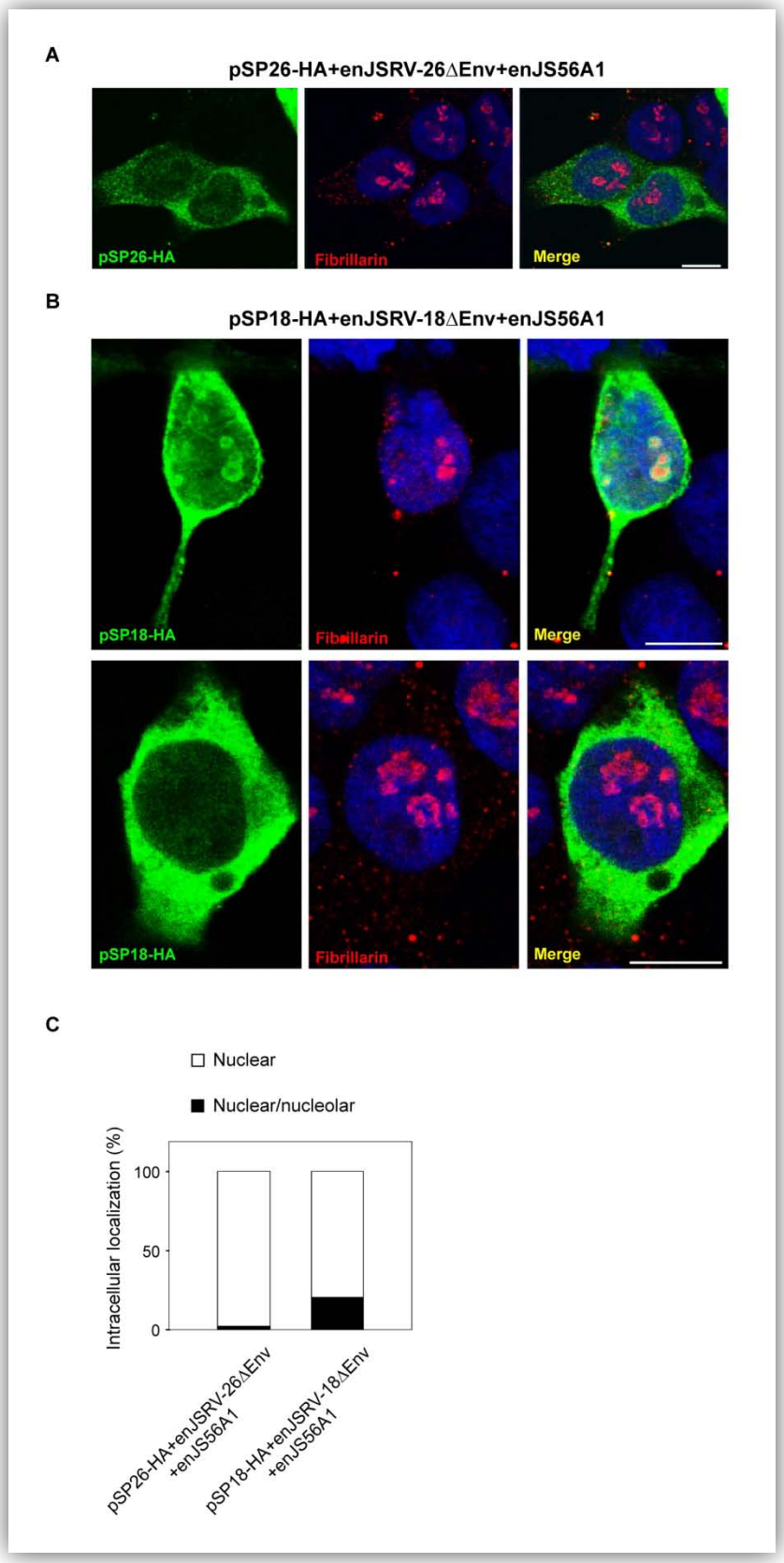


Figure 30| **The functional SP56 does not relocalize SP26.** (A and B) 293T cells were transfected with the indicated plasmids and analysed by confocal microscopy employing antibodies towards the HA epitope (green) and fibrillarin (red). Confocal experiments of cells transfected with pSP18-HA, enJSRV-18ΔEnv and enJS56A1 expression plasmids, displaying both nuclear and nuclear/nucleolar staining are showed. Scale bars, 10 μm. (C) Percentage of 293T cells expressing SP26 or SP18, in presence of enJS56A1 and the correspondent Env mutants, displaying nuclear or nuclear/nucleolar localization. Approximately fifty cells were counted for each experiment.

4.2.4 SP26 does not show dominant negative effects over enJS56A1 Gag expression

The data presented above indicate that the A6D mutation affects the intracellular localization of SP26 and it is not rescued by SP56. However, it was possible that the A6D mutation could influence SP56 intracellular localization, thereby favouring JLR escape. As already mentioned, JSE-SP was found to enhance Gag protein synthesis and viral particle release (Caporale et al., 2009; Hofacre et al., 2009; Nitta et al., 2009). Thus, in our system, it was feasible to speculate that the A6D mutation in SP26 might exert dominant negative effects on the expression of enJS56A1 Gag protein, reducing the levels of defective Gag and promoting enJSRV-26 JLR escape. If this was the case, increasing amounts of SP26 would decrease the amount of enJS56A1 Gag. To test this hypothesis, the levels of enJS56A1 Gag were compared by western blotting in 293T cells co-transfected with enJS56A1 and increasing amounts of expression plasmids for SP26 or SP18 (Fig. 31).

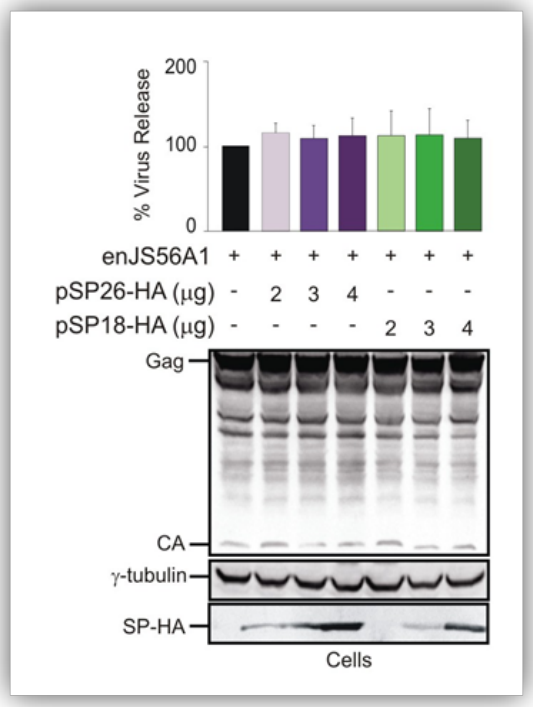


Figure 31| SP26 does not affect enJS56A1 Gag expression. Western blot of cell extracts (cells) of 293T cells transfected with a fixed amount (2 μg) of enJS56A1 expression plasmid and increasing amounts of enJSRV-18 and enJSRV-26 SPs expression plasmids, as indicated in the panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA), the HA epitope (to detect SPs), or γ-tubulin as a loading control. Levels of Gag protein extracted from the lysate of transfected cells were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by enJS56A1 (black bars) were arbitrarily set as 100%. Note that these values are not significantly different, suggesting that neither pSP-26HA nor pSP-18HA affect expression of transdominant Gag proteins. Error bars indicate standard errors.

The presence of SP18 did not alter the level of enJS56A1 Gag proteins (Fig. 31, cfr. green bars and black bar). Interestingly, essentially the same results were obtained in presence of SP26 (Fig. 31, cfr. purple bars and black bar,), ruling out the hypothesis that the latter would act as a dominant negative on the expression of enJS56A1 Gag proteins.

4.2.5 The A6D mutation in enJSRV-26 Env impairs at least some of the functions of its SP

The results collected thus far demonstrate that the A6D mutation in SP26 does not impair the synthesis of transdominant Gag. However, it was possible that the altered localization of SP26 could affect protein expression by enJSRV-26, thereby promoting JLR escape. To test this hypothesis, interference assays were carried out in 293T cells by co-expressing enJSRV-26ΔEnv or enJSRV-18ΔEnv mutants and SP26 or SP18 (pSP26-HA and pSP18-HA, respectively), in presence or absence of enJS56A1 (Fig. 32).

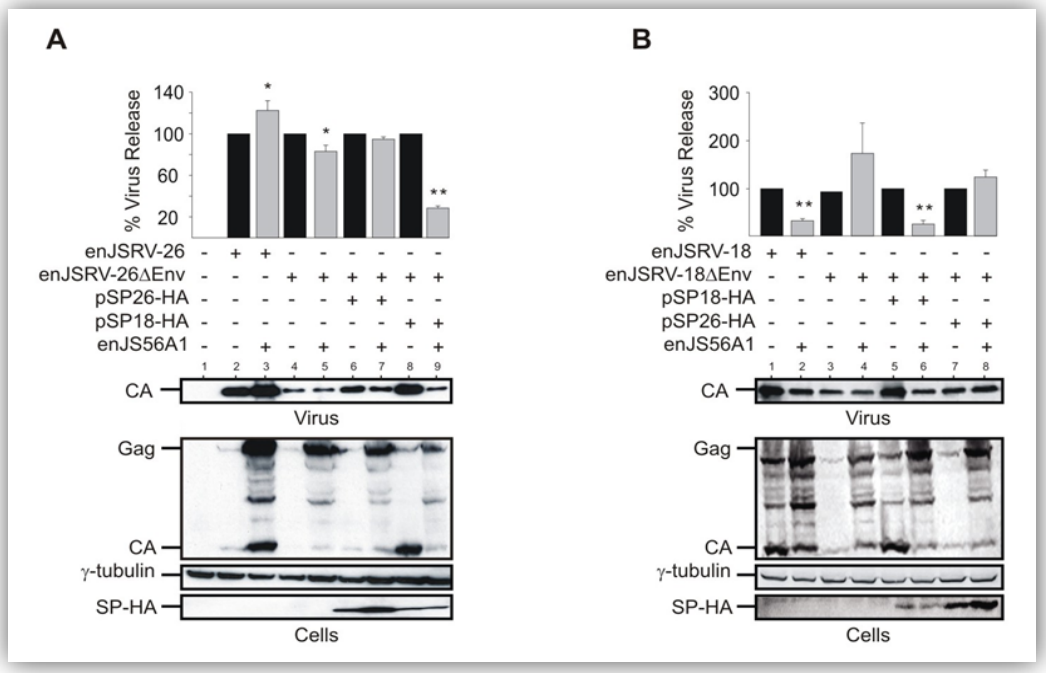


Figure 32| The A6D mutation in enJSRV-26 Env affects the ability of its SP to enhance viral particle release. (A and B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA), the HA epitope (to detect the SPs), or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by enJSRV-26, enJSRV-18 or related mutants in absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P<0.05$) or two ($P<0.01$) asterisks.

enJSRV-26ΔEnv was able to escape JLR if SP26 was provided *in trans* (Fig. 32A, lanes 6 and 7), while it was impaired in presence of SP18 (Fig. 32A, lanes 8 and 9). Similar results were obtained with enJSRV-18ΔEnv that, in presence of SP26, was able to escape the restriction induced by enJS56A1 (Fig. 32B, lanes 7 and 8). Interestingly, viral particle release from enJSRV26ΔEnv was strongly enhanced by SP18 compared to SP26 (Fig. 32A, lanes 4, 6 and 8). Moreover, both enJSRV26ΔEnv and enJSRV18ΔEnv mutants were able to escape JLR when expressed by themselves (Fig. 32A, lanes 4 and 5; Fig. 32B, lanes 3 and 4), suggesting that JLR escape might be due to a relative loss of SP26 function.

As already mentioned in paragraph 1.6.6, betaretroviruses assemble in the pericentriolar area and their Env proteins are involved in intracellular Gag trafficking and particle release (Arnaud et al., 2007b; Sfakianos and Hunter, 2003). The interference assays presented above were carried out in absence of functional Env protein and could therefore lead to misinterpretations of results. To this end, the same experiments as described above were performed in 293T cells by co-expressing enJSRV-26ΔEnv or enJSRV-18ΔEnv mutants in presence of enJSRV-26, enJSRV-18 or enJS56A1 Env glycoproteins (enJSEnv26, enJSEnv18 and enJSEnv56, respectively) (Fig. 33).

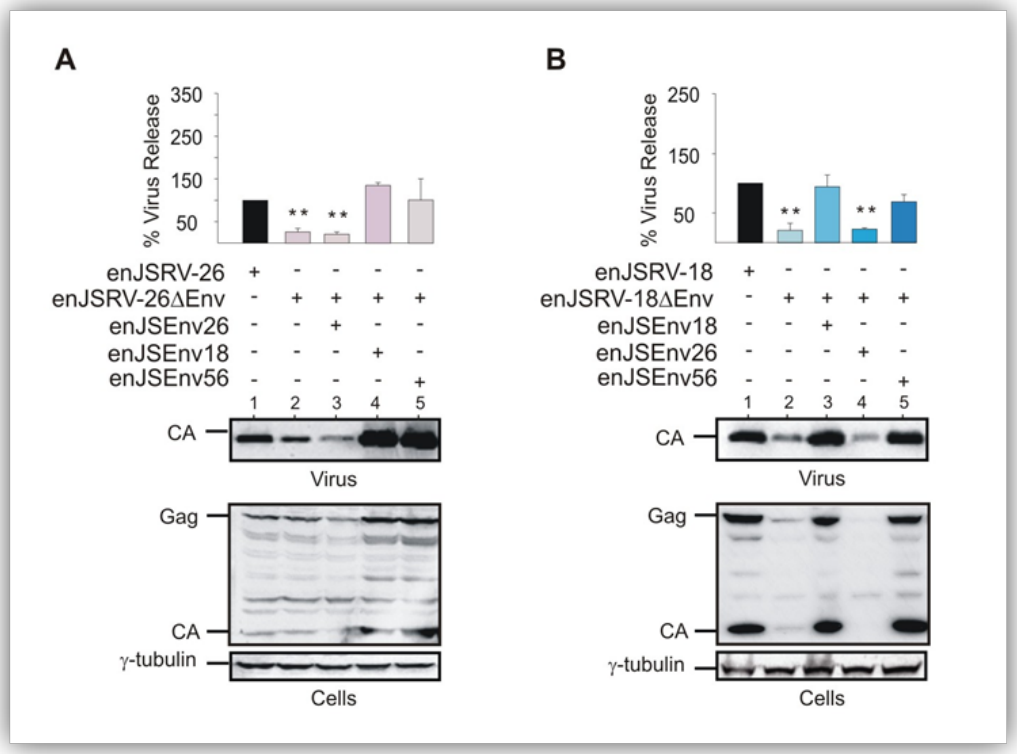


Figure 33| enJS56A1 Env enhances enJSRV-26 expression. Legend on next page.

(From previous page) (A and B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA) and γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by the wild-type enJSRV-26 or enJSRV-18 were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with two ($P<0.01$) asterisks.

As expected, both enJSEnv18 and enJSEnv56 were able to enhance Gag protein synthesis and viral production from enJSRV-26ΔEnv mutant (Fig. 33A, lanes 2, 4 and 5). Conversely, similar levels of viral particle and Gag protein were observed from enJSRV26ΔEnv mutant in presence or absence of enJSEnv26 (Fig. 33A, lanes 2 and 3). Moreover, the levels of Gag protein and viral particle released from enJSRV18ΔEnv mutant were increased in presence of enJSEnv18 and enJSEnv56 compared to when expressed with enJSEnv26 (Fig. 33B, lanes 3, 4 and 5). These results further confirm the hypothesis that the A6D mutation may impair some of the functions of SP26, thereby favouring JLR escape.

4.2.6 The ability of enJSRV-26 to escape JLR can be attributed to a relative lack of function of its SP

The data presented above demonstrate that JLR escape occurs even in absence of signal peptide (Fig. 32A, lanes 4 and 5; Fig. 32B, lanes 3 and 4), suggesting that the A6D mutation may confer loss of SP26 function. If this was the case, JLR escape would take place in presence of a heterologous, and thus not related, signal peptide. To this end, an expression plasmid for enJSRV-18 Env was generated by replacing its signal peptide with the one of the human preprotrypsin protein (enJS18Env-PPT). Western blot analyses were then carried out on supernatant and cell lysates of 293T cells co-transfected with enJSRV-26ΔEnv or enJSRV-18ΔEnv mutant in presence or absence of enJS18Env-PPT (Fig. 34).

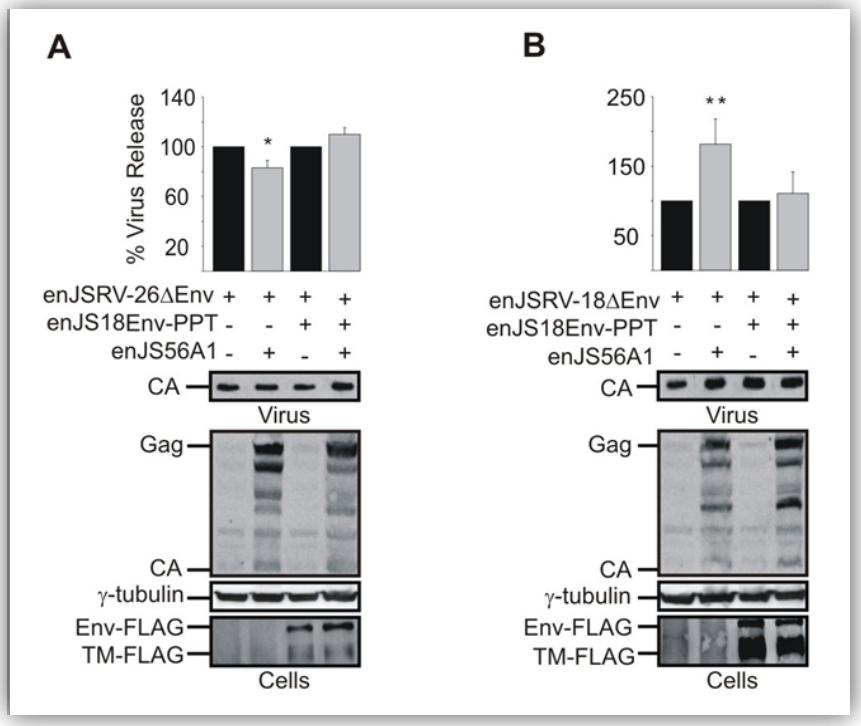


Figure 34| The A6D mutation in enJSRV-26 Env determines a relative lack of function of its SP. (A and B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA), the FLAG epitope (to detect the Env), or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by enJSRV-26ΔEnv or enJSRV-18ΔEnv in absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P<0.05$) or two ($P<0.01$) asterisks.

As shown in figure 34, both enJSRV-26ΔEnv and enJSRV-18ΔEnv were found to escape JLR when enJS18Env-PPT was provided *in trans*. In addition, JLR escape occurred in presence of enJSRV-26ΔEnv and enJSRV-18ΔEnv mutants expressed by themselves (i.e., in absence of SP26 or SP18). These data conclusively demonstrate that JLR escape is due to a loss of SP26 function.

4.2.7 The presence of SP56 is necessary for enJSRV-26 to elude JLR

Recent studies revealed that the JSE-SP activity is mediated by its SPRE, resulting in cytoplasmic accumulation of unspliced viral RNAs and increase of viral particle release (Caporale et al., 2009; Hofacre et al., 2009; Nitta et al., 2009). The results presented thus far show that SP26 is not able to enhance Gag synthesis nor viral exit from enJSRV-26ΔEnv, unlike SP56 or SP18. However, in the presence of enJS56A1, enJSRV-26 always exhibits increased levels of Gag and

viral particles. Thus, it was possible that the functional SP56 might be involved in enhancing enJSRV-26 expression.

In order to test this hypothesis, a heterologous system, such as the Rev-RRE-dependent HIV-1 Gag-Pol expression vector (pNLgagSty330; termed pHIV1-RRE in this study) (Felber et al., 1989; Sfakianos and Hunter, 2003), was employed and constructs were derived by replacing the HIV-1 RRE with the SPREs of enJS56A1 (pHIV-SPRE56), enJSRV-18 (pHIV-SPRE18), and enJSRV-26 (pHIV-SPRE26). These constructs were co-transfected with either pSP18-HA, pSP26-HA, or pSP56-HA, and the release of HIV particles in the supernatants was measured by ELISA (Fig. 35). The levels of Gag released in the supernatants of cells transfected with pHIV-SPRE56, pHIV-SPRE18 or pHIV-SPRE26 did not change in presence or absence of SP26 (pSP26-HA), while they increased considerably (~7 fold) when co-expressed with SP18 or SP56 (Fig. 35). In addition, both SP18 and SP56 were able to increase the levels of Gag expressed by pHIV-SPRE26 (Fig. 35). These results demonstrate that the A6D mutation affects the ability of SP26 to bind SPRE-26 and ultimately increase Gag protein synthesis.

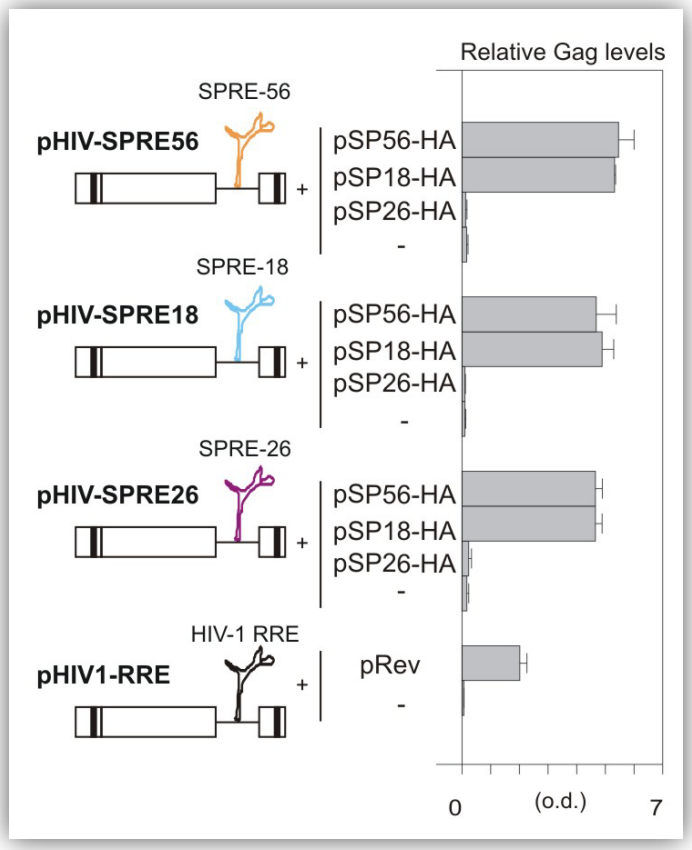


Figure 35 | SP18 and SP56, but not SP26, act as post-transcriptional regulators of viral gene expression. Legend on next page.

(From previous page) HIV Gag ELISAs were performed on supernatants of 293T cells transfected with HIV-1 Gag-Pol expression derived plasmids (pHIV-SPRE26, pHIV-SPRE18, and pHIV-SPRE56) in presence or absence of expression plasmids for enJSRV-26, enJSRV-18, and enJS56A1 SPs (pSP26-HA, pSP18-HA, and pSP56-HA). Controls included supernatants of cells transfected with HIV-1 Gag-Pol expression plasmid containing HIV-1 RRE (pHIV1-RRE) in presence or absence of HIV-1 Rev (pRev). SPRE, signal peptide-responsive element; RRE, Rev-responsive element.

In addition, these data indicate that SPRE26 interacts with both SP56 and SP18, reinforcing the notion that the inability of enJSRV-26 to increase Gag protein level is due to the defect in its SP. Indeed, bioinformatic analysis showed that enJSRV-18 and enJSRV-26 SPREs are predicted to possess an identical RNA secondary structure (Fig. 36).

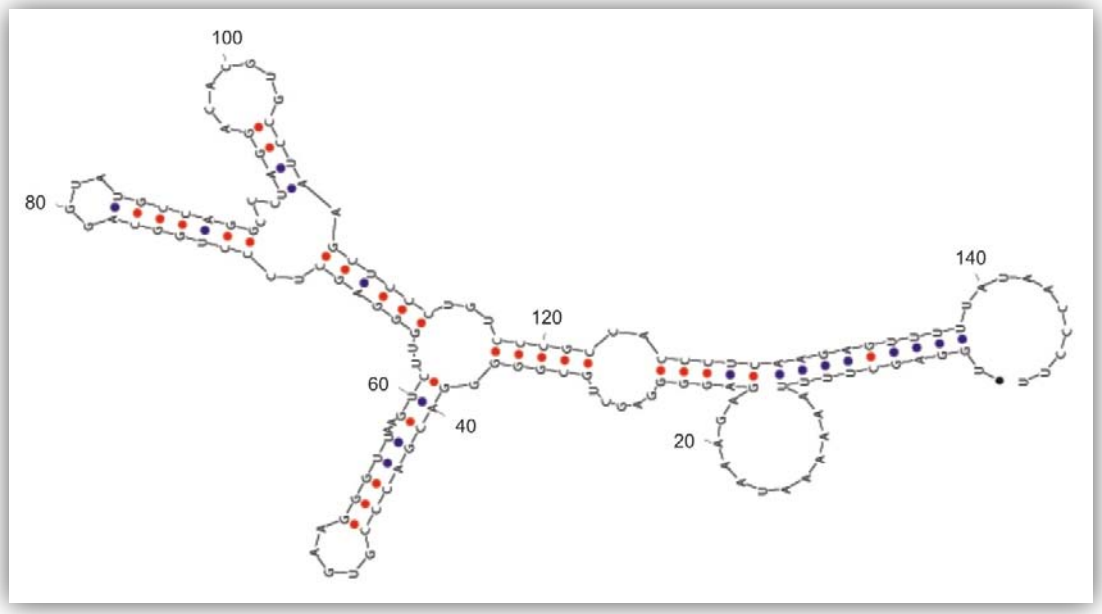


Figure 36 | Predicted RNA secondary structure of enJSRV-26 and enJSRV-18 SPREs. Mfold program (version 3.3) predicted the RNA secondary structure of enJSRV-18 and enJSRV-26 SPREs (ΔG=-54). SPREs include the last fifty nucleotides of *env* and the proximal one hundred and nineteen nucleotides of U3 (note that there is an overlapping of twenty-one nucleotides between the 3' *env* and U3). Numbers refer to the nucleotide residues of enJSRV-26 and enJSRV-18 *env* deposited in GenBank, under accession numbers EF680297 and EF680301, respectively.

Overall, the results illustrated above demonstrate that SP56 is involved in increasing enJSRV-26 expression. Thus, it was possible that this enhancement by SP56 could favour enJSRV-26 JLR escape. If this was the case, in absence of SP56, enJSRV-26 would not be able to elude the restriction mechanism induced by enJS56A1. In order to experimentally address this hypothesis, an enJS56A1 mutant (enJS56A1-4CTE) was generated by replacing the enJS56A1 Env glycoprotein (including its SP) with four M-PMV CTE repeats (Bray et al., 1994). Note that, in enJS56A1-4CTE, the expression of the transdominant Gag is controlled by the four CTEs and is therefore SP-independent. Interference assays

were then carried out with enJSRV-26 and the wild-type enJS56A1 or the enJS56A1-4CTE mutant (Fig. 37).

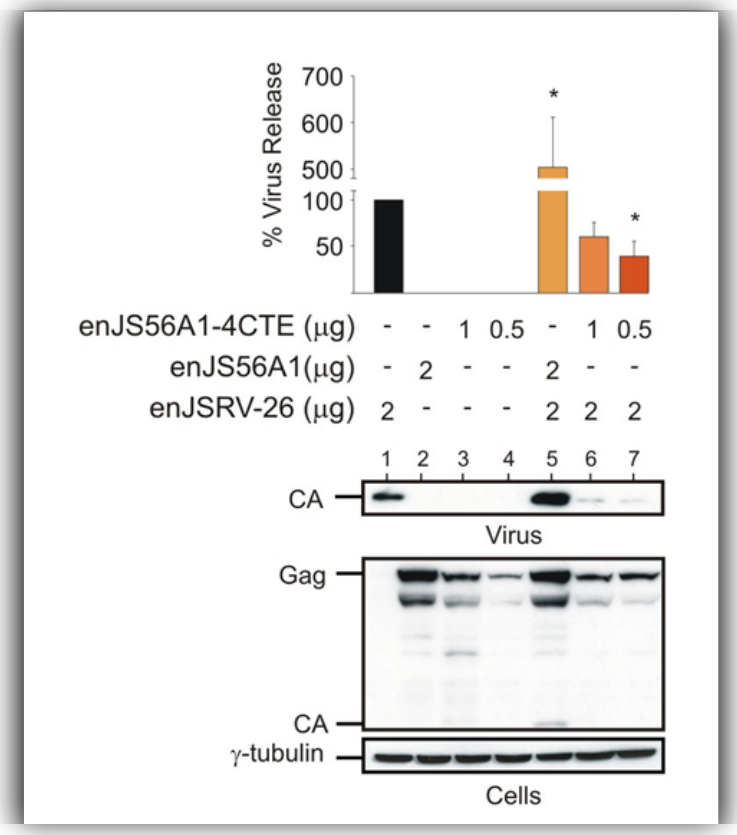


Figure 37| enJS56A1 Env plays a key role in enJSRV-26 JLR escape. Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated. Membranes were incubated with antibodies against the JSRV major capsid protein (CA) and γ-tubulin as a loading control. Note that enJS56A1-4CTE lacks the viral Env and is able to block enJSRV-26. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by enJSRV-26 were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P<0.05$) asterisk.

In presence of enJS56A1-4CTE, viral particle release by enJSRV-26 was impaired considerably (Fig. 37, lanes 1, 6 and 7), indicating that the expression of enJS56A1 Env, and therefore its functional SP, is necessary to enJSRV-26 to escape JLR.

4.2.8 The ratio between enJS56A1 and enJSRV-26 Gag proteins is critical for JLR escape

As already mentioned before (paragraph 3.1), the restriction mechanism induced by enJS56A1 depends on Gag-Gag interactions. The R21W mutation in enJS56A1 Gag seems to alter the overall surface of the protein, resulting in defective particles that cannot traffic efficiently and are subsequently degraded by the

proteasome. Interestingly, the enJS56A1 defect is transdominant over JSRV and enJSRVs; in other words, when co-expressed in the same cell, enJS56A1 Gag protein forms multimers with JSRV and enJSRVs Gag, which are targeted to the proteasomal machinery and subsequently degraded (Arnaud et al., 2007b; Mura et al., 2004; Murcia et al., 2007). Thus, one feasible hypothesis to explain enJSRV-26 JLR escape could be an insufficient amount of enJS56A1 Gag protein available to form multimers and target enJSRV-26 Gag to degradation, thus preventing viral particle release by the latter. If this was the case, increasing amount of enJS56A1 Gag proteins would be able to block enJSRV-26 virus exit.

To test this hypothesis, 293T cells were co-transfected with different ratios of enJSRV-26 and enJS56A1 expression plasmids, and viral particle release was assessed by western blotting (Fig. 38A).

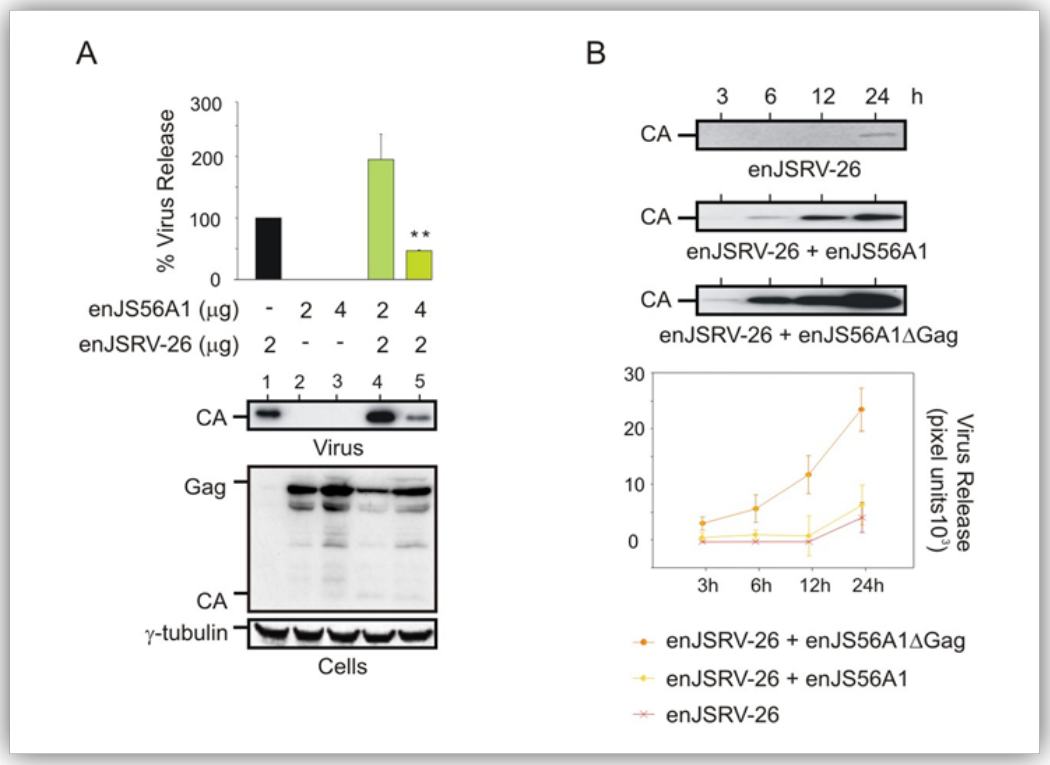


Figure 38| The ratio between enJS56A1 and enJSV-26 Gag is critical for JLR escape. (A) Western blot of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA) and γ-tubulin as a loading control. Note that over-expression of enJS56A1 restricts viral particle release of enJSRV-26. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The value obtained by enJSRV-26 (black bar) was arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with two ($P<0.01$) asterisks. (B) Western blot analysis of concentrated supernatants of 293T cells transfected with the indicated plasmids and analysed at 3, 6, 12, and 24 hours post-transfection. The bottom panel represents the quantification of blots by chemifluorescence as described for panel A. Values are expressed as arbitrary pixel units derived from three independent experiments.

As expected, enJSRV-26 was able to escape JLR when the ratio of the transfected expression plasmids was 1 to 1 (Fig. 38A, lane 4). However, when this ratio was 1 to 2, viral particle release from enJSRV-26 was blocked (Fig. 38A, lane 5), confirming the hypothesis formulated above, and suggesting that JLR is directly related to the relative ratio between transdominant and functional Gag.

These results were further confirmed by interference assays performed with enJSRV-26 in presence of the wild-type enJS56A1 or an enJS56A1 mutant deleted of Gag (enJS56A1ΔGag), and assessing viral particle release at various time points (Fig. 38B). The levels of enJSRV-26 viral particles release were higher in presence of the enJS56A1ΔGag mutant than the wild-type enJS56A1 (Fig. 38B, middle and bottom panels). These results further reinforced the notion that enJSRV-26 JLR escape relies on the ratio between functional and transdominant Gag proteins.

4.3 Discussion

The results presented in the previous chapter demonstrated that the main determinant of JLR escape is the A6D mutation in SP26. Here, it was found that the A6D mutation affects the intracellular localization of SP26. Indeed, the D6 residue lies within predicted NLS and ARM motifs, which have been shown to be important for the intracellular localization and function of JSE-SP (Caporale et al., 2009; Hofacre et al., 2009). Confocal microscopy analysis revealed that SP26 displays a diffuse staining within the nucleus and the cytoplasm of transfected cells, with no accumulation in the nucleolus. The presence of a hydrophilic residue, such as the aspartic acid (D), may alter the overall conformation of SP26, thereby affecting its intracellular localization. In particular, it is possible that only the nucleolar localization is altered, while SP26 may still passively diffuse into the nucleus, given the small size of this protein (approximately 17kDa, on SDS-page gel).

As already mentioned in paragraph 4.1, the nucleolus plays an important role during the course of viral infections. Indeed, several studies have shown that many viruses, including DNA viruses, retroviruses and RNA viruses, localize to the nucleolus to take over host cell functions and favour viral replication (Wang et al., 2010). The HIV-1 Rev protein, for instance, localizes predominantly to the nucleolus (Kubota et al., 1999), where it promotes the nucleocytoplasmic export of viral unspliced RNAs (Daelemans et al., 2004). Deletions of its NLS domain result in the disruption of nucleolar localization and in the inability of Rev to regulate viral gene expression (Cochrane et al., 1990). It is therefore possible to hypothesise that the inability of SP26 to reach the nucleolus may affect its ability to participate in the nucleocytoplasmic trafficking of viral unspliced RNAs. It is noteworthy that the staining pattern observed for SP26 resembled what was previously obtained with a JSE-SP mutant deleted of its NLS (pJSESPΔNLS-HA) (Caporale et al., 2009). Indeed, the JSESPΔNLS-HA mutant localizes both in the nucleus and in the cytoplasm but it is excluded from the nucleolus of transfected cells, and is unable to function as a post-transcriptional regulator of viral gene expression (Caporale et al., 2009). Overall data seem to suggest that, most likely, SP26 does not enhance Gag synthesis and viral particle release due to its inability to bind its SPRE. Indeed, confocal analysis demonstrate that the intracellular localization of SP26 is not affected by the

presence of its SPRE, unlike SP18 (Fig. 29) and JSE-SP (Caporale et al., 2009; Nitta et al., 2009).

JSE-SP possesses HIV Rev-like functions and, similarly to HIV Rev, it favours nucleocytoplasmic export of full-length viral RNAs by binding the SPRE located at the 3' end of the viral genome (Hofacre et al., 2009; Nitta et al., 2009). Unlike JSE-SP, SP26 is not able to enhance Gag protein synthesis and viral particle release. However, this "defect" allows enJSRV-26 to elude enJS56A1 restriction. As already discussed, JLR depends on the interaction between defective and functional Gag, even though the exact stoichiometry between these proteins is currently unknown. A previous study revealed that enJS56A1 is still able to block JSRV even when the latter is over expressed in co-transfection assays (Mura et al., 2004). Conversely, the results presented in this chapter show that the higher the levels of enJSRV-26, the more efficient is the ability of this virus to elude enJS56A1-induced restriction. It was therefore possible that enJSRV-26 might rescue the defective enJS56A1 Gag during JLR. This would explain the increased amount of viral particle released from enJSRV-26 in presence of the transdominant virus. Unfortunately, all the attempts to differentially tag enJSRV-26 and enJS56A1 Gag proteins resulted in altered phenotypes of the wild-type viruses and, therefore, this point could not be experimentally addressed. However, western blot analysis indicates that enJSRV-26 is not totally exempted from JLR, as a certain amount of its Gag protein is restricted by enJS56A1. Thus, it is feasible to assume that enJSRV-26 does not rescue the defect of the transdominant Gag proteins.

As already mentioned (paragraph 1.6.6), betaretroviruses assemble in the pericentriolar area, where their Env glycoproteins facilitate intracellular Gag trafficking and viral particle release (Arnaud et al., 2008; Sfakianos and Hunter, 2003). However, enJSRV-18 and enJSRV-26 Env mutants (enJSRV-26ΔEnv and enJSRV-18ΔEnv) always showed significant amounts of viral particle release. These results could be due to the high transcriptional levels of the expression plasmids for enJSRV-26ΔEnv and enJSRV-18ΔEnv driven by a CMV promoter or, as postulated by others, to the presence of a CTE in JSRV (and enJSRVs) RNA, which can function independently from SPRE (Hofacre et al., 2009; Nitta et al., 2003).

enJSRV-26 JLR escape can be overcome by increasing the levels of transdominant Gag. These results suggest that JLR escape depends on the stoichiometry between enJS56A1 and enJSRV-26 Gag which, in turn, is regulated by the SPs of these viruses. Data obtained by HIV ELISA seem to suggest that SP26 does not facilitate RNA export, even if more experiments are necessary to formally prove this point. Moreover, it appears that the SPREs of enJS56A1 and enJSRV-26 respond equally well to SP56. Overall, the data presented in this chapter support a model where, when co-expressed in the same cell, the SPREs of enJS56A1 and enJSRV-26 compete for the only functional SP (i.e., SP56), resulting in an increased synthesis of enJSRV-26 and reduced levels of enJS56A1 Gag proteins (Fig. 39).

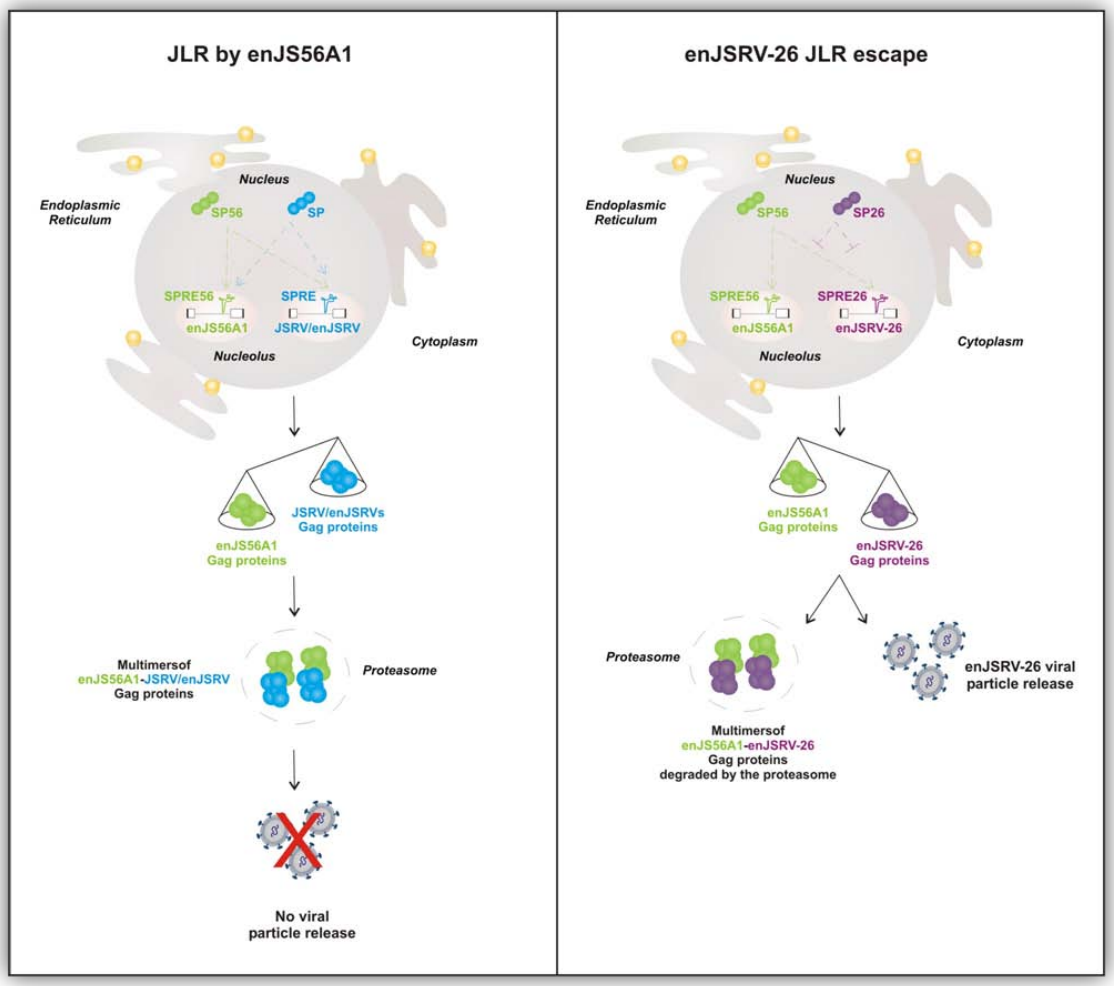


Figure 39 | Model of enJSRV-26 JLR escape. The ability of enJSRV-26 to elude JLR restriction is dependent on the impaired function of its SP. Consequently, SPRE26 and SPRE56 compete for the only functional SP (i.e., SP56), resulting in a reduced expression of transdominant Gag.

Chapter V

The transdominant enJSRV proviruses are amplified in the sheep genome

5.1 Introduction

It is difficult to correlate results obtained in *in vitro* experiments with the evolution of enJSRVs that, *in vivo,* lead to the selection of transdominant proviruses and viruses able to escape JLR. Previous studies suggest that enJS56A1 was positively selected during sheep domestication, most probably as it conferred advantages in protecting the host against infections by related exogenous retroviruses (Arnaud et al., 2007a). Interestingly, another enJSRV locus, enJSRV-20, was also found to possess the same R21W mutation in Gag that confers the defective and transdominant phenotype to enJS56A1 (Arnaud et al., 2007a). enJS56A1 and enJSRV-20 are 99% identical at the nucleotide level and possess intact open reading frames for all of the retroviral genes but *orf-x* (Arnaud et al., 2007a; Palmarini et al., 2000). enJS56A1 contains a two base pairs deletion in *pol* that causes a frameshift and yields a shorter protein than enJSRV-20 and the exogenous $JSRV_{21}$ Pol (Arnaud et al., 2007a; Palmarini et al., 2000). enJS56A1 and enJSRV-20 possess identical 3' genomic flanking regions; however, enJSRV-20 contains a portion of an *env* gene immediately before the 5'LTR (Fig. 11). Overall, these findings lead to hypothesise that enJSRV-20 arose from various processes of recombination between enJS56A1 and other proviruses, rather than independent mutations.

It has been estimated that enJS56A1 and enJSRV-20 entered the sheep genome within the last 3 MYA, during speciation of *Ovis* (Fig. 12). The exogenous enJSRV-like virus from which these transdominant proviruses derived possessed the "wild-type" R residue at position 21 in Gag when it first entered host genome, in order to replicate and infect host germ line. Only subsequently, around the time of sheep domestication, the "transdominant" enJS56A1 genotype harbouring W21 appeared in the host genome and became fixed in the host population (Arnaud et al., 2007a).

During evolution, transdominant proviruses might have been positively selected for the ability to interfere with related exogenous pathogenic retroviruses and enJSRVs already colonizing sheep genome (Arnaud et al., 2007a). The results presented thus far indicate that the relative ratio between defective transdominant and functional Gag could determine the efficiency of JLR. Sheep genome contains an overwhelming majority of enJSRV loci with functional Gag

(i.e., harbouring an R at position 21) (Fig. 11). Interestingly, the BAC clones containing enJS56A1 were found to be overrepresented in the sheep genomic BAC library used to clone the known enJSRV loci. Indeed, 22% of BAC positive for enJSRV sequences contained the enJS56A1 provirus, whereas the expected frequency was 3.7% (Arnaud et al., 2007a).

The overrepresentation of particular clones could be due to artefacts related to the construction and the screening of the library or to the amplification of the genomic region containing enJS56A1. In order to address this point, the copy number variation of the transdominant enJS56A1 was assessed in the genomic DNA of wild and domestic sheep by quantitative polymerase chain reaction (qPCR) assays.

5.2 Results

5.2.1 The chromosomal region containing the enJS56A1 provirus is amplified within the genome of domestic sheep

The relative gene dosage of the transdominant enJS56A1 was determined by employing locus-specific primers annealing in the host genomic DNA and the provirus (Fig. 40).

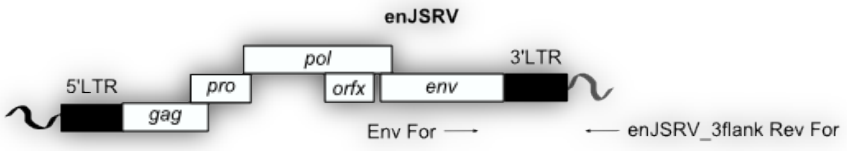


Figure 40| **Schematic representation of an enJSRV provirus.** Primers used for qPCR assays are indicated by arrows.

Samples analysed included wild sheep (*O. dalli*, n=1; *O. canadensis*, n=2; *O. ammon*, n=4; *O. vignei*, n=1), the Mediterranean Mouflon (*Ovis orientalis musimon*, n=4), and different breeds of domestic sheep (*O. aries*), including Dorset (n=1), Suffolk (n=1), Texel (n=10), Jacob (n=2), Red Masai (n=2), Merino (n=3), Xalda (n=2), Rambouillet (n=1), Soay (n=3), Norway (n=2) and Finsheep (n=3). DNA from animals within the genera *Budorcas* (*B. taxicolor*, n=2) and *Pseudois* (*P. nayaur*, n=1) was used as a negative control (Fig. 41).

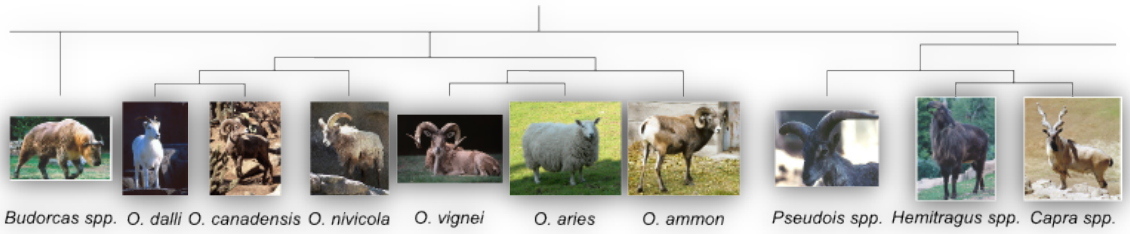


Figure 41| **Representative species within the Caprinae subfamily.** Simplified phylogenetic tree (branch length are not shown to scale) of representative species belonging to the Caprinae subfamily used in this study. Images of the various animal species were kindly provided by Brent Huffman (http://www.ultimateungulate.com) and Wolfgang Dreier. The tree was derived from (Hernandez-Fernandez and Vrba, 2005).

Moreover, relative gene dosages of enJSRV-6 (a provirus fixed throughout the *Ovis* genus) and the insertionally polymorphic enJSRV-18 (present in most but not all domestic sheep) (Arnaud et al., 2007a) were used as additional controls. The data were expressed as the ratio between the estimated copy number of target genes and reference genes in each genomic DNA sample. In order to

increase the accuracy of the results, two host genes, *β-actin* and *SOX9*, were used as internal controls to standardize the qPCR assays.

As expected, when *β-actin* was used as the reference gene, the dosage of *SOX9* did not change significantly across domestic and wild sheep. The modern enJSRV-18 provirus showed low variations among domestic sheep, whereas enJSRV-6 displayed modest divergences between the samples tested. In contrast, the transdominant enJS56A1 revealed major differences between the samples analyzed, with clear indications of genomic amplification in some domestic sheep breeds. It is noteworthy that, within the *Ovis* genus, wild sheep phylogenetically closer to *O. aries*, such as *O. ammon* and *O. vignei*, displayed a higher variation in enJS56A1 copy number compared to those phylogenetically more distant from domestic sheep, such as *O. dalli* and *O. canadensis* (Fig. 42).

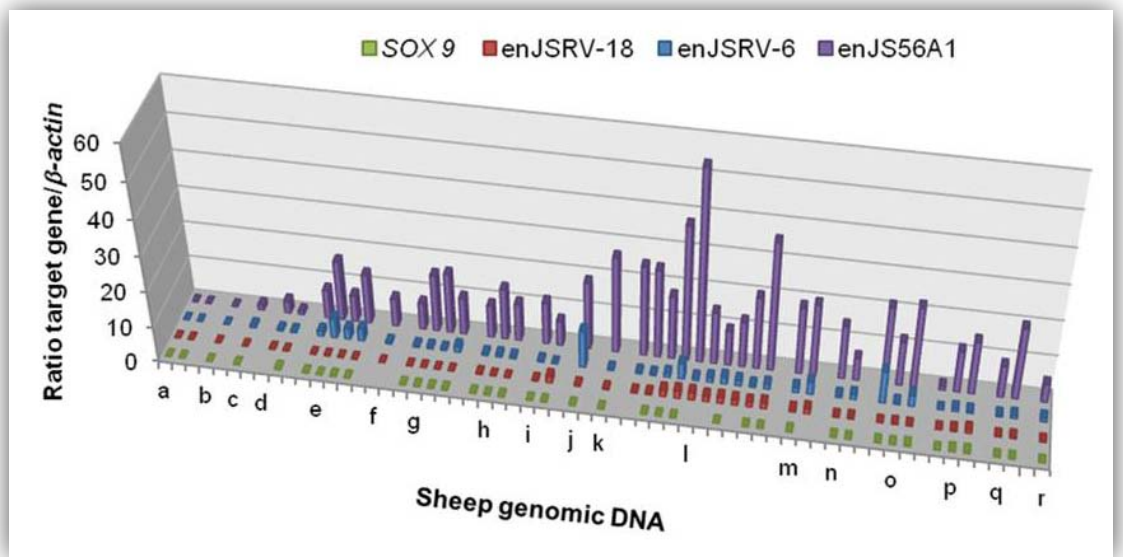


Figure 42 | Copy number variation of the enJS56A1 provirus relative to the *β-actin* gene in wild and domestic sheep. Graph representing the copy number variation of the transdominant enJS56A1 (purple), enJSRV-6 (blue), enJSRV-18 (red) and *SOX9* (green) loci relative to the *β-actin* gene. Each bar represents a single animal and each letter represents a different species/breed. The absence of a bar indicates that the assay could not be performed due to limited amount of the DNA sample available. Samples tested included genomic DNA collected from *B. taxicolor* (a), *P. nayaur* (b), *O. dalli* (c), *O. canadensis* (d), *O. ammon* (e), *O. vignei* (f), *Ovis orientalis musimon* (g), and various breeds of the domestic sheep (*O. aries*), such as Soay (h), Norway (i), Dorset (j), Suffolk (k), Texel (l), Jacob (m), Red Masai (n), Finsheep (o), Merino (p), Xalda (q), and Rambouillet (r). Note that enJS56A1 is amplified within the genome of domestic sheep.

A similar pattern was observed when *SOX9* was employed as reference gene to determine the copy number variation of enJS56A1, enJSRV-6 and enJSRV-18 in the genomic DNA of wild and domestic sheep (Fig. 43).

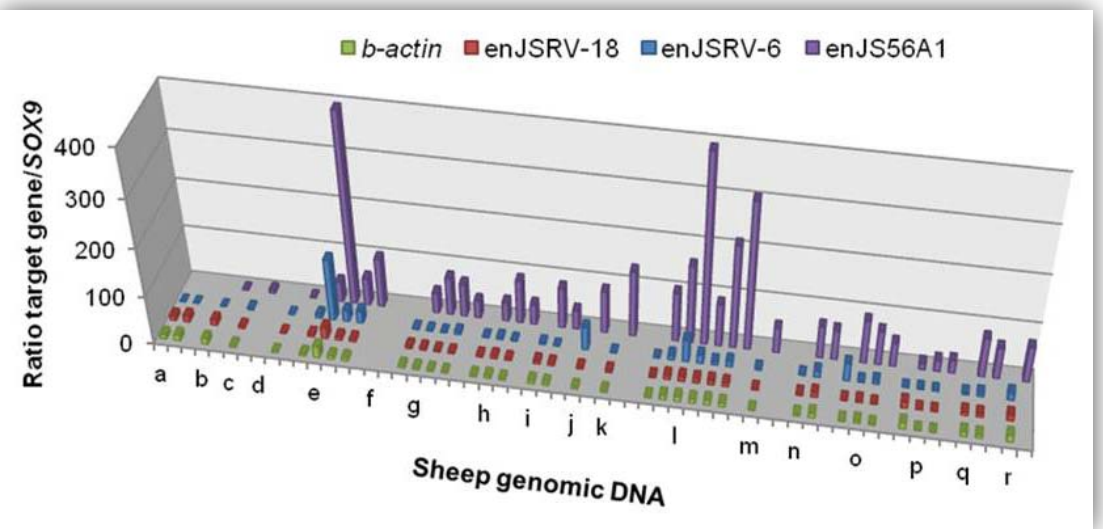


Figure 43| Copy number variation of enJS56A1 relative to the *SOX9* gene in wild and domestic sheep. Graph representing the copy number variation of the transdominant enJS56A1 (purple), enJSRV-6 (blue), enJSRV-18 (red) and *β-actin* (green) loci relative to the *SOX9* gene. Each bar represents a single animal and each letter represents a different species/breed. The absence of a bar indicates that the assay could not be performed due to exhaustion of the DNA sample. Samples tested included genomic DNA collected from *B. taxicolor* (a), *P. nayaur* (b), *O. dalli* (c), *O. canadensis* (d), *O. ammon* (e), *O. vignei* (f), *Ovis orientalis musimon* (g), and various breeds of the domestic sheep (*O. aries*), such as Soay (h), Norway (i), Dorset (j), Suffolk (k), Texel (l), Jacob (m), Red Masai (n), Finsheep (o), Merino (p), Xalda (q), and Rambouillet (r). The transdominant enJS56A1 is amplified within the genome of domestic sheep.

Curiously, all the loci examined displayed a higher gene dosage compared to those obtained with *β-actin* used as the reference gene. Nevertheless, it is noteworthy that the relative gene dosage of enJS56A1 was maintained in each sample, confirming the notion that the transdominant provirus is amplified within the genome of domestic sheep, regardless of the control gene used in the analysis.

FISH analysis on metaphase chromosomes derived from domestic sheep (Fig. 44) showed that at least one of the copies of the transdominant enJS56A1 provirus (enJSRV-20) maps exactly to the same chromosomal location as enJS56A1 in chromosome 6 (6q13). Interestingly, as already mentioned, enJS56A1 and enJSRV-20 share identical 3' genomic flanking regions (Arnaud et al., 2007a). Thus, overall data suggest that genome amplification has driven the amplification of enJS56A1-like proviruses with identical genomic flanking regions in any given animal analyzed in this study. FISH data were performed in a collaborative study by Giulia Pia di Meo in Leopoldo Iannuzzi laboratory in Naples (Italy).

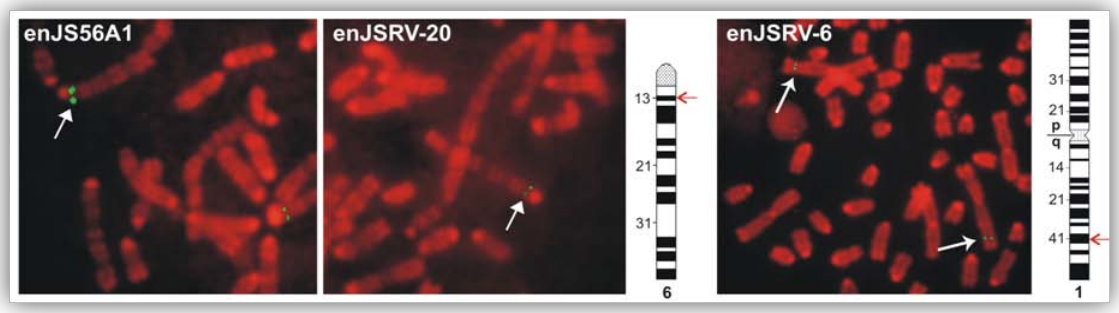


Figure 44| Fluorescent *in situ* hybridization of metaphase R-banded chromosomes derived from a Merino sheep (mixed breed). Fluorescent probes were derived from BAC clones containing enJSRV-20, enJS56A1 or enJSRV-6 as described in chapter 2. The green fluorescent signals (indicated by arrows) are specific for the two transdominant proviruses (both located on chromosome 6, band 6q13) and the enJSRV-6 locus (situated on chromosome 1, band 1q41). Ideograms of *Ovis aries* chromosomes with R-banding patterns are also shown.

As already mentioned (paragraph 5.1), it was previously speculated that enJSRV-20 arose by a process of recombination/gene conversion with enJS56A1. These two proviruses can be distinguished by minor nucleotide sequence differences (twenty-three nucleotides along the entire genome) and the 5' flanking region, but they share identical 3' flanking regions (Arnaud et al., 2007a). The qPCR assays performed in this study uses a forward primer designed in a conserved region in *env* and a reverse primer located in the 3' flanking region of enJS56A1. Consequently, all the enJS56A1-like proviruses sharing identical 3' flanking regions (and not only the enJS56A1 provirus *per se*) might have been amplified from the genomic DNA of all the samples analyzed. In order to address this point, PCR experiments were performed on those BAC clones resulted positive for enJSRV-20 and already characterized in a previous study (Arnaud et al., 2007a), employing the same primers and conditions used to amplify enJS56A1 by qPCR in the present work.

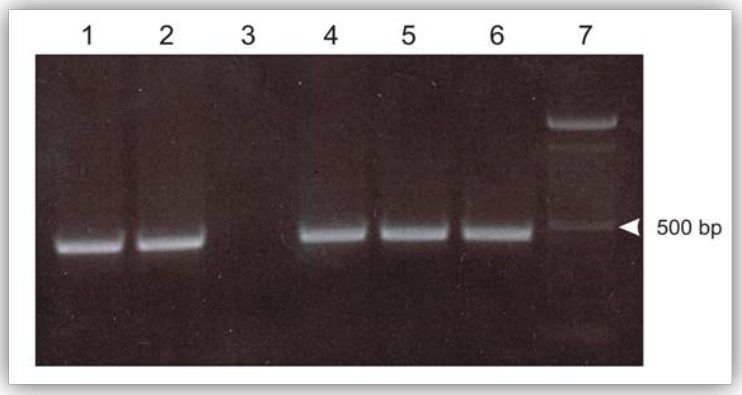


Figure 45| Amplification of the enJSRV-20 locus using enJS56A1-specific PCR primers. PCR fragments amplified from BAC clones previously showed to be positive for enJSRV-20 (81J8 and 63N7; lanes 1 and 2) or enJS56A1 (106U3, 40N10 and 14C5; lanes 4-6). Control primers for PCR (lane 3) and 500 bp ladder (lane 7) are also shown.

As shown in figure 45, enJSRV-20 was amplified from all the BAC clones analyzed confirming that, during qPCR assays, enJS56A1-like proviruses with identical genomic flanking regions were amplified in all the samples tested.

5.2.2 Frequency of the W21 residue in the Gag of enJS56A1-like proviruses in the genome of domestic and wild sheep

Previous results obtained in our laboratory demonstrated that the W21 residue in enJS56A1 and enJSRV-20 Gag became fixed during sheep domestication (Arnaud et al., 2007a). The data presented thus far show the presence of multiple copies of enJS56A1-like proviruses in the genome of domestic sheep. Hence, the next step was to assess the relative frequency of R21 and W21 codons in the Gag of proviruses present in the genome of representative wild and domestic sheep. To this end, the 5' *gag* coding region of enJS56A1-like proviruses (including enJSRV-20) was amplified from the genomic DNA of wild sheep (*O. dalli*, n=1; *O. canadensis*, n=1; *O. ammon*, n=1), the Mediterranean Mouflon (*O. orientalis musimon*, n=1), and two different breeds of domestic sheep (Texel, n=2; Merino, n=1). The PCR products were then cloned and at least forty individual clones for each sample were sequenced (Table 2).

Table 2| Relative frequency of the wild-type arginine (R) or the transdominant tryptophan (W) residue at position 21 in the Gag of enJS56A1-like proviruses[a].

Genus	enJS56A1	enJSRV-20
O. dalli (c)	100% R	-
O. canadensis (d)	100% R	100% R
O. ammon (e)	90% W, 10% R	-
O. orientalis (g)	73% W, 27% R	100% W
O. aries (l)	100% W	100% W
O. aries (l)	100% W	100% W
O. aries (p)	100% W	100% W

[a] The 5' gag coding region of enJS56A1 and enJSRV-20 was amplified by PCR from genomic DNA collected from the species indicated in the table. Note that letters in parentheses (c, d, e, etc.) refer to the code used for figure 42. PCR products were cloned into the pCR4-TOPO vector (Invitrogen), and forty individual clones for each PCR product were sequenced to determine the relative presence of an R or W residue at position 21 in Gag.

The codon corresponding to the Gag R21 residue was detected in 100% of the PCR clones derived from the amplification of DNA collected from *O. dalli* and *O. canadensis*, which are the species phylogenetically more distant from domestic sheep among those analyzed in this study, and for which there was no evidence

of genomic amplification (Arnaud et al., 2007a). Conversely, and as expected, 100% of the PCR clones derived from the amplification of DNA collected from *O. aries* was found to possess the W21 codon, indicating that all the copies of enJS56A1-like proviruses are transdominant within the genome of domestic sheep. Finally, in *O. ammon* and *O. orientalis*, which are the species phylogenetically closer to *O. aries* among those analyzed in this study, the great majority of enJS56A1-like proviruses was found to possess the transdominant phenotype (Table 2).

5.2.3 Analysis of copy number variation of transdominant enJS56A1-like proviruses in OPA affected sheep

Previous studies conducted in our laboratory suggested that the W21 residue in enJS56A1 Gag might have been positively selected, during evolution, for the ability to interfere with related exogenous and pathogenic retroviruses (Arnaud et al., 2007a). Interestingly, in the field, most JSRV infected sheep do not develop OPA during their commercial lifespan (Caporale et al., 2005). Thus, it could be speculated that animals infected with JSRV and that develop lung tumours possess a lower copy number variation of transdominant enJS56A1-like proviruses compared to those that do not develop OPA.

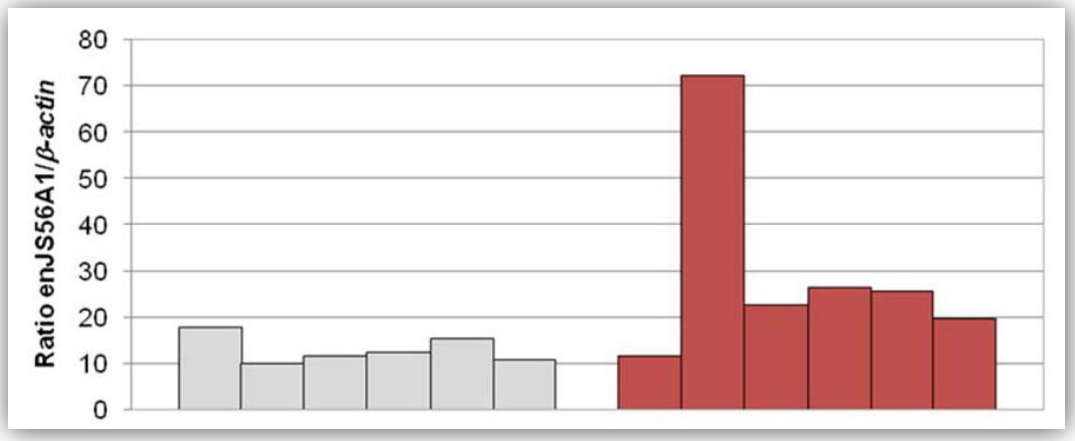


Figure 46| Copy number variation of the transdominant enJS56A1 in the genomic DNA of normal and JSRV infected sheep. Graph representing the copy number variation of the transdominant enJS56A1 relative to the *β-actin* gene. Each bar represents a different animal. Samples tested included genomic DNA collected from seven healthy (in grey) and seven OPA affected (in red) Blackface sheep (*O. aries*).

In order to address this point, a pilot experiment was performed to evaluate the gene dosage of transdominant proviruses in healthy (n=7) and OPA affected (n=7) Blackface sheep by qPCR, as already described. OPA affected sheep did not

display a lower copy number of enJS56A1-like proviruses compared to healthy animals (Fig. 46), ruling out the hypothesis that the genomic amplification of transdominant proviruses *per se* may determine OPA outcome.

5.3 Discussion

The results presented in this chapter reinforce the notion that transdominant proviruses have been positively selected by the host during sheep domestication. In particular, these data demonstrate that the chromosomal region containing enJS56A1-like proviruses is amplified in the sheep genome, resulting in the generation of multiple copies of transdominant proviruses. Previous studies conducted in our laboratory suggested that these enJS56A1-like proviruses might have been co-opted because they provided adaptive advantages to the host (Arnaud et al., 2007a; Arnaud et al., 2007b; Mura et al., 2004; Murcia et al., 2007). In particular, in accordance with a "protective" role in sheep, it was speculated that a second transdominant provirus, enJSRV-20, most likely arose by processes of recombination and/or gene conversion with enJS56A1 (Arnaud et al., 2007a). The results presented in this chapter suggest that the chromosomal location containing enJS56A1 has been amplified several times, especially in some breeds of domestic sheep, thereby further supporting the idea that sheep domestication has contributed to the selection and amplification of transdominant proviruses.

Previous work conducted by Frédérick Arnaud and colleagues demonstrate that enJS56A1 possessed the "wild-type" R residue at position 21 in Gag, when it first entered sheep genome. Only subsequently, the "transdominant" enJS56A1 genotype harbouring W21 appeared in the host and became fixed around the time of sheep domestication (Arnaud et al., 2007a). In line with these findings, by determining the relative frequencies of W21 residues in the Gag of enJS56A1-like proviruses in both wild and domestic sheep, we confirmed that transdominant proviruses became fixed in the host genome around sheep domestication, further corroborating previous results (Arnaud et al., 2007a).

The challenge of this study was to identify meaningful differences in the copy number variation of transdominant enJSRV proviruses between wild and domestic sheep by qPCR assays. The accuracy of this technique greatly relies on the use of multiple valid control genes for normalization, as well as on the size of the fragments amplified from the genomic DNA. Indeed, it has been shown that amplicon size is a critical factor for qPCR assays, as smaller fragments increase the levels of specificity. Optimal PCR products should be less than one

hundred fifty base pairs in length (Varga and James, 2006). In our assays, however, PCR products were clearly too long (between four hundred and thirty-eight and five hundred and seventy-seven base pairs) than usually recommended, due to the limited possibilities of designing specific primers that would not include the LTR regions[25]. Thus, amplicons length might have reduced primers binding specificity, raising the possibility that the copy number variations of all the loci investigated in this study might have been underestimated. Moreover, reverse primers were designed assuming that all the copies of enJS56A1-like proviruses that integrated into the sheep genome shared the same 3' flanking region, which might not have been necessarily the case, and might have, therefore, further contributed to underestimate the gene dosage of transdominant proviruses.

Furthermore, it is important to bear in mind that one of the limits of our qPCR method is the reliability of the reference genes employed in the assay. Indeed, the absolute copy number estimated for all the loci examined in this study greatly differed depending on the reference gene used as internal control. However, reassuringly, the relative gene dosage of the various enJSRV proviruses in wild and domestic sheep was generally maintained. A possible explanation could be that the sheep genome possesses different copies of *β-actin* and *SOX9* genes due to the presence, for example, of pseudogenes that could affect the gene dosage estimation of the loci examined.

Another critical point to consider is that the qPCR assays presented in this study assessed the gene dosage, and not the expression, of the transdominant proviruses. Therefore, the twenty or more copies of enJS56A1-like proviruses detected in some animals may not be necessarily (all) expressed. Indeed, although it is possible that the genomic amplification of transdominant proviruses in the sheep genome may be directly proportional to the level of expression of these loci, one cannot rule out the hypothesis that some of these copies may be (genetically and/or epigenetically) silenced.

Finally, the pilot experiment on the evaluation of gene copy number variation of transdominant enJSRVs in healthy and OPA affected sheep did not confirm our hypothesis on a protective role played by these proviruses against the

[25] LTRs are similar at each end in every provirus as well as among different enJSRV loci.

development of lung tumours. The analysis of a much larger number of samples will be necessary to firmly address this point. In addition, it would be interesting to include a group of JSRV infected sheep that do not develop OPA, in order to assess whether there is any direct correlation between JSRV infection and OPA resistance.

Chapter VI

Conclusions

The evolutionary interplay between endogenous and exogenous retroviruses has always been a difficult subject to study in outbreed animal species, because of the lack of reliable *in vivo* models. For example, the hypothesis that ERVs may protect the host against infections by related exogenous retroviruses came almost exclusively from experiments performed in laboratory mice (Kozak et al., 1984). Sheep, on the other hand, represent an extremely fascinating model to study these evolutionary mechanisms in natural conditions, because of the co-existence, in the genome of these animal species, of at least twenty-seven enJSRV loci highly related to the exogenous and pathogenic JSRV. Moreover, most of the enJSRV proviruses are biologically active and are still invading the sheep genome (Arnaud et al., 2007a).

Over time, sheep and enJSRVs have engaged in mutualistic relationships, in which both of them have gained mutual benefits. Sheep have contributed to the maintenance of enJSRVs by transmission to subsequent generations; enJSRVs, on the other hand, have played critical roles in host survival. Experiments conducted by Tom Spencer and our group demonstrated indeed that, in absence of functional enJSRV Env proteins, ewes abort at the very early stages of pregnancy (Dunlap et al., 2006b).

Several studies suggest that enJSRVs might have also provided protection against infections by related exogenous retroviruses. Similarly to murine Fv4, expression of enJSRVs Env glycoproteins can block, *in vitro*, cell entry of related exogenous retroviruses by receptor interference (Spencer et al., 2003). Moreover, the transdominant enJS56A1 can interfere, *in vitro*, with the late replication steps of JSRV, by a unique mechanism known as JLR. The main determinant of JLR was mapped to the W21 residue in enJS56A1 Gag, which substitutes a R21 well conserved in betaretroviruses (Mura et al., 2004). The R21W mutation confers to enJS56A1 Gag a defective phenotype that is transdominant over JSRV as well as other enJSRVs (Arnaud et al., 2007a; Mura et al., 2004). enJS56A1 and JSRV Gag molecules seemingly form chimeric multimers that cannot traffic properly and are further degraded by the proteasome (Arnaud et al., 2007b; Murcia et al., 2007).

The results presented in this thesis cover relevant biological aspects of enJSRVs biology that were previously unknown, particularly how the simultaneous

expression of transdominant and functional Gag proteins can lead to JLR escape. In order to better characterize the molecular bases of enJSRV-26 JLR escape, we performed a variety of *in vitro* experiments, in which we compared the phenotype of the escape mutant to that of enJSRV-18. Indeed, at the nucleotide level, enJSRV-18 is 98% identical to enJSRV-26, but it is blocked by the transdominant enJS56A1. Because JLR impairs viral exit and due to the lack of a robust *in vitro* replication system for JSRV and enJSRVs in ovine cells, the experiments performed in this thesis were mainly based on transfections of 293T cells with virus-expressing plasmids, followed by western blot analysis of cell lysates and viral pellets obtained from supernatants. Within this experimental approach, we performed interference assays with chimeric mutants to map the molecular determinant of JLR escape. We found that a single point mutation (A6D) in enJRV/JSRV Env can alter the intracellular localization and function of its SP. We also revealed that the A6D mutation in Env confer to enJSRV/JSRV the ability to escape JLR. Furthermore, we proposed a model for JLR escape that strictly depends on the intracellular kinetics of functional and defective SPs. The experimental data generated in this thesis suggest that the ARM and NLS domains of enJSRV/JSRV SP are critical in determining the balance between transdominant and wild-type Gag and, therefore, the outcome of JLR. Finally, we unveiled that the chromosomal region containing enJS56A1 has been amplified several times during domestication, particularly in some breeds of domestic sheep.

A key point to address in this chapter is the biological relevance of the work presented in this thesis. The importance of JLR escape can be considered from different perspectives. From a functional point of view, we showed that a single point mutation (A6D) impairs the biological function of SP26. This observation is directly connected to the ability of enJSRV/JSRV SPs to act as post-transcriptional regulators of viral gene expression. Previous work conducted by Marco Caporale and colleagues demonstrated that, similarly to HIV-1 Rev, enJSRVs and JSRV SPs are critical for the nucleocytoplasmic export of viral unspliced RNAs (Caporale et al., 2009). We showed that the A6D substitution in SP26 affects its intracellular localization. Although we do not know whether enJSRV/JSRV SP localizes to the nucleolus to take over host cell functions, we assume that, most likely, this is the site where binding between SP and viral unspliced RNAs occurs. This assumption is supported by studies demonstrating

that the nucleolus is the site where HIV-1 Rev multimerizes and it is likely involved in nucleocytoplasmic transport (Daelemans et al., 2004; Michienzi et al., 2006).

Within the context of the evolutionary interplay between host and pathogen, the emergence of animals harbouring transdominant proviruses must have exerted selective pressure for the appearance of exogenous viruses able to escape JLR. Exogenous viruses could have used different ways to evade the restriction activity exerted by transdominant proviruses. For example, they might have "run away" by acquiring different tissue tropism. Most likely, this represents the strategy adopted by JSRV to avoid JLR: replicating in tissues where interfering enJSRVs are not highly expressed. Indeed, enJSRVs are primarily detected in the genital tract of sheep (Palmarini et al., 1996b; Palmarini et al., 2000; Palmarini et al., 2001a; Spencer et al., 1999), while JSRV is abundantly expressed in proliferating type 2 pneumocytes of sheep lung (Murgia et al., 2011). Several lines of evidences seem to indicate that, over time, endogenous and exogenous sheep betaretroviruses acquired different tissue tropism due to their LTR regions (Palmarini et al., 2000). We speculate that the ancestor of the modern circulating JSRV was initially expressed in the genital tract of sheep. Subsequently, during the process of endogenization and, most likely, in coincidence with the appearance of transdominant proviruses, some exogenous JSRV-like viruses might have diverted their tropism from the genital tract towards the lung, in order to escape the restriction mechanism exerted by transdominant enJSRVs.

In line with this, it is worth noting that enJSRV-26 possesses a thirty-three base pairs deletion in its LTRs, which might have contributed, besides the A6D mutation in SP26, to escape the restriction mechanism exerted by transdominant proviruses.

From an evolutionary perspective, this study has provided important information on the dynamic interplay between host and pathogen. Previous studies estimated that enJS56A1 integrated into the sheep genome between 0.9 to 1.8 MYA (Palmarini et al., 2000). Most likely, the exogenous virus from which enJS56A1 derives possessed the wild-type R residue at position 21 in Gag, in order to replicate and successfully infect host germ line (Arnaud et al., 2007a;

Murcia et al., 2007). Only subsequently, the transdominant enJS56A1 genotype, harbouring W21, appeared in the host genome and became fixed around the time of sheep domestication, approximately 0.9 MYA (Arnaud et al., 2007a). We reason that, with domestication, a relatively large number of animals were suddenly kept in restricted spaces, and this likely facilitated the spread of infectious agents more easily than before. Under these circumstances, sheep with transdominant proviruses might have had a selective advantage over animals not harbouring them. It is possible that the driving force that influenced the fixation of transdominant enJSRVs in the genome of domestic sheep was their role in protecting the host against infections by related exogenous retroviruses. Experiments conducted in our laboratory indicate indeed that transdominant proviruses can interfere *in vitro* with the late steps of JSRV/enJSRV viral replication.

An important question to address is whether JLR influences the outcome of JSRV infections *in vivo*. Our preliminary results seem to indicate that this is not the case, even though more studies are needed to formally prove this point. However, before completely ruling out this hypothesis, certain considerations should be made. First, OPA is a very rare disease: it has been shown that only a minority of JSRV infected sheep develop the tumour in natural conditions (Caporale et al. , 2005). Second, OPA is a very slow disease in naturally infected animals, which takes up to several months/years before developing (Caporale et al., 2005). Third, low levels of virus can be detected in peripheral leucocytes and lymphoid organs of infected animals with or without clinical OPA (Holland et al., 1999). These observations lead to hypothesize that JSRV uses cells of the lymphoreticular system as main reservoir of infection, without the requirement to infect the target cells (proliferating type 2 pneumocytes) for transformation. A recent study conducted by Claudio Murgia and colleagues suggests that OPA is triggered by JSRV infection of proliferating type 2 pneumocytes in active division, which are abundantly present in young lambs during post-natal development, and in adults after lung injury induced, for example, by respiratory pathogens (Murgia et al., 2011). However, the mechanisms involved in viral maintenance for such long periods of incubation in the host remain still to be elucidated, and it is tempting to speculate that JLR can potentially play a role in keeping the spread of the virus “on hold” within infected animals. Indeed, the fact that the chromosomal region containing enJS56A1 has been

amplified several times strongly support the notion that JLR has provided some benefits to the host.

enJSRVs are abundantly expressed in the genital tract of sheep. A recent study conducted by Frédérick Arnaud and colleagues in our laboratory suggests that enJSRVs viral tropism might have been influenced by host restriction factors, such as oBST-2. Support to this hypothesis comes from the observation that enJSRVs are expressed in the luminal and glandular epithelia of ovine uterus, while both isoforms of oBST-2 are mostly found in the stroma (Arnaud et al., 2010). Since the duplication of the oBST-2 gene predated the initial invasion of enJSRVs in the sheep genome, it has been proposed that oBST-2 might have been one of the selective forces that confined enJSRVs within specific areas of the reproductive tract, where these cellular restriction factors were not expressed at all, or at very low levels (Arnaud et al., 2010).

In response to IFNτ, intact enJSRV loci are able to release viral particles into the uterine lumen of pregnant ewes, and potentially infect ovine conceptus (Black et al., 2010b). The presence of transdominant proviruses in the genital tract of sheep may therefore prevent unrestrained integrations, potentially deleterious for the developing conceptus. The bovine genome does not contain enJSRVs (Hecht et al., 1996), thus bovine blastocysts were utilized in the aforementioned study to unequivocally assess the potential infectivity of enJSRVs in recipient ovine uteri. Phylogenetic analyses revealed that transdominant proviruses, such as enJ56A1 and enJSRV-20, were rarely recovered from ovine endometria, while enJSRV-26 escape mutants-like (i.e., harbouring the A6D mutation in the SP) were found in two of the four bovine conceptuses analyzed. In addition, the great majority of enJSRV sequences, amplified from the endometria and uterine flushes of recipient ewes and transferred bovine embryos, clustered mainly with the youngest enJSRV loci (Black et al., 2010a). Perhaps, the low abundance of enJS56A1-like proviruses is necessary to promote *de novo* integrations of intact enJSRV loci in the sheep genome that may render redundant the function provided by older proviruses (e.g., role of enJSRV Env in placental morphogenesis). However, it is important to bear in mind that any retroviral integration is potentially mutagenic and, if uncontrolled, may jeopardize host survival.

The presence of JLR escape mutants, together with the genomic amplification of transdominant proviruses in the sheep genome, illustrate really well the dynamic between host and pathogen. Phylogenetic analyses suggest that enJSRV-26 might have been derived from the closely related enJSRV-18 provirus (Arnaud et al., 2007a). Under these circumstances, it is feasible to speculate that both the A6D mutation in SP26 and the LTR deletion might have conferred selective advantages to enJSRV-26 (and enJSRV26-like), including JLR escape. The recent integration of enJSRV-26 in the sheep genome strongly suggests that the interplay between endogenous and exogenous sheep betaretroviruses is still ongoing. Interestingly, enJSRV-26 was found only in one Texel ram and may represent a unique integration event (Arnaud et al., 2007a). Thus, the most feasible scenario could be that an exogenous retrovirus, closely related to enJSRV-26, is still circulating (or it has been circulating) within sheep population.

Host-pathogen interaction is modelled as a typical "arms race", in which each partner gains advantage over the other by maximizing its own fitness at the other expenses. Co-evolutionary processes favour rapid rates of evolution and are driven by recombinations that lead to constant natural selection for adaptation and counter-adaptation. This "back-and-forth" interplay has been highly dynamic and contributed to rapid changes in viral and host strategies, with each "species" rushing to evolve the upper hand in the interaction in a never ending struggle or, in the words of Lewis Carroll's Red Queen, «*it takes all the running you can do, to keep in the same place*»[26]. The studies presented in this thesis have provided further insights on how exogenous and endogenous sheep betaretroviruses have interacted with their hosts during evolution.

[26] The Red Queen hypothesis (Van Valen, 1973) is named after the Red Queen's race in Lewis Carroll's "Through the Looking-Glass, and What Alice Found There" (1871), in which the Red Queen states: «*it takes all the running you can do, to keep in the same place*». In evolutionary biology, this sentence can be translated as "continuous adaptation is needed in order for a pathogen to maintain its relative fitness by escaping host defences, or for the host to counteract pathogen infections".

Appendix

Oligonucleotides employed in this study

Table 3 | List of oligonucleotides employed for cloning and site-directed mutagenesis.

Primer	Sequence (5'→3')	Plasmid[a]
enJS26-3'E18 For	ATAATTTGTTTAGCTCCTTGCCTTATTCGT	enJS26-3'Env18
enJS26-3'E18 Rev	ACGAATAAGGCAAGGAGCTAAACAAATTAT	
enJS18-3'E26 For	ATAATTTGTCTAGCTCCTTGTCTTATTCGT	enJS18-3'Env26
enJS18-3'E26 Rev	ACGAATAAGACAAGGAGCTATACAAATTAT	
enJS26Stop1Env For	GTGTTTTTCCACAGGTAACCGAAGCGCCGCGATG	enJSRV-26ΔEnv
enJS26Stop1Env Rev	CATCGCGGCGCTTCGGTTACCTGTGGAAAAACAC	
enJS26Stop2Env For	CACATCAAATGCAACGCTAAACACTGAGCGAGCCCAC	
enJS26Stop2Env Rev	GTGGGCTCGCTCAGTGTTTAGCGTTGCATTTGATGTG	
enJS18Stop1Env For	GTGTTTTTCCACAGGTAACCGAAGCGCCGCGC	enJSRV-18ΔEnv
enJS18Stop1Env Rev	GCGCGGCGCTTCGGTTACCTGTGGAAAAACAC	
enJS18Stop2Env For	CACATCAAATGCATCGCTAAACACTGAGTGAGCCCAC	
enJS18Stop2Env Rev	GTGGGCTCACTCAGTGTTTAGCGATGCATTTGATGTG	
enJS26-D6A For	ATGCCGAAGCGCCGCGCTGGATTCCGGAAAGG	enJSRV26-EnvD6A
enJS26-D6A Rev	CCTTTCCGGAATCCAGCGCGGCGCTTCGGCAT	
enJS18-A6D For	ATGCCGAAGCGCCGCGATGGATTCCGGAAAGG	enJSRV18EnvA6D/ JSRV-EnvA6D
enJS18-A6D Rev	CCTTTCCGGAATCCATCGCGGCGCTTCGGCAT	
pSP26-HA For	CCCAGACTACGCTTGCGGGGGACGACCCGTGAAGGG TTAAGTCTTGGGAGCT	pSP26-HA
pSP26-HA Rev	ACGTCGTATGGGTAAGCCGCCCCGTTTTGTATCCGCT GTAACAATAATAATAA	
pSP18-HA For	CCCAGACTACGCTTGCGGGGGACGACCCGTGAAGGG TTAAGTCTTGGGAGCT	pSP18-HA
pSP18-HA Rev	ACGTCGTATGGGTAAGCTGCCCCGTTTTGTATCCGCT GTAACAATAATAATAA	
pSP56-HA For	CCCAGACTACGCTTGCGGGGGACGACCCGTGAAGGG TTAAGTCTTGGGAGCT	pSP56-HA
pSP56-HA Rev	ACGTCGTATGGGTAAGCCGCCCCGTTTTGTATCCGCT GTAACAATAATAATAA	
enJS56Stop1Gag For	GTTGAGAGTATAAATTAAGGACAGACGCATAGTC	enJS56A1ΔGag
enJS56Stop1Gag Rev	GACTATGCGTCTGTCCTTAATTTATACTCTCAAC	
enJS56Stop2Gag For	CATATGTTATCTGTATAATTAAAACATTGGGG	
enJS56Stop2Gag Rev	CCCCAATGTTTTAATTATACAGATAACATATG	
pHIV-SPRE26 For	AGAGGAATTCTAACCGAAGCGCCGCGATGGATTC	pHIV-SPRE26
pHIV-SPRE26 Rev	AAAATCTAGAGCTGATACCTTGCTTTATTGTGC	
pHIV-SPRE18 For	AGAGGAATTCTAACCGAAGCGCCGCGCTGGATTC	pHIV-SPRE18
pHIV-SPRE18 Rev	AAAATCTAGAGCTGATACCTTGCTTTATTGTGC	

Table 4 | List of oligonucleotides employed for qPCR assays.

Primer	Sequence (5'→3')	Plasmid[a]
Env For	ATAAAGAGAGGGGAGCTGCG	enJS56A1
enJS56A1_3flank Rev	GGAAGGATCTGAAACGTGGA	
Env For	ATAAAGAGAGGGGAGCTGCG	enJSRV-6
enJSRV6_3flank Rev	CAGGGGAATAACTGGTGCTACCT	
Env For	ATAAAGAGAGGGGAGCTGCG	enJSRV-18
enJSRV18_3flank Rev	CAAGTGCCAGAGCCCAGAGCCA	
β-actin For	ATCATGTTTGAGACCTTCAACACCCC	β-actin
β-actin Rev	CCAGGAAGGAAGGCTGGAAGAGAGC	
SOX9 For	CCTAGCTTTTCTTGCAGCC	SOX9
SOX9 Rev	GCATTCCCCAGACAGATTTC	

[a] Name of the resulting expression plasmids.

Bibliography

Alberti, A., Murgia, C., Liu, S.L., Mura, M., Cousens, C., Sharp, M., Miller, A.D. and Palmarini, M. (2002) Envelope-induced cell transformation by ovine betaretroviruses. *J Virol*, 76, 5387-5394.

Amé, J.C., Spenlehauer, C. and de Murcia, G. (2004) The PARP superfamily. *Bioessays*, 26, 882-893.

Anderson, J.L., Campbell, E.M., Wu, X., Vandegraaff, N., Engelman, A. and Hope, T.J. (2006) Proteasome inhibition reveals that a functional preintegration complex intermediate can be generated during restriction by diverse TRIM5 proteins. *J Virol*, 80, 9754-9760.

Antony, J.M., Ellestad, K.K., Hammond, R., Imaizumi, K., Mallet, F., Warren, K.G. and Power, C. (2007) The human endogenous retrovirus envelope glycoprotein, syncytin-1, regulates neuroinflammation and its receptor expression in multiple sclerosis: a role for endoplasmic reticulum chaperones in astrocytes. *J Immunol*, 179, 1210-1224.

Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prevost, M.C., Allen, T.D. and Charneau, P. (2007) HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *Embo J*, 26, 3025-3037.

Arnaud, F., Black, S.G., Murphy, L., Griffiths, D.J., Neil, S.J, Spencer, T.E. and Palmarini, M. (2010) Interplay between ovine bone marrow stromal cell antigen 2/tetherin and endogenous retroviruses. *J Virol*, 84, 4415-4425.

Arnaud, F., Caporale, M., Varela, M., Biek, R., Chessa, B., Alberti, A., Golder, M., Mura, M., Zhang, Y.P., Yu, L., Pereira, F., Demartini, J.C., Leymaster, K., Spencer, T.E. and Palmarini M. (2007a) A paradigm for virus-host coevolution: sequential counter-adaptations between endogenous and exogenous retroviruses. *PLoS Pathog*, 3, e170.

Arnaud, F., Murcia, P.R. and Palmarini, M. (2007b) Mechanisms of late restriction induced by an endogenous retrovirus. J Virol, 81, 11441-11451.

Arnaud, F., Varela, M., Spencer, T.E. and Palmarini, M. (2008) Coevolution of endogenous betaretroviruses of sheep and their host. *Cell Mol Life Sci*, 65, 3422-3432.

Askjaer, P., Jensen, T.H., Nilsson, J., Englmeier, L. and Kjems, J. (1998) The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J Biol Chem*, 273, 33414-33422.

Bai, J., Zhu, R.Y., Stedman, K., Cousens, C., Carlson, J., Sharp, J.M. and DeMartini, J.C. (1996) Unique long terminal repeat U3 sequences distinguish exogenous Jaagsiekte sheep retroviruses associated with ovine pulmonary carcinoma from endogenous loci in the sheep genome. *J Virol*, 70, 3159-3168.

Baldwin, D.N. and Linial, M.L. (1998) The roles of Pol and Env in the assembly pathway of human foamy virus. *J Virol*, 72, 3658-3665.

Balvay, L., Lopez Lastra, M., Sargueil, B., Darlix, J.L. and Ohlmann, T. (2007) Translational control of retroviruses. *Nat Rev Microbiol*, 5, 128-140.

Baltimore, D. (1970) RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*, 226, 1209-1211.

Bannert, N., Fiebig, U. and Hohn, O. (2010) Retroviral particles, proteins and genomes. *Retroviruses*. R. Kurth and N. Bannert. Nopen reading frameolk, Caister Academic Press: 71-106.

Baribaud, F., Maillard, I., Vacheron, S., Brocker, T., Diggelmann, H. and Acha-Orbea, H. (1999) Role of dendritic cells in the immune response induced by mouse mammary tumor virus superantigen. *J Virol* 73, 8403–8410.

Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). 1983. *Science*, 220, 868-871.

Basu, V.P., Song, M., Gao, L., Rigby, S.T., Hanson, M.N. and Bambara, R.A. (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res*, 134, 19-38.

Basyuk, E., Boulon, S., Skou Pedersen, F., Bertrand, E. and Vestergaard Rasmussen S. (2005) The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. *J Mol Biol,* 354, 330-339.

Bazer, F.W., Burghardt, R.C., Johnson, G.A., Spencer, T.E. and Wu, G. (2008) Interferons and progesterone for establishment and maintenance of pregnancy: interactions among novel cell signaling pathways. *Reprod Biol*, 8, 179-211.

Belshaw, R., Watson, J., Katzourakis, A., Howe, A., Woolven-Allen, J., Burt, A. and Tristem, M. (2007) Rate of recombinational deletion among human endogenous retroviruses. *J Virol,* 81, 9437-9442.

Berkhout, B., Arts, K. and Abbink, T.E. (2011) Ribosomal scanning on the 5'-untranslated region of the human immunodeficiency virus RNA genome. *Nucleic Acids Res*, 39, 5232-5244.

Berlioz, C. and Darlix, J.L. (1995) An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors. *J Virol*, 69, 2214-22122.

Berthoux, L., Sebastian, S., Sokolskaja, E. and Luban, J. (2005) Cyclophilin A is required for TRIM5{alpha}-mediated resistance to HIV-1 in Old World monkey cells. *Proc Natl Acad Sci USA,* 102, 14849-14853.

Bertrand, P., Côté, M., Zheng, Y.M., Albritton, L.M. and Liu, S.L. (2008) Jaagsiekte sheep retrovirus utilizes a pH-dependent endocytosis pathway for entry. *J Virol*, 82, 2555-2559.

Best, S., Le Tissier, P., Towers, G. and Stoye, J.P. (1996) Positional cloning of the mouse retrovirus restriction gene Fv1. *Nature*, 382, 826-829.

Bhadra, S., Lozano, M.M., Payne, S.M. and Dudley, J.P. (2006) Endogenous MMTV proviruses induce susceptibility to both viral and bacterial pathogens. *PLoS Pathog,* 2, e128.

Bi, W., Huang, W., Whitworth, D.J., Deng, J.M., Zhang, Z., Behringer, R.R. and de Crombrugghe, B. (2001) Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. *Proc Natl Acad Sci USA*, 98, 6698-6703.

Bick, M.J., Carroll, J.W., Gao, G., Goff, S.P., Rice, C.M. and MacDonald, M.R. (2003) Expression of the zinc-finger antiviral protein inhibits alphavirus replication. *J Virol*, 77, 11555-11562.

Bieche, I., Olivi, M., Champeme, M.H., Vidaud, D., Lidereau, R. and Vidaud, M. (1998) Novel approach to quantitative polymerase chain reaction using real-time detection: application to the detection of gene amplification in breast cancer. *Int J Cancer*, 78, 661-666.

Bieniasz, P.D., Weiss, R.A. and McClure, M.O. (1995) Cell cycle dependence of foamy retrovirus infection. *J Virol*, 69, 7295-7299.

Bieniasz, P.D. (2004) Intrinsic immunity: a front-line defense against viral attack. *Nat Immunol*, 5, 1109-1115.

Bittner, J.J. (1936) Some possible effects of nursing on the mammary gland tumor incidence in mice. *Science*, 84, 162.

Black, S.G., Arnaud, F., Burghardt, R.C., Satterfield, M.C., Fleming, J.A., Long, C.R., Hanna, C., Murphy, L., Biek, R., Palmarini., M, and Spencer, T.E. (2010a) Viral particles of endogenous betaretroviruses are released in the sheep uterus and infect the conceptus trophectoderm in a transspecies embryo transfer model. *J Virol*, 84, 9078-9085.

Black, S.G., Arnaud, F., Palmarini, M. and Spencer, T.E. (2010b) Endogenous retroviruses in trophoblast differentiation and placental development. *Am J Reprod Immunol,* 64, 255-264.

Blaise, S., de Parseval, N., Benit, L. and Heidmann, T. (2003) Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci USA*, 100, 13013-13018.

Blond, .JL., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F. and Cosset, F.L. (2000) An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. *J Virol*, 74, 3321–3329.

Bodem, J., Schied, T., Gabriel, R., Rammling, M. and Rethwilm, A. (2011) Foamy virus nuclear RNA export is distinct from that of other retroviruses. *J Virol,* 85, 2333-2341.

Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V. and Cullen, B.R. (2006) Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc Natl Acad Sci USA*, 103, 8780-8785.

Bolinger, C. and Boris-Lawrie, K. (2009) Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. *Retrovirology*, 6, 8.

Bowles, J., Schepers, G. and Koopman, P. (2000) Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol*, 227, 239-255.

Braaten, D. and Luban, J. (2001) Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells. *EMBO J*, 20, 1300-1309.

Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K.T., Rekosh, D. and Hammarskjold, M.L. (1994) A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc Natl Acad Sci USA*, 91, 1256-1260.

Byun, H., Halani, N., Mertz, J.A., Ali, A.F., Lozano, M.M. and Dudley, J.P. (2010) Retroviral Rem protein requires processing by signal peptidase and retrotranslocation for nuclear function. *Proc Natl Acad Sci USA*, 107, 12287-12292.

Caporale, M., Centorame, P., Giovannini, A., Sacchini, F., Di Ventura, M., De las Heras, M. and Palmarini, M. (2005) Infection of lung epithelial cells and induction of pulmonary adenocarcinoma is not the most common outcome of naturally occurring JSRV infection during the commercial lifespan of sheep. *Virology*, 338, 144-153.

Caporale, M., Arnaud, F., Mura, M., Golder, M., Murgia, C. and Palmarini, M. (2009) The signal peptide of a simple retrovirus envelope functions as a posttranscriptional regulator of viral gene expression. *J Virol*, 83, 4591-4604.

Carlton, J.G. and Martin-Serrano, J. (2007) Parallels between cytokinesis and retroviral budding: a role for the ESCRT machinery. *Science*, 316, 1908-1912.

Cartellieri, M., Herchenröder, O., Rudolph, W., Heinkelein, M., Lindemann, D., Zentgraf, H. and Rethwilm, A. (2005) N-terminal Gag domain required for foamy virus particle assembly and export. *J Virol*, 79, 12464-12476.

Casartelli, N., Sourisseau, M., Feldmann, J., Guivel-Benhassine, F., Mallet, A., Marcelin, A.G., Guatelli, J. and Schwartz, O. (2010) Tetherin restricts productive HIV-1 cell-to-cell transmission. *PLoS Pathog*, 17, e1000955.

Chen, H., Lilley, C.E., Yu, Q., Lee, D.V., Chou, J., Narvaiza, I., Landau, N.R. and Weitzman, M.D. (2006) APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr Biol*, 16, 480-485.

Cherepanov, P., Maertens, G.N. and Hare, S. (2011) Structural insights into the retroviral DNA integration apparatus. *Curr Opin Struct Biol*, 21, 249-256.

Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, I., Kao, R.R., Pemberton, J.M., Beraldi, D., Stear, M.J., Alberti, A., Pittau, M., Iannuzzi, L., Banabazi, M.H., Kazwala, R.R., Zhang, Y.P., Arranz, J.J., Ali, B.A., Wang, Z., Uzun, M., Dione, M.M., Olsaker, I., Holm, L.E., Saarma, U., Ahmad, S., Marzanov, N., Eythorsdottir, E., Holland, M.J., Ajmone-Marsan, P., Bruford, M.W., Kantanen, J., Spencer, T.E. and Palmarini, M. (2009) Revealing the history of sheep domestication using retrovirus integrations. *Science*, 324, 532-536.

Chow, Y.H., Alberti, A., Mura, M., Pretto, C., Murcia, P., Albritton, L.M. and Palmarini, M. (2003) Transformation of rodent fibroblasts by the jaagsiekte sheep retrovirus envelope is receptor independent and does not require the surface domain. *J Virol*, 77, 6341-6350.

Clarke, M. and Spudich, J.A. (1977) Nonmuscle contractile proteins: the role of actin and myosin in cell motility and shape determination. *Annu Rev Biochem*, 46, 797-822.

Cochrane, A.W., McNally, M.T. and Mouland, A.J. (2006) The retrovirus RNA trafficking granule: from birth to maturity. *Retrovirology*, 3, 18.

Cochrane, A.W., Perkins, A. and Rosen, C.A. (1990) Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. *J Virol*, 64, 881-885.

Conticello, S.G., Thomas, C.J., Petersen-Mahrt, S.K. and Neuberger, M.S. (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol*, 22, 367-377.

Côté, M., Zheng, Y.M., Albritton, L.M. and Liu, S.L. (2008) Fusogenicity of Jaagsiekte sheep retrovirus envelope protein is dependent on low pH and is enhanced by cytoplasmic tail truncations. *J Virol*, 82, 2543-2554.

Côté, M., Zheng, Y.M. and Liu, S.L. (2009) Receptor binding and low pH coactivate oncogenic retrovirus envelope-mediated fusion. *J Virol*, 83, 11447-11455.

Cousens, C., Maeda, N., Murgia, C., Dagleish, M.P., Palmarini, M. and Fan, H. (2007) In vivo tumorigenesis by Jaagsiekte sheep retrovirus (JSRV) requires Y590 in Env TM, but not full-length orfX open reading frame. *Virology*, 367, 413-421.

D'Souza, V. and Summers, M.F. (2005) How retroviruses select their genomes. *Nat Rev Microbiol*, 3, 643-655.

Daelemans, D., Costes, S.V., Cho, E.H., Erwin-Cohen, R.A., Lockett, S. and Pavlakis, G.N. (2004) In vivo HIV-1 Rev multimerization in the nucleolus and cytoplasm identified by fluorescence resonance energy transfer. *J Biol Chem*, 279, 50167-50175.

Dalton, A.J., Melnick, J.L., Bauer, H., Beaudreau, G., Bentvelzen, P., Bolognesi, D., Gallo, R., Graffi, A., Haguenau, F., Heston, W., Huebner, R., Todaro, G. and Heine, U.I. (1974) The case for a family of reverse transcriptase viruses: Retraviridae. *Intervirology*, 4, 201-206.

Dawson, M., Done, S.H., Venables, C. and Jenkins, C.E. (1990) Maedi-visna and sheep pulmonary adenomatosis: a study of concurrent infection. *Br Vet J*, 146, 531-538.

Deaton, A.M. and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev*, 25, 1010-1022.

Deffaud, C. and Darlix, J.L. (2000a) Characterization of an internal ribosomal entry segment in the 5 leader of murine leukemia virus env RNA. *J Virol*, 74, 846-850.

Deffaud, C. and Darlix, J.L. (2000b) Rous sarcoma virus translation revisited: characterization of an internal ribosome entry segment in the 5leader of the genomic RNA. *J Virol*, 74, 11581-11588.

De las Heras, M., González, L. and Sharp, J.M. (2003) Pathology of ovine pulmonary adenocarcinoma. *Curr Top Microbiol Immunol*, 275, 25-54.

Delelis, O., Saïb, A. and Sonigo, P. (2003) Biphasic DNA synthesis in spumaviruses. *J Virol*, 77, 8141-8146.

DeMartini, J.C., Bishop, J.V., Allen, T.E., Jassim, F.A., Sharp, J.M., De las Heras, M., Voelker, D.R. and Carlson, J.O. (2001) Jaagsiekte sheep retrovirus proviral clone JSRV(JS7), derived from the JS7 lung tumor cell line, induces ovine pulmonary carcinoma and is integrated into the surfactant protein A gene. *J Virol*, 75, 4239-4246.

DeMartini, J.C., Rosadio, R.H. and Lairmore, M.D. (1988) The etiology and pathogenesis of ovine pulmonary carcinoma (sheep pulmonary adenomatosis). *Vet Microbiol*, 17, 219–236.

Denny, P., Swift, S., Brand, N., Dabhade, N., Barton, P. and Ashworth, A. (1992) A conserved family of genes related to the testis determining gene, SRY. *Nucleic Acids Res*, 20, 2887.

Denner, J. (2010) Endogenous retroviruses. *Retroviruses*. R. Kurth and N. Bannert. Nopen reading frameolk, Caister Academic Press: 35-70.

Deaton, A.M. and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev*, 25, 1010-1022.

Derse, D., Hill, S.A., Princler, G., Lloyd, P. and Heidecker, G. (2007) Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. *Proc Natl Acad Sci USA,* 104, 2915-2920.

Duda, A., Stange, A., Lüftenegger, D., Stanke, N., Westphal, D., Pietschmann, T., Eastman, S.W., Linial, M.L., Rethwilm, A. and Lindemann, D. (2004) Prototype foamy virus envelope glycoprotein leader peptide processing is mediated by a furin-like cellular protease, but cleavage is not essential for viral infectivity. *J Virol*, 78, 13865-13870.

Dultz, E., Hildenbeutel, M., Martoglio, B., Hochman, J., Dobberstein, B. and Kapp, K. (2008) The signal peptide of the mouse mammary tumor virus Rem protein is released from the endoplasmic reticulum membrane and accumulates in nucleoli. *J Biol Chem*, 283, 9966-9976.

Dunlap, K.A., Palmarini, M. and Spencer, T.E. (2006a) Ovine endogenous betaretroviruses (enJSRVs) and placental morphogenesis. *Placenta Suppl A*, S135–S140.

Dunlap, K.A., Palmarini, M., Varela, M., Burghardt, R.C., Hayashi, K., Farmer, J.L. and Spencer, T.E. (2006b) Endogenous retroviruses regulate periimplantation placental growth and differentiation. *Proc Natl Acad Sci USA*, 103, 14390-14395.

Dunn, B.M., Goodenow, M.M., Gustchina, A. and Wlodawer, A. (2002) Retroviral proteases. *Genome Biol*, 3, REVIEWS3006.

Dupressoir, A., Marceau, G., Vernochet, C., Benit, L., Kanellopoulos, C., Sapin, V. and Heidmann, T. (2005) Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc Natl Acad Sci USA*, 102, 725-730.

Dupressoir, A., Vernochet, C., Bawa, O., Harper, F., Pierron, G., Opolon, P. and Heidmann, T. (2009) Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. *Proc Natl Acad Sci USA*, 106, 12127-12132.

Ellermann, V. and Bang, O. (1908) Experimentelle Leukämie bei Hühnern. *Zentralbl Bakteriol Parasitenkd Infectionskr Hyg Abt Orig*, 46, 595-609.

Emmott, E. and Hiscox, J.A. (2009) Nucleolar targeting: the hub of the matter. *EMBO Rep*, 10, 231-238.

Engelman, A. (2010) Reverse transcription and integration. *Retroviruses*. R. Kurth and N. Bannert. Nopen reading frameolk, Caister Academic Press: 129-160.

Enssle, J., Jordan, I., Mauer, B. and Rethwilm, A. (1996) Foamy virus reverse transcriptase is expressed independently from the Gag protein. *Proc Natl Acad Sci USA*, 93, 4137-4141.

Esnault, C., Millet, J., Schwartz, O. and Heidmann, T. (2006) Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. *Nucleic Acids Res,* 34, 1522-1531.

Esnault, C., Priet, S., Ribet, D., Heidmann, O. and Heidmann, T. (2008) Restriction by APOBEC3 proteins of endogenous retroviruses with an extracellular life cycle: ex vivo effects and in vivo "traces" on the murine IAPE and human HERV-K elements. *Retrovirology*, 5, 75.

Evans, D.T., Serra-Moreno, R., Singh, R.K. and Guatelli, J.C. (2010) BST-2/tetherin: a new component of the innate immune response to enveloped viruses. *Trends Microbiol*, 18, 388-396.

Felber, B.K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T. and Pavlakis, G.N. (1989) Rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci USA*, 86, 1495-1499.

Fitzpatrick, K., Skasko, M., Deerinck, T.J., Crum, J., Ellisman, M.H. and Guatelli, J. (2010) Direct restriction of virus release and incorporation of the interferon-induced protein BST-2 into HIV-1 particles. *PLoS Pathog*, 6, e1000701.

Fontoura, B.M., Faria, P.A. and Nussenzveig, D.R. (2005) Viral interactions with the nuclear transport machinery: discovering and disrupting pathways. *IUBMB Life*, 57, 65-72.

Fornerod, M., Ohno, M., Yoshida, M. and Mattaj, I.W. (1997) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell*, 90, 1051-1060.

Forshey, B.M., von Schwedler, U., Sundquist, W.I. and Aiken, C. (2002) Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J Virol*, 76, 5667-5677.

Foster, J.W., Dominguez-Steglich, M.A., Guioli, S., Kwok, C., Weller, P.A., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I.D., Goodfellow, P.N. and et al. (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature*, 372, 525-530.

Franke, E.K., Yuan, H.E. and Luban, J. (1994) Specific incorporation of cyclophilin A into HIV-1 virions. *Nature*, 372, 359-362.

Freed, E.O. (2001) HIV-1 replication. *Somat Cell Mol Genet*, 26, 13-33.

Freed, E.O. (2002) Viral late domains. *J Virol*, 76, 4679-4687.

Freed, E.O. (2004) Mechanisms of enveloped virus release. *Virus Res*, 106, 85-86.

Friend, C. (1957) Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia. *J Exp Med*, 105, 307-318.

Gao, G., Guo, X. and Goff, S.P. (2002) Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein. *Science*, 297, 1703-1706.

Garrels, J.I. and Gibson, W. (1976) Identification and characterization of multiple forms of actin. *Cell*, 9, 793-805.

Geiselhart, V., Bastone, P., Kempf, T., Schnölzer, M. and Löchelt, M. (2004) Furin-mediated cleavage of the feline foamy virus Env leader protein. *J Virol*, 78, 13573-13581.

Gierasch, L.M. (1989) Signal sequences. *Biochemistry*, 28, 923-390.

Gifford, R.J., Katzourakis, A., Tristem, M., Pybus, O.G., Winters, M. and Shafer, R.W. (2008) A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. *Proc Natl Acad Sci USA*, 105, 20362-20367.

Gifford, R. and Tristem, M. (2003) The evolution, distribution and diversity of endogenous retroviruses. *Virus Genes,* 26, 291-315.

Godornes, C., Leader, B.T., Molini, B.J., Centurion-Lara, A. and Lukehart, S.A. (2007) Quantitation of rabbit cytokine mRNA by real-time RT-PCR. *Cytokine*, 38, 1-7.

Goff, S.P. (2007) Retroviridae: the retroviruses and their replication. *Fields virology*. D.M. Knipe and P.M. Howley. Philadelphia, Lippincott Williams and Wilkins: 1999–2069.

Golovkina, T.V., Chervonsky, A., Dudley, J.P. and Ross, S.R. (1992) Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. *Cell*, 69, 637-645.

Golovkina, T.V., Dudley, J.P. and Ross, S.R. (1998) B and T cells are required for mouse mammary tumor virus spread within the mammary gland. *J Immunol,* 161, 2375-2382.

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

Göttlinger, H.G and Weissenhorn, W. (2010) Assembly and release. *Retroviruses*. R. Kurth and N. Bannert. Nopen reading frameolk, Caister Academic Press: 187-216.

Gottwein, E., Bodem, J., Müller, B., Schmechel, A., Zentgraf, H. and Kräusslich, H.G. (2003) The Mason-Pfizer monkey virus PPPY and PSAP motifs both contribute to virus release. *J Virol*, 77, 9474-9485.

Grego, E., De Meneghi, D., Alvarez, V., Benito, A.A., Minguijón, E., Ortín, A., Mattoni, M., Moreno, B., Pérez de Villarreal, M., Alberti, A., Capucchio, M.T., Caporale, M., Juste, R., Rosati, S. and De las Heras, M. (2008) Colostrum and milk can transmit jaagsiekte retrovirus to lambs. *Vet Microbiol*, 130, 247-257.

Griffiths, D.J., Martineau, H.M. and Cousens, C. (2010) Pathology and pathogenesis of ovine pulmonary adenocarcinoma. *J Comp Pathol*, 142, 260-283.

Gross, L. (1951) "Spontaneous" leukemia developing in C3H mice following inoculation in infancy, with AK-leukemic extracts, or AK-embryos. *Proc Soc Exp Biol Med*, 76, 27-32.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature*, 346, 245-250.

Guo, X., Ma, J., Sun, J. and Gao, G. (2007) The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc Natl Acad Sci USA,* 104, 151-156.

Habermann, A., Krijnse-Locker, J., Oberwinkler, H., Eckhardt, M., Homann, S., Andrew, A., Strebel, K. and Kräusslich, H.G. (2010) CD317/tetherin is enriched in the HIV-1 envelope and downregulated from the plasma membrane upon virus infection. *J Virol*, 84, 4646-4658.

Hamard-Peron, E. and Muriaux, D. (2011) Retroviral matrix and lipids, the intimate interaction. *Retrovirology*, 7, 8-15.

Harel, J., Rassart, E. and Jolicoeur, P. (1981) Cell cycle dependence of synthesis of unintegrated viral DNA in mouse cells newly infected with murine leukemia virus. *Virology*, 110, 202-207.

Harris, J.R. (1991) The evolution of placental mammals. *FEBS Lett*, 295, 3-4.

Hartl, M.J., Bodem, J., Jochheim, F., Rethwilm, A., Rösch, P. and Wöhrl, B.M. (2011) Regulation of foamy virus protease activity by viral RNA: a novel and unique mechanism among retroviruses. *J Virol*, 85, 4462-4469.

Hassanin, A. and Douzery, E.J. (2003) Molecular and morphological phylogenies of ruminantia and the alternative position of the moschidae. *Syst Biol,* 52, 206-228.

Hauser, H., Lopez, L.A., Yang, S.J., Oldenburg, J.E., Exline, C.M., Guatelli, J.C. and Cannon, P.M. (2010) HIV-1 Vpu and HIV-2 Env counteract BST-2/tetherin by sequestration in a perinuclear compartment. *Retrovirology*, 7, 51.

Hecht, S.J., Stedman, K.E., Carlson, J.O. and DeMartini, J. C. (1996) Distribution of endogenous type B and type D sheep retrovirus sequences in ungulates and other mammals. *Proc Natl Acad Sci USA,* 93, 3297-3302.

Herbreteau, C.H., Weill, L., Décimo, D., Prévôt, D., Darlix, J.L., Sargueil, B. and Ohlmann, T. (2005) HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. *Nat Struct Mol Biol*, 12, 1001-1007.

Herschhorn, A. and Hizi, A. (2010) Retroviral reverse transcriptases. *Cell Mol Life Sci*, 67, 2717-2747.

Hernández, F.M. and Vrba, E.S. (2005) A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. *Biol Rev Camb Philos Soc*, 80, 269-302.

Higuchi, R., Dollinger, G., Walsh, P.S. and Griffith, R. (1992) Simultaneous amplification and detection of specific DNA sequences. *Biotechnology (NY)*, 10, 413-417.

Hilditch, L., Matadeen, R., Goldstone, D.C., Rosenthal, P.B., Taylor, I.A. and Stoye, J.P. (2011) Ordered assembly of murine leukemia virus capsid protein on lipid nanotubes directs specific binding by the restriction factor, Fv1. *Proc Natl Acad Sci USA*, 108, 5771-5776.

Hofacre, A., Nitta, T. and Fan, H. (2009) Jaagsiekte sheep retrovirus encodes a regulatory factor, Rej, required for synthesis of Gag protein. *J Virol*, 83, 12483-12498.

Hofmann, W., Schubert, D., LaBonte, J., Munson, L., Gibson, S., Scammell, J., Ferrigno, P. and Sodroski, J. (1999) Species-specific, postentry barriers to primate immunodeficiency virus infection. *J Virol*, 73, 10020-10028.

Holland, M.J., Palmarini, M., Garcia-Goti, M., Gonzalez, L., McKendrick, I., De Las Heras, M. and Sharp, J.M. (1999) Jaagsiekte retrovirus is widely distributed both in T and B lymphocytes and in mononuclear phagocytes of sheep with naturally and experimentally acquired pulmonary adenomatosis. *J Virol*, 73, 4004-4008.

Holmes, R.K., Malim, M.H. and Bishop, K.N. (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem Sci*, 32, 118-128.

Huebner, R.J. and Todaro, G.J. (1969) Oncogenes of RNA tumor viruses as determinants of cancer. *Proc Natl Acad Sci USA*, 64, 1087-1094.

Hulme, A.E., Perez, O. and Hope, T.J. (2011) Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc Natl Acad Sci USA*, 108, 9975-9980.

Hunter, E. (1997) Viral entry and receptors. *Retroviruses*. J.M. Coffin, S.H. Hughes and H.E. Varmus. New York, Cold Spring Harbour laboratory Press: 71-120.

Hunter, T. and Garrels, J.I. (1977) Characterization of the mRNAs for alpha-, beta- and gamma-actin. *Cell*, 12, 767-781.

ISCNDB2000. (2001) International system for chromosome nomenclature of domestic bovids. *Cytogenet Cell Genet*, 92, 283-299.

Ikeda, H. and Odaka, T. (1983) Cellular expression of murine leukemia virus gp70-related antigen on thymocytes of uninfected mice correlates with Fv-4 gene-controlled resistance to Friend leukemia virus infection. *Virology*, 128, 127-139.

Ikeda, H., Laigret, F., Martin, M.A. and Repaske, R. (1985) Characterization of a molecularly cloned retroviral sequence associated with Fv-4 resistance. *J Virol*, 55, 768-777.

Ikeda, T., Abd El Galil, K.H., Tokunaga, K., Maeda, K., Sata, T., Sakaguchi, N., Heidmann, T. and Koito, A. (2011) Intrinsic restriction activity by apolipoprotein B mRNA editing enzyme APOBEC1 against the mobility of autonomous retrotransposons. *Nucleic Acids Res*, 39, 5538-5554.

Indik, S., Günzburg, W.H., Salmons, B. and Rouault, F. (2005) A novel, mouse mammary tumor virus encoded protein with Rev-like properties. *Virology*, 337, 1-6.

Izard, J.W. and Kendall, D.A. (1994) Signal peptides: exquisitely designed transport promoters. *Mol Microbiol*, 13, 765-773.

Jarrett, W.F., Crawford, E.M., Martin, W.B. and Davie, F. (1964a) A virus-like particle associated with leukemia (lymphosarcoma). *Nature*, 202, 567-569.

Jarrett, W.F., Martin, W.B, Crighton, G.W., Dalton, R.G and Stewart, M.F. (1964b) Transmission experiments with leukemia (lymphosarcoma). *Nature*, 202, 566-567.

Jern, P. and Coffin, J.M. (2008) Effects of retroviruses on host genome function. *Annu Rev Genet*, 42, 709-732.

Jern, P., Sperber, G.O. and Blomberg, J. (2005) Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. *Retrovirology*, 2, 50.

Johnson, S.F. and Telesnitsky, A. (2010) Retroviral RNA dimerization and packaging: the what, how, when, where, and why. *PLoS Pathog*, 6, e1001007.

Johnson, W.E. and Sawyer, S.L. (2009) Molecular evolution of the antiretroviral TRIM5 gene. *Immunogenetics*, 61, 163-176.

Juretzek, T., Holm, T., Gartner, K., Kanzler, S., Lindemann, D., Herchenroder, O., Picard-Maureau, M., Rammling, M., Heinkelein, M. and Rethwilm, A. (2004) Foamy virus integration. *J Virol*, 78, 2472-2477.

Kalter, S.S., Helmke, R.J., Heberling, R.L., Panigel, M., Fowler, A.K., Strickland, J.E. and Hellman, A. (1973) Brief communication: C-type particles in normal human placentas. *J Natl Cancer Inst*, 50, 1081-1084.

Katzourakis, A., Tristem, M., Pybus, O.G. and Gifford, R.J. (2007) Discovery and analysis of the first endogenous lentivirus. *Proc Natl Acad Sci USA*, 104, 6261-6265.

Kawakami, T.G., Huff, S.D., Buckley, P.M., Dungworth, D.L., Synder, S.P. and Gilden, R.V. (1972) C-type virus associated with gibbon lymphosarcoma. *Nat New Biol*, 235, 170-171.

Kerns, J.A., Emerman, M. and Malik, H.S. (2008) Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. *PLoS Genet*, 4, e21.

Korman, A.J., Bourgarel, P., Meo, T. and Rieckhof, G.E. (1992) The mouse mammary tumor virus long terminal repeat encodes a type II transmembrane glycoprotein. *EMBO J,* 11, 1901–1905.

Kost, B., Bao, Y.Q. and Chua, N.H. (2002) Cytoskeleton and plant organogenesis. *Philos Trans R Soc Lond B Biol Sci*, 357, 777-789.

Kozak, C.A., Gromet, N.J., Ikeda, H. and Buckler, C.E. (1984) A unique sequence related to the ecotropic murine leukemia virus is associated with the Fv-4 resistance gene. *Proc Natl Acad Sci USA*, 81, 834-837.

Kubista, M., Andrade, J.M., Bengtsson, M., Forootan, A., Jonak, J., Lind, K., Sindelka, R., Sjoback, R., Sjogreen, B., Strombom, L., Stahlberg, A. and Zoric, N. (2006) The real-time polymerase chain reaction. *Mol Aspects Med*, 27, 95-125.

Kubota, S., Copeland, T.D. and Pomerantz, R.J. (1999) Nuclear and nucleolar targeting of human ribosomal protein S25: common features shared with HIV-1 regulatory proteins. *Oncogene*, 18, 1503-1514.

Kutay, U., Ahnert-Hilger, G., Hartmann, E., Wiedenmann, B. and Rapoport, T.A. (1995) Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. *EMBO J*, 14, 217-223.

Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 5, 105-132.

Larsson, E., Kato, N. and Cohen, M. (1989) Human endogenous proviruses. *Curr Top Microbiol Immunol*, 148, 115-132.

LaRue, R.S., Jónsson, S.R., Silverstein, K.A., Lajoie, M., Bertrand, D., El-Mabrouk, N., Hötzel, I., Andrésdóttir, V., Smith, T.P. and Harris, R.S. (2008) The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC Mol Biol*, 9, 104.

Lavie, L., Medstrand, P., Schempp, W., Meese, E. and Mayer, J. (2004) Human endogenous retrovirus family HERV-K(HML-5): status, evolution, and reconstruction of an ancient betaretrovirus in the human genome. *J Virol*, 78, 8788-8798.

Leavitt, J., Bushar, G., Kakunaga, T., Hamada, H., Hirakawa, T., Goldman, D. and Merril, C. (1982) Variations in expression of mutant beta actin accompanying incremental increases in human fibroblast tumorigenicity. *Cell*, 28, 259-268.

Lehmann-Che, J., Renault, N., Giron, M.L., Roingeard, P., Clave, E., Tobaly-Tapiero, J., Bittoun, P., Toubert, A., de The, H. and Saib, A. (2007) Centrosomal latency of incoming foamy viruses in resting cells. *PLoS Pathog*, 3, e74.

Lenzini, E., Leszl, A., Artifoni, L., Casellato, R., Tenconi, R. and Baccichetti, C. (1988) Partial duplication of 17 long arm. *Ann Genet*, 31, 175-180.

Leroux, C., Girard, N., Cottin, V., Greenland, T., Mornex, J.F. and Archer, F. (2007) Jaagsiekte Sheep Retrovirus (JSRV): from virus to lung cancer in sheep. *Vet Res*, 38, 211-228.

Lewinski, M.K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., Berry, C.C., Ecker, J.R. and Bushman, F.D. (2006) Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog*, 2, e60.

Lewis, P., Hensel, M. and Emerman, M. (1992) Human immunodeficiency virus infection of cells arrested in the cell cycle. *Embo J*, 11, 3053-3058.

Li, X., Krishnan, L., Cherepanov, P. and Engelman, A. (2011) Structural biology of retroviral DNA integration. *Virology*, 411, 194-205.

Li, Y., Bor, Y.C., Misawa, Y., Xue, Y., Rekosh, D. and Hammarskjold, M.L. (2006) An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature*, 443, 234-237.

Lilly, F. (1967) Susceptibility to two strains of Friend leukemia virus in mice. *Science*, 155, 461-462.

Lilly, F. and Steeves, R.A. (1973) B-tropic Friend virus: a host-range pseudotype of spleen focus-forming virus (SFFV). *Virology*, 55, 363-370.

Lindemann, D., Pietschmann, T., Picard-Maureau, M., Berg, A., Heinkelein, M., Thurow, J., Knaus, P., Zentgraf, H. and Rethwilm, A. (2001) A particle-associated glycoprotein signal peptide essential for virus maturation and infectivity. *J Virol*, 75, 5762-5771.

Lindemann, D. and Rethwilm, A. (2011) Foamy virus biology and its application for vector development. *Viruses*, 3, 561-585.

Löchelt, M., Romen, F., Bastone, P., Muckenfuss, H., Kirchner, N., Kim, Y.B., Truyen, U., Rösler, U., Battenberg, M., Saib, A., Flory, E., Cichutek, K. and Münk, C. (2005) The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proc Natl Acad Sci USA,* 102, 7982-7987.

Long, Q., Bengra, C., Li, C., Kutlar, F. and Tuan, D. (1998) A long terminal repeat of the human endogenous retrovirus ERV-9 is located in the 5' boundary area of the human beta-globin locus control region. *Genomics*, 54, 542-555.

Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) Protein measurement with the folin phenol reagent. *J Biol Chem* 193, 265–275.

Luban, J. (2010) Retroviral restriction factors. *Retroviruses*. R. Kurth and N. Bannert. Nopen reading frameolk, Caister Academic Press: 285-308.

Lydyard, P., Whelan, A. and Fanger, M. (2004) *Instant Notes: Immunology*. 2nd edition. 332 pp., Bios/Taylor & Francis, New York and Abingdon, 2004, ISBN 1-8599-6039-1.

Maeda, N., Fu, W., Ortín, A., De las Heras, M. and Fan, H. (2005) Roles of the Ras-MEK-mitogen-activated protein kinase and phosphatidylinositol 3-kinase-Akt-mTOR pathways in Jaagsiekte sheep retrovirus-induced transformation of rodent fibroblast and epithelial cell lines. *J Virol*, 79, 4440-4450.

Maeda,N., Palmarini, M., Murgia,C. and Fan, H. (2001) Direct transformation of rodent fibroblasts by Jaagsiekte sheep retrovirus DNA. *Proc Natl Acad Sci USA*, 98, 4449–4454.

Maksakova, I.A., Mager, D.L. and Reiss, D. (2008) Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. *Cell Mol Life Sci*, 65, 3329-3347.

Maksakova, I.A., Romanish, M.T., Gagnier, L., Dunn, C.A., van de Lagemaat, L.N. and Mager, D.L. (2006) Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet,* 2, e2.

Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V. and Cullen, B.R. (1989) The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*, 338, 254-257.

Malmquist, W.A., Krauss, H.H., Moulton, J.E. and Wandera, J.G. (1972) Morphologic study of virus-infected lung cell cultures from sheep pulmonary adenomatosis (Jaagsiekte). *Lab Invest*, 26, 528–533.

Mangeney, M., Renard, M., Schlecht-Louf, G., Bouallaga, I., Heidmann, O., Letzelter, C., Richaud, A., Ducos, B. and Heidmann, T. (2007) Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc Natl Acad Sci USA*, 104, 20534-20539.

Mar, J.C., Kimura, Y., Schroder, K., Irvine, K.M., Hayashizaki, Y., Suzuki, H., Hume, D. and Quackenbush, J. (2009) Data-driven normalization strategies for high-throughput quantitative RT-PCR. *BMC Bioinformatics*, 10, 110.

Martin, W.B., Scott, F.M., Sharp, J.M., Angus, K.W. and Norval, M. (1976) Experimental production of sheep pulmonary adenomatosis (Jaagsiekte). *Nature*, 264, 183-185.

Martins, H. and Villesen, P. (2011) Improved integration time estimation of endogenous retroviruses with phylogenetic data. *PLoS One*, 6, e14745.

McClintock, B. (1950) The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA*, 36, 344-355.

McGee-Estrada, K. and Fan, H. (2006) In vivo and in vitro analysis of factor binding sites in Jaagsiekte sheep retrovirus long terminal repeat enhancer sequences: roles of HNF-3, NF-I, and C/EBP for activity in lung epithelial cells. *J Virol*, 80, 332-341.

McGee-Estrada, K., Palmarini, M. and Fan, H. (2002) HNF-3beta is a critical factor for the expression of the Jaagsiekte sheep retrovirus long terminal repeat in type II pneumocytes but not in Clara cells. *Virology*, 292, 87-97.

Mehle, A., Strack, B., Ancuta, P., Zhang, C., McPike, M. and Gabuzda, D. (2004) Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J Biol Chem*, 279, 7792-7798.

Melikyan, G.B. (2008) Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology*, 5, 111.

Mermer, B., Felber, B.K., Campbell, M. and Pavlakis, G.N. (1990) Identification of trans-dominant HIV-1 rev protein mutants by direct transfer of bacterially produced proteins into human cells. *Nucleic Acids Res*, 18, 2037-2044.

Mertz, J.A., Chadee, A.B., Byun, H., Russell, R. and Dudley, J.P. (2009) Mapping of the functional boundaries and secondary structure of the mouse mammary tumor virus Rem-responsive element. *J Biol Chem*, 284, 25642-25652.

Mertz, J.A., Simper, M.S., Lozano, M.M., Payne, S.M. and Dudley, J.P. (2005) Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. *J Virol*, 79, 14737-14747.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., Keith, J.C. Jr and McCoy, J.M. (2000) Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*, 403, 785-789.

Michienzi, A., De Angelis, F.G., Bozzoni, I. and Rossi J.J. (2006) A nucleolar localizing Rev binding element inhibits HIV replication. *AIDS Res Ther*, 3, 13.

Miyakawa, K., Ryo, A., Murakami, T., Ohba, K., Yamaoka, S., Fukuda, M., Guatelli, J. and Yamamoto, N. (2009) BCA2/Rabring7 promotes tetherin-dependent HIV-1 restriction. *PLoS Pathog*, 5, e1000700.

Mothes, W., Boerger, A.L., Narayan, S., Cunningham, J.M. and Young, J.A. (2000) Retroviral entry mediated by receptor priming and low pH triggering of an envelope glycoprotein. *Cell*, 103, 679-689.

Mothes, W. and Uchil, P. D (2010) Retroviral entry and uncoating. *Retroviruses*. R. Kurth and N. Bannert. Nopen reading frameolk, Caister Academic Press: 107-128.

Müller, S., Möller, P., Bick, M.J., Wurr, S., Becker, S., Günther, S. and Kümmerer, B.M. (2007) Inhibition of filovirus replication by the zinc finger antiviral protein. *J Virol*, 81, 2391-2400.

Mura, M., Murcia, P., Caporale, M., Spencer, T.E., Nagashima, K., Rein, A. and Palmarini, M. (2004) Late viral interference induced by transdominant Gag of an endogenous retrovirus. *Proc Natl Acad Sci USA*, 101, 11117-11122.

Murcia, P.R., Arnaud, F. and Palmarini, M. (2007) The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of Jaagsiekte sheep retrovirus Gag. *J Virol*, 81, 1762-1772.

Murgia, C., Caporale, M., Ceesay, O., Di Francesco, G., Ferri, N., Varasano, V., De las Heras, M. and Palmarini, M. (2011) Lung adenocarcinoma originates from retrovirus infection of proliferating type 2 pneumocytes during pulmonary post-natal development or tissue repair. *PLoS Pathog*, 7, e1002014.

Muriaux, D. and Darlix, J.L. (2010) Properties and functions of the nucleocapsid protein in virus assembly. *RNA Biol*, 7, 744-753.

Murray, S.M., Picker, L.J., Axthelm, M.K., Hudkins, K., Alpers, C.E. and Linial, M.L. (2008) Replication in a superficial epithelial cell niche explains the lack of pathogenicity of primate foamy virus infections. *J Virol*, 82, 5981-5985.

Mustafa, F., Lozano, M. and Dudley, J.P. (2000) C3H mouse mammary tumor virus superantigen function requires a splice donor site in the envelope gene. *J Virol*, 74, 9431-9440.

Narayan, S. and Young, J.A. (2004) Reconstitution of retroviral fusion and uncoating in a cell-free system. *Proc Natl Acad Sci USA*, 101, 7721-7726.

Neil, S.J., Zang, T. and Bieniasz, P.D. (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, 451, 425-430.

Netzer, K.O., Rethwilm, A., Maurer, B. and ter Meulen, V. Identification of the major immunogenic structural proteins of human foamy virus. *J Gen Virol*, 71, 1237-1241.

Nicot, C., Harrod, R.L., Ciminale, V. and Franchini, G. (2005) Human T-cell leukemia/lymphoma virus type 1 nonstructural genes and their functions. *Oncogene*, 24, 6026-6034.

Nisole, S., Lynch, C., Stoye, J.P. and Yap, M.W. (2004) A Trim5-cyclophilin A fusion protein found in owl monkey kidney cells can restrict HIV-1. *Proc Natl Acad Sci USA*, 101, 13324-13328.

Nitta, T., Hofacre, A., Hull, S. and Fan, H. (2009) Identification and mutational analysis of a Rej response element in Jaagsiekte sheep retrovirus RNA. *J Virol*, 83, 12499-12511.

Ogert, R.A., Lee, L.H. and Beemon, K.L. (1996) Avian retroviral RNA element promotes unspliced RNA accumulation in the cytoplasm. *J Virol*, 70, 3834-3843.

Ohlmann, T., Lopez-Lastra, M. and Darlix, J.L. (2000) An internal ribosome entry segment promotes translation of the simian immunodeficiency virus genomic RNA. *J Biol Chem*, 275, 11899-11906.

Onafuwa-Nuga, A.A., Telesnitsky, A. and King, S.R. (2006) 7SL RNA, but not the 54-kd signal recognition particle protein, is an abundant component of both infectious HIV-1 and minimal virus-like particles. *Rna*, 12, 542-546.

Ono A. (2010) Relationships between plasma membrane microdomains and HIV-1 assembly. *Biol Cell*, 102, 335-350.

Ono, M., Yasunaga, T., Miyata, T. and Ushikubo, H. (1986) Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. *J Virol*, 60, 589-598.

Ortín, A., Minguijón, E., Dewar, P., García, M., Ferrer, L.M., Palmarini, .M, Gonzalez, L., Sharp, J.M. and De las Heras, M. (1998) Lack of a specific immune response against a recombinant capsid protein of Jaagsiekte sheep retrovirus in sheep and goats naturally affected by enzootic nasal tumour or sheep pulmonary adenomatosis. *Vet Immunol Immunopathol*, 61, 229–237.

Paetzel, M., Karla, A., Strynadka, N.C. and Dalbey, R.E. (2002) Signal peptidases. *Chem Rev*, 102, 4549-4580.

Palmarini, M., Cousens, C., Dalziel, R.G., Bai, J., Stedman, K., DeMartini, J.C. and Sharp, J.M. (1996a) The exogenous form of Jaagsiekte retrovirus is specifically associated with a contagious lung cancer of sheep. *J Virol,* 70, 1618-1623.

Palmarini, M. and Fan, H. (2001) Retrovirus-induced ovine pulmonary adenocarcinoma, an animal model for lung cancer. *J Natl Cancer Inst,* 93, 1603-1614.

Palmarini, M. and Fan, H. (2003) Molecular biology of jaagsiekte sheep retrovirus. *Curr Top Microbiol Immunol*, 275, 81-115.

Palmarini, M., Gray, C.A., Carpenter, K., Fan, H., Bazer, F.W. and Spencer, T.E. (2001a) Expression of endogenous betaretroviruses in the ovine uterus: effects of neonatal age, estrous cycle, pregnancy, and progesterone. *J Virol*, 75, 11319–11327.

Palmarini, M., Hallwirth, C., York, D., Murgia, C., de Oliveira, T., Spencer, T. and Fan, H. (2000) Molecular cloning and functional analysis of three type D endogenous retroviruses of sheep reveal a different cell tropism from that of the highly related exogenous Jaagsiekte sheep retrovirus. *J Virol*, 74, 8065–8076.

Palmarini, M., Holland, M.J., Cousens, C., Dalziel, R.G. and Sharp, J.M. (1996b) Jaagsiekte retrovirus establishes a disseminated infection of the lymphoid tissues of sheep affected by pulmonary adenomatosis. *J Gen Virol*, 77, 2991-2998.

Palmarini, M., Maeda, N., Murgia, C., De-Fraja, C., Hofacre, A. and Fan, H. (2001b) A phosphatidylinositol 3-kinase docking site in the cytoplasmic tail of the Jaagsiekte sheep retrovirus transmembrane protein is essential for envelope-induced transformation of NIH 3T3 cells. *J Virol*, 75, 11002-11009.

Palmarini, M., Murgia, C. and Fan, H. (2002) Spliced and prematurely polyadenylated Jaagsiekte sheep retrovirus-specific RNAs from infected or transfected cells. *Virology*, 294, 180-188.

Palmarini, M., Sharp, J.M., De las Heras, M. and Fan, H. (1999a) Jaagsiekte sheep retrovirus is necessary and sufficient to induce a contagious lung cancer in sheep. *J Virol,* 73, 6964-6972.

Palmarini, M., Sharp, J.M., Lee, C., Fan, H. (1999b) In vitro infection of ovine cell lines by Jaagsiekte sheep retrovirus. *J Virol*, 73, 10070-10078.

Patthy, L. (1999) Genome evolution and the evolution of exon-shuffling--a review. *Gene*, 238, 103-114.

Payen, E.J. and Cotinot, C.Y. (1994) Sequence evolution of SRY gene within Bovidae family. *Mamm Genome*, 5, 723-725.

Payne, L.N. and Chubb, R.C. (1968) Studies on the nature and genetic control of an antigen in normal chick embryos which reacts in the COFAL test. *J Gen Virol*, 3, 379-391.

Perez-Caballero, D., Zang, T., Ebrahimi, A., McNatt, M.W., Gregory, D.A., Johnson, M.C. and Bieniasz, P.D. (2009) Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell*, 139, 499-511.

Perk, K., Hod, I. and Nobel, T.A. (1971) Pulmonary adenomatosis of sheep (Jaagsiekte). I. Ultrastructure of the tumor. *J Natl Cancer Inst*, 46, 525–537.

Perk, K., Michalides, R., Spiegelman, S. and Schlom, J. (1974) Biochemical and morphologic evidence for the presence of an RNA tumor virus in pulmonary carcinoma of sheep (Jaagsiekte). *J Natl Cancer Inst*, 53, 131-135.

Perkovic, M., Schmidt, S., Marino, D., Russell, R.A., Stauch, B., Hofmann, H., Kopietz, F., Kloke, B.P., Zielonka, J., Ströver, H., Hermle, J., Lindemann, D., Pathak, V.K., Schneider, G., Löchelt, M., Cichutek, K. and Münk, C. (2009) Species-specific inhibition of APOBEC3C by the prototype foamy virus protein bet. *J Biol Chem*, 284, 5819-5826.

Perl, A., Rosenblatt, J.D., Chen, I.S., DiVincenzo, J.P., Bever, R., Poiesz, B.J. and Abraham, G.N. (1989) Detection and cloning of new HTLV-related endogenous sequences in man. *Nucleic Acids Res*, 17, 6841-6854.

Pertel, T., Hausmann, S., Morger, D., Züger, S., Guerra, J., Lascano, J., Reinhard, C., Santoni, F.A., Uchil, P.D., Chatel, L., Bisiaux, A., Albert, M.L., Strambio-De-Castillia, C., Mothes, W., Pizzato, M., Grütter, M.G. and Luban, J. (2011) TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature*, 472, 361-365.

Pi, W., Zhu, X., Wu, M., Wang, Y., Fulzele, S., Eroglu, A., Ling, J. and Tuan, D. (2010) Long-range function of an intergenic retrotransposon. *Proc Natl Acad Sci USA*, 107, 12992-12997.

Pietschmann, T., Heinkelein, M., Heldmann, M., Zentgraf, H., Rethwilm, A. and Lindemann, D. (1999) Foamy virus capsids require the cognate envelope protein for particle export. *J Virol*, 73, 2613-2621.

Pike, L.J. (2006) Rafts defined: a report on the Keystone Symposium on Lipid Rafts and Cell Function. *J Lipid Res*, 47, 1597-1598.

Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. and Gallo, R.C. (1980) Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci USA*, 77, 7415-7419.

Pollard, V.W. and Malim, M.H. (1998) The HIV-1 Rev protein. *Annu Rev Microbiol*, 52, 491-532.

Popovici, V., Goldstein, D.R., Antonov, J., Jaggi, R., Delorenzi, M. and Wirapati, P. (2009) Selecting control genes for RT-QPCR using public microarray data. *BMC Bioinformatics*, 10, 42.

Purcell, D.F. and Martin, M.A. (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol*, 67, 6365-6378.

Quimby, B.B. and Dasso, M. (2003) The small GTPase Ran: interpreting the signs. *Curr Opin Cell Biol*, 15, 338-344.

Rabson, A.B. and Graves, B.J. (1997) Synthesis and processing of viral RNA. *Retroviruses*. J.M. Coffin, S.H. Hughes and H.E. Varmus. New York, Cold Spring Harbour Laboratory Press: 205-261.

Rai, S.K., Duh, F.M., Vigdorovich, V., Danilkovitch-Miagkova, A., Lerman, M.I. and Miller, A.D. (2001) Candidate tumor suppressor HYAL2 is a glycosylphosphatidylinositol (GPI)-anchored cell-surface receptor for Jaagsiekte sheep retrovirus, the envelope protein of which mediates oncogenic transformation. *Proc Natl Acad Sci USA*, 98, 4443-4448.

Redmond, S., Peters, G. and Dickson, C. (1984) Mouse mammary tumor virus can mediate cell fusion at reduced pH. *Virology*, 133, 393-402.

Renault, N., Tobaly-Tapiero, J., Paris, J., Giron, M.L., Coiffic, A., Roingeard, P. and Saïb, A. (2011) A nuclear export signal within the structural Gag protein is required for prototype foamy virus replication. *Retrovirology*, 8, 6.

Rethwilm, A. (2010) Molecular biology of foamy viruses. *Med Microbiol Immunol*, 199, 197-207.

Rhee, S.S., Hui, H.X. and Hunter, E. (1990) Preassembled capsids of type D retroviruses contain a signal sufficient for targeting specifically to the plasma membrane. *J Virol*, 64, 3844-3852.

Rhee, S.S. and Hunter, E. (1990) A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a type C retrovirus. *Cell*, 63, 77-86.

Riviere, L., Darlix, J.L. and Cimarelli, A. (2010) Analysis of the viral elements required in the nuclear import of HIV-1 DNA. *J Virol*, 84, 729-739.

Rosati, S., Pittau, M., Alberti, A., Pozzi, S., York, D.F., Sharp, J.M. and Palmarini, M. (2000) An accessory open reading frame (orf-x) of jaagsiekte sheep retrovirus is conserved between different virus isolates. *Virus Res*, 66, 109-116.

Ross, S.R. (2009) Are viruses inhibited by APOBEC3 molecules from their host species? *PLoS Pathog*, 5, e1000347.

Ross, S.R., Schofield, J.J., Farr, C.J. and Bucan, M. (2002) Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *Proc Natl Acad Sci USA*, 99, 12386-12390.

Rous, P. (1911) A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *J Exp Med*, 13, 397-411.

Rowe, H.M. and Trono, D. (2011) Dynamic control of endogenous retroviruses during development. *Virology*, 411, 273-287.

Sarafianos, S.G., Das, K., Tantillo, C., Clark, A.D. Jr, Ding, J., Whitcomb, J.M., Boyer, P.L., Hughes, S.H. and Arnold, E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J*, 20, 1449-1461.

Sawyer, S.L., Wu, L.I., Emerman, M. and Malik, H.S. (2005) Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci USA,* 102, 2832-2837.

Sayah, D.M., Sokolskaja, E., Berthoux, L. and Luban, J. (2004) Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature*, 430, 569-573.

Sharp, J.M. and DeMartini, J.C. (2003) Natural history of JSRV in sheep. *Curr Top Microbiol Immunol,* 275, 55-79.

Sharp, J.M. and Herring, A.J. (1983) Sheep pulmonary adenomatosis: demonstration of a protein which cross- reacts with the major core proteins of Mason-Pfizer monkey virus and mouse mammary tumour virus. *J Gen Virol,* 64, 2323–2327.

Shaw, K.L., Lindemann, D., Mulligan, M.J. and Goepfert, P.A. (2003) Foamy virus envelope glycoprotein is sufficient for particle budding and release. *J Virol*, 77, 2338-2348.

Schmidt, A. and Hall, M.N. (1998) Signaling to the actin cytoskeleton. *Annu Rev Cell Dev Biol*, 14, 305-338.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R. and Bushman, F. (2002) HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, 110, 521-529.

Sfakianos, J.N. and Hunter, E. (2003) M-PMV capsid transport is mediated by Env/Gag interactions at the pericentriolar recycling endosome. *Traffic,* 4, 671-680.

Sheehy, A.M., Gaddis, N.C., Choi, J.D. and Malim, M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, 418, 646-650.

Snyder, S.P., DeMartini, J.C., Ameghino, E. and Caletti, E. (1983) Coexistence of pulmonary adenomatosis and progressive pneumonia in sheep in the central sierra of Peru. *Am J Vet Res*, 44, 1334-1348.

Sokolskaja, E., Berthoux, L. and Luban, J. (2006) Cyclophilin A and TRIM5alpha independently regulate human immunodeficiency virus type 1 infectivity in human cells. *J Virol*, 80, 2855-2862.

Song, B., Gold, B., O'Huigin, C., Javanbakht, H., Li, X., Stremlau, M., Winkler, C., Dean, M. and Sodroski, J. (2005) The B30.2(SPRY) domain of the retroviral restriction factor TRIM5alpha exhibits lineage-specific length and sequence variation in primates. *J Virol*, 79, 6111-6121.

Spencer, T.E., Johnson, G.A., Bazer, F.W. and Burghardt, R.C. (2004) Implantation mechanisms: insights from the sheep. *Reproduction*, 128, 657-668.

Spencer, T.E., Mura, M., Gray, C.A., Griebel, P.J. and Palmarini, M. (2003) Receptor usage and fetal expression of ovine endogenous betaretroviruses: implications for coevolution of endogenous and exogenous retroviruses. *J Virol,* 77, 749-753.

Spencer, T.E., Stagg, A.G., Joyce, M.M., Jenster, G., Wood, C.G., Bazer, F.W., Wiley, A.A. and Bartol, F.F. (1999) Discovery and characterization of endometrial epithelial messenger ribonucleic acids using the ovine uterine gland knockout model. *Endocrinology,* 140, 4070–4080.

Stange, A., Lüftenegger, D., Reh, J., Weissenhorn, W. and Lindemann, D. (2008) Subviral particle release determinants of prototype foamy virus. *J Virol*, 82, 9858-9869.

Stanke, N., Stange, A., Lüftenegger, D., Zentgraf, H. and Lindemann, D. (2005) Ubiquitination of the prototype foamy virus envelope glycoprotein leader peptide regulates subviral particle release. *J Virol*, 79, 15074-15083.

Stauber, R., Gaitanaris, G.A. and Pavlakis, G.N. (1995) Analysis of trafficking of Rev and transdominant Rev proteins in living cells using green fluorescent protein fusions: transdominant Rev blocks the export of Rev from the nucleus to the cytoplasm. *Virology*, 213, 439-449.

Steeves, R. and Lilly, F. (1977) Interactions between host and viral genomes in mouse leukemia. *Annu Rev Genet*, 11, 277-296.

Stevenson, M. (2000) HIV nuclear import: What's the flap? *Nat Med*, 6, 626-628.

Stoye, J.P. (2001) Endogenous retroviruses: still active after all these years? *Curr Biol,* 11, R914-916.

Stoye, J.P. (2009) Proviral protein provides placental function. *Proc Natl Acad Sci USA*, 106, 11827-118228.

Strebel, K., Luban, J. and Jeang, K.T. (2009) Human cellular restriction factors that target HIV-1 replication. *BMC Med*, 7, 48.

Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P. and Sodroski, J. (2004) The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature*, 427, 848-853.

Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., Diaz-Griffero, F., Anderson, D.J., Sundquist, W.I. and Sodroski, J. (2006) Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci USA*, 103, 5514-5519.

Stremlau, M., Perron, M., Welikala, S. and Sodroski, J. (2005) Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. *J Virol,* 79, 3139-3145.

Suzuki, Y. and Craigie, R. (2007) The road to chromatin-nuclear entry of retroviruses. *Nat Rev Microbiol*, 5, 187-196.

Swanstrom, R. and Wills, J.W. (1997) Synthesis, assembly, and processing of viral proteins. *Retroviruses*. J.M. Coffin, S.H. Hughes and H.E. Varmus. New York, Cold Spring Harbour Laboratory Press: 263-364.

Takagi, S. (2003) Actin-based photo-orientation movement of chloroplasts in plant cells. *J Exp Biol*, 206, 1963-1969.

Tarlinton, R.E., Meers, J. and Young, P.R. (2006) Retroviral invasion of the koala genome. *Nature*, 442, 79-81.

Telesnitsky, A. and Goff, S.P. (1997) Reverse transcriptase and the generation of retroviral DNA. *Retroviruses*. J.M. Coffin, S.H. Hughes and H.E. Varmus. New York, Cold Spring Harbour Laboratory Press: 121-160.

Temin, H.M. and Rubin, H. (1958) Characteristics of an assay for Rous sarcoma virus and Rous sarcoma cells in tissue culture. *Virology*, 6, 669-688.

Temin, H.M. (1964) Nature of the provirus of Rous sarcoma. *Natl Cancer Inst Monogr,* 17, 557-570.

Temin, H.M. and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*, 226, 1211-1213.

Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C.T., Sodroski, J. and Göttlinger, H.G. (1994) Functional association of cyclophilin A with HIV-1 virions. *Nature*, 372, 363-365.

Theilen, G.H., Gould, D., Fowler, M. and Dungworth, D.L. (1971) C-type virus in tumor tissue of a woolly monkey (Lagothrix spp.) with fibrosarcoma. *J Natl Cancer Inst*, 47, 881-889.

Towers, G.J. (2007) The control of viral infection by tripartite motif proteins and cyclophilin A. *Retrovirology*, 4, 40.

Trobridge, G.D., Miller, D.G., Jacobs, M.A., Allen, J.M., Kiem, H.P., Kaul, R. and Russell, D.W. (2006) Foamy virus vector integration sites in normal human cells. *Proc Natl Acad Sci USA*, 103, 1498-1503.

Tustin, R.C. (1969) Ovine Jaagsiekte. *J S Afr Vet Med Assoc*, 40, 3–23.

Uchil, P.D., Quinlan, B.D., Chan, W.T., Luna, J.M. and Mothes, W. (2008) TRIM E3 ligases interfere with early and late stages of the retroviral life cycle. *PLoS Pathog*, 4, e16.

Vagner, S., Waysbort, A., Marenda, M., Gensac, M.C., Amalric, F. and Prats, A.C. (1995) Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor. *J Biol Chem*, 270, 20376-20383.

Vallée, H. and Carré, H. (1904) Sur la nature infectieuse de l'anémie du cheval. *C. R. Hebd Seances Acad Sci Ser D Sci Nat*, 139, 331-333.

Van Damme, N., Goff, D., Katsura, C., Jorgenson, R.L., Mitchell, R., Johnson, M.C., Stephens, E.B. and Guatelli, J. (2008) The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe*, 3, 245-252.

Van Valen, L. (1973) A new evolutionary law. *Evol Theory*, 1, 1–30.

Varela, M., Chow, Y.H., Sturkie, C., Murcia, P. and Palmarini, M. (2006) Association of RON tyrosine kinase with the Jaagsiekte sheep retrovirus envelope glycoprotein. *Virology*, 350, 347-357.

Varela, M., Spencer, T.E., Palmarini, M. and Arnaud, F. (2009) Friendly viruses: the special relationship between endogenous retroviruses and their host. *Ann N Y Acad Sci*, 1178, 157-172.

Vardabasso, C., Manganaro, L., Lusic, M., Marcello, A. and Giacca, M. (2008) The histone chaperone protein Nucleosome Assembly Protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription. *Retrovirology*, 5, 8.

Varmus, H.E., Weiss, R.A, Friis, R.R., Levinson, W. and Bishop, J.M. (1972) Detection of avian tumor virus-specific nucleotide sequences in avian cell DNAs (reassociation kinetics-RNA tumor viruses-gas antigen-Rous sarcoma virus, chick cells). *Proc Natl Acad Sci USA*, 69, 20-24.

Verwoerd, D.W., Williamson, A.L. and De Villiers, E.M. (1980) Aetiology of Jaagsiekte: transmission by means of subcellular fractions and evidence for the involvement of a retrovirus. *Onderstepoort J Vet Res*, 47, 275–280.

Villarreal, L.P. (1997) On viruses, sex, and motherhood. *J Virol*, 71, 859-865.

Vogt, P.K. (1997) Historical introduction to the general properties of retroviruses. *Retroviruses*. J.M. Coffin, S.H. Hughes and H.E. Varmus. New York, Cold Spring Harbour Laboratory Press: 1-25.

Vogt, V.M. (1997) Retroviral virions and genomes. *Retroviruses*. J.M. Coffin, S.H. Hughes and H.E. Varmus. New York, Cold Spring Harbour laboratory Press: 27-69.

Voisset, C., Weiss, R.A. and Griffiths, D.J. (2008) Human RNA "rumor" viruses: the search for novel human retroviruses in chronic disease. *Microbiol Mol Biol Rev*, 72, 157-196.

von Heijne, G. (1985) Signal sequences. The limits of variation. *J Mol Biol,* 184, 99-105.

von Heijne, G. (1986) Towards a comparative anatomy of N-terminal topogenic protein sequences. *J Mol Biol*, 189, 239-242.

von Heijne, G. (1990) The signal peptide. *J Membr Biol*, 115, 195-201.

Vorobev, L.N. and Ia, V. (1963) [Movement of Nitella Flexilis Protoplasm in the Presence of Local and Spreading Potentials]. *Biofizika*, 8, 575-578.

Wang, G. and Mulligan, M.J. (1999) Comparative sequence analysis and predictions for the envelope glycoproteins of foamy viruses. *J Gen Virol*, 80, 245-254.

Wang, L., Ren, X.M., Xing, J.J. and Zheng, A.C. (2010) The nucleolus and viral infection. *Virol Sin, 2*5, 151-157.

Weaver, R.F. and Weissmann, C. (1979) Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15 S beta-globin mRNA precursor and mature 10 s beta-globin mRNA have identical map coordinates. *Nucleic Acids Res*, 7, 1175-1193.

Weinberg, J.B., Matthews, T.J., Cullen, B.R. and Malim, M.H. (1991) Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J Exp Med*, 174, 1477-1482.

Weis, K. (2003) Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell*, 112, 441-451.

Weiss, R.A. (2006) The discovery of endogenous retroviruses. *Retrovirology*, 3, 67.

Weiss, R.A. and Payne, L.N. (1971) The heritable nature of the factor in chicken cells which acts as a helper virus for Rous sarcoma virus. *Virology*, 45, 508-515.

Wick, S.M. (1991) Spatial aspects of cytokinesis in plant cells. *Curr Opin Cell Biol*, 3, 253-260.

Wiegand, H.L., Coburn, G.A., Zeng, Y., Kang, Y., Bogerd, H.P. and Cullen, B.R. (2002) Formation of Tap/NXT1 heterodimers activates Tap-dependent nuclear mRNA export by enhancing recruitment to nuclear pore complexes. *Mol Cell Biol*, 22, 245-256.

Wilhelm, J. and Pingoud, A. (2003) Real-time polymerase chain reaction. *Chembiochem*, 4, 1120-1128.

Wilk, T., Geiselhart, V., Frech, M., Fuller, S.D., Flügel, R.M. and Löchelt, M. (2001) Specific interaction of a novel foamy virus Env leader protein with the N-terminal Gag domain. *J Virol*, 75, 7995-8007.

Wright, E.M., Snopek, B. and Koopman, P. (1993) Seven new members of the Sox gene family expressed during mouse development. *Nucleic Acids Res*, 21, 744.

Wu, X., Anderson, J.L., Campbell, E.M., Joseph, A.M. and Hope, T.J. (2006) Proteasome inhibitors uncouple rhesus TRIM5alpha restriction of HIV-1 reverse transcription and infection. *Proc Natl Acad Sci USA,* 103, 7465-7470.

York, D.F., Vigne, R., Verwoerd, D.W. and Querat, G. (1991) Isolation, identification, and partial cDNA cloning of genomic RNA of Jaagsiekte retrovirus, the etiological agent of sheep pulmonary adenomatosis. *J Virol,* 65, 5061-5067.

York, D.F., Vigne, R., Verwoerd, D.W. and Querat, G. (1992) Nucleotide sequence of the Jaagsiekte retrovirus, an exogenous and endogenous type D and B retrovirus of sheep and goats. *J Virol*, 66, 4930-4939.

Yu, S.F., Eastman, S.W. and Linial, M.L. (2006) Foamy virus capsid assembly occurs at a pericentriolar region through a cytoplasmic targeting/retention signal in Gag. *Traffic*, 7; 966-977.

Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P. and Yu, X.F. (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science*, 302, 1056-1060.

Zaitseva, M., Peden, K. and Golding, H. (2003) HIV coreceptors: role of structure, posttranslational modifications, and internalization in viral-cell fusion and as targets for entry inhibitors. *Biochim Biophys Acta*, 1614, 51-61.

Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. and Charneau, P. (2000) HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*, 101, 173-185.

Zhang, F., Landford, W.N., Ng, M., McNatt, M.W., Bieniasz, P.D. and Hatziioannou, T. (2011) SIV Nef proteins recruit the AP-2 complex to antagonize Tetherin and facilitate virion release. *PLoS Pathog*, 7, e1002039.

Journal of Virology

The Signal Peptide of a Recently Integrated Endogenous Sheep Betaretrovirus Envelope Plays a Major Role in Eluding Gag-Mediated Late Restriction

Alessia Armezzani, Frédérick Arnaud, Marco Caporale, GiuliaPia di Meo, Leopoldo Iannuzzi, Claudio Murgia and Massimo Palmarini

J. Virol. **2011, 85(14):7118. DOI: 10.1128/JVI.00407-11. Published Ahead of Print 18 May 2011.**

JOURNAL OF VIROLOGY, July 2011, p. 7118–7128 Vol. 85, No. 14
0022-538X/11/$12.00 doi:10.1128/JVI.00407-11

The Signal Peptide of a Recently Integrated Endogenous Sheep Betaretrovirus Envelope Plays a Major Role in Eluding Gag-Mediated Late Restriction▽

Alessia Armezzani,[1] Frédérick Arnaud,[1,2] Marco Caporale,[1,3] Giulia Pia di Meo,[4] Leopoldo Iannuzzi,[4] Claudio Murgia,[1] and Massimo Palmarini[1]*

MRC, University of Glasgow Centre for Virus Research, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom[1]; EPHE, Université de Lyon 1, INRA, UMR754, École Nationale Vétérinaire de Lyon, IFR 128, Lyon, France[2]; Istituto G. Caporale, Teramo, Italy[3]; and ISPAAM, National Research Council (CNR), Naples, Italy[4]

Received 28 February 2011/Accepted 3 May 2011

The exogenous and pathogenic Jaagsiekte sheep retrovirus (JSRV) coexists with highly related and biologically active endogenous retroviruses (enJSRVs). The endogenous enJS56A1 locus possesses a defective Gag polyprotein which blocks the late replication steps of related exogenous and endogenous retroviruses by a mechanism known as *J*SRV *l*ate *r*estriction (JLR). Conversely, enJSRV-26, which most likely integrated into the sheep genome less than 200 years ago, is able to escape JLR. In this study, we demonstrate that the ability of enJSRV-26 to escape JLR is due to a single-amino-acid substitution in the signal peptide (SP) of its envelope glycoprotein. We show that enJSRV-26 SP does not localize to the nucleolus, unlike the functional SPs of related exogenous and endogenous sheep betaretroviruses. In addition, enJSRV-26 SP function as a posttranscriptional regulator of viral gene expression is impaired. enJSRV-26 JLR escape relies on the presence of the functional enJS56A1 SP. Moreover, we show that the ratio between enJSRV-26 and enJS56A1 Gag is critical to elude JLR. Interestingly, we found that the domestic sheep has acquired, by genome amplification, several copies of the enJS56A1 provirus. These data further reinforce the notion that transdominant enJSRV proviruses have been positively selected in domestic sheep, and that the coevolution between endogenous and exogenous sheep betaretroviruses and their host is still occurring.

Retroviruses must integrate their genome into the host genomic DNA to replicate successfully. As a consequence of their peculiar replication cycle, retroviruses exist in nature as exogenous retroviruses, transmitted horizontally from infected to uninfected host like any other virus, and endogenous retroviruses (ERVs). ERVs derive from the infection of the host germ line during evolution and are transmitted vertically from generation to generation like any other Mendelian gene (15, 18). ERVs colonize the genome of all vertebrates studied to date, where they represent a significant percentage of the DNA of their host species (e.g., ~8% of the human and mouse genomes) (15, 18).

During evolution, most ERVs have accumulated genetic defects and lost the ability to express proteins and/or infectious viruses. However, some ERVs have been coopted by their hosts because they fulfill useful functions (4, 5, 12–14, 18). In addition, some ERVs protect the host against the infection of related exogenous pathogenic retroviruses. In chickens and mice, for example, it has been shown that the expression of Env glycoproteins by some ERVs can saturate the receptors used by related exogenous retroviruses to gain entry into the cell (42).

Domestic sheep provide a fascinating model for studying the interplay between retroviruses and their host. The sheep genome harbors at least 27 copies of endogenous betaretroviruses (enJSRVs) that are highly related to the exogenous and pathogenic Jaagsiekte sheep retrovirus (JSRV) (1, 2, 4, 29, 30, 39). enJSRVs have been integrating into the genome of their host throughout the evolution of the *Caprinae* for the last 5 to 7 million years (i.e., sheep, goats, and related species) (2).

enJSRVs (or at least some of the enJSRV proviruses) can be considered to be in symbiosis with their host, as they play an essential part in the reproductive biology of sheep and interfere with the replication cycle of related exogenous retroviruses (4, 11, 12, 34, 35, 39). A transdominant provirus, enJS56A1, blocks JSRV replication by a unique mechanism that we termed *J*SRV *l*ate *r*estriction (JLR) (25). enJS56A1 possesses a defective Gag protein that does not traffic properly to the pericentriolar area, where newly formed viral particles assemble and use the recycling endosomes to exit from cells (3, 26). Interestingly, the defect of enJS56A1 is transdominant over JSRV as well as other enJSRVs. In other words, enJS56A1 Gag forms multimers with JSRV Gag, which consequently cannot traffic properly and subsequently are degraded by the proteasomal machinery of the cell (3, 26).

The main determinant of JLR is a tryptophan residue (W) at position 21 in enJS56A1 Gag, which replaces an arginine (R) that is well conserved in betaretroviruses (25). We showed that enJS56A1 possessed an arginine residue in Gag at position 21 when originally integrated into the host genome. Subsequently, the transdominant enJS56A1 with the W21 Gag residue ap-

* Corresponding author. Mailing address: MRC, University of Glasgow Centre for Virus Research, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow Q61 1QH, United Kingdom. Phone: 141-330-2541. Fax: 141-330-2271. E-mail: massimo.palmarini@glasgow.ac.uk.

▽ Published ahead of print on 18 May 2011.

peared in the closest relatives of domestic sheep and then became fixed after domestication (2).

Interestingly, we also have identified five enJSRV proviruses (enJSRV-7, enJSRV-15, enJSRV-16, enJSRV-18, and enJSRV-26) with an intact genomic organization that are able to produce viral particles *in vitro* (2). These loci are insertionally polymorphic in domestic sheep, in other words, they are present in only some individuals/breeds. These observations suggest that the original integration of these proviruses occurred after domestication (i.e., in the last 10,000 years). In particular, enJSRV-26 probably integrated into the host germ line less than 200 years ago. Remarkably, this virus possesses the unique ability to escape the restriction induced by enJS56A1 (2). enJSRV-26 has been detected in the germ line of a single sheep to date, suggesting that an enJSRV-26-like exogenous retrovirus still circulates within the sheep population.

Sheep betaretroviruses have allowed us to witness sequential counteradaptations between endogenous and exogenous retroviruses and represent an ideal model to study the arms race between virus and host over long evolutionary periods. In this study, we investigated the molecular mechanisms followed by enJSRV-26 to elude the restriction induced by enJS56A1. Using a variety of approaches, we demonstrate that a single amino acid substitution in the signal peptide of the enJSRV-26 Env confers the ability of this virus to escape JLR. We and others have shown previously that the signal peptide (SP) of sheep and mouse betaretrovirus envelope glycoproteins is a multifunctional protein acting as a post-transcriptional regulator of viral gene expression (7, 8, 17). Here, we demonstrate that the SP of enJSRV-26 lacks at least some of these functions. In addition, we show that JLR escape depends on the ratio between enJSRV-26 and enJS56A1 Gag. Interestingly, we also obtain evidence suggesting that the transdominant proviruses are amplified within the genome of domestic sheep.

MATERIALS AND METHODS

Plasmids. pCMV4JS21, pCMV5-enJS26, pCMV5-enJS18, and pCMV2en56A1 express the full-length $JSRV_{21}$ molecular clone and the endogenous enJSRV-26, enJSRV-18, and enJS56A1, respectively. These plasmids have been described previously (2, 28, 30). All of the chimeras/mutants employed in this study were derived from the plasmids listed above and are schematically represented in Fig. 1. Specific details for the cloning procedures of any of the plasmids described below are available upon request. Mutants were obtained by site-directed mutagenesis using the QuikChange kit (Stratagene), as suggested by the manufacturer. Chimeras were derived by swapping, respectively, the full-length (enJS26-Env18 and enJS18-Env26), the 5′-end (enJS26-5′Env18 and enJS18-5′Env26), or the 3′-end (enJS26-3′Env18 and enJS18-3′Env26) *env* between enJSRV-26 and enJSRV-18. Mutants enJSRV-26ΔEnv and enJSRV-18ΔEnv contain two nonsense mutations in the first and third methionines of their respective Env glycoproteins. Mutant enJS56A1ΔGag contains two nonsense mutations replacing the first and third methionines of the enJS56A1 Gag. enJS56A1-4CTE contains four copies of the constitutive transport element (CTE) of Mason-Pfizer monkey virus (M-PMV) at the 3′ end of *env*. The M-PMV CTE was derived from pSarm4, as already described, and was a gift from Eric Hunter (32). Single mutants enJSRV26-EnvD6A, enJSRV18-EnvA6D, and JSRV-EnvA6D express the full-length proviruses with a single point mutation in their Env glycoprotein at position 6. Plasmids penJSEnv26 and penJSEnv18 express the Env of enJSRV-26 and enJSRV-18, respectively. These mutants were obtained by deleting *gag*, *pro*, *pol*, and *orf-x* from the plasmids encoding their respective full-length proviruses. pSP26-HA, pSP18-HA, and pSP56-HA encode the signal peptide of enJSRV-26, enJSRV-18, and enJS56A1, respectively, tagged with the hemagglutinin (HA) epitope. In penJS18Env-PPT, the signal peptide of enJSRV-18 Env was replaced by the signal peptide of the human preprotrypsin, followed by the FLAG epitope fused at the N terminus of the enJSRV-18 surface domain (SU) of Env. pNLgagSty330 and pRev were kindly provided by Barbara Felber and have been described already (16, 21). pNLgagSty330 is a Rev- and Tat-dependent HIV-1 Gag-Pol expression plasmid and, for simplicity, is termed pHIV1-RRE in this study. pRev is an expression plasmid for HIV-1 Rev. The expression plasmid for HIV-1 Tat was a gift from Mauro Giacca and has been described elsewhere (37). pHIV-SPRE26, pHIV-SPRE18, and pHIV-SPRE56 were obtained by replacing the HIV-1 Rev-responsive element (RRE) in pHIV1-RRE with the signal peptide-responsive elements (encompassing *env* and the 3′ untranslated region [UTR]) of enJSRV-26, enJSRV-18, and enJS56A1, respectively.

Cell cultures, transfections, and viral preparations. 293T and COS cells were cultured in Dulbecco's modified Eagle medium (Gibco) and supplemented with 10% fetal bovine serum at 37°C, 5% CO_2, and 95% humidity. Virus preparations were obtained by transient transfections of 293T cells with the appropriate plasmids using the Calphos mammalian transfection kit (Clontech). The empty vector pcDNA3.1 (Invitrogen) was used to calibrate the amount of DNA used in each experiment. Cell supernatants were collected at 48 h posttransfection, and viral particles were concentrated by ultracentrifugation as already described (8, 30, 31). For the analysis of intracellular proteins, cells were lysed by standard techniques, as described previously (38).

Western blotting and immunoprecipitation. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting were performed on concentrated viral particles and cell lysates (200 μg of protein extracts) as previously described (8, 30). enJSRV Gag was detected using a rabbit polyclonal serum against the JSRV major capsid protein (CA) (26). A rabbit polyclonal serum toward the JSRV transmembrane protein (TM) was employed to detect Env proteins. Signal peptides tagged with the HA epitope were detected with a mouse monoclonal anti-HA antibody (Abcam), while γ-tubulin was detected with a rabbit polyclonal antibody (Sigma). Membranes were exposed to the appropriate peroxidase-conjugated secondary antibodies and further developed by chemiluminescence using ECL Plus (Amersham). Levels of CA associated with viral particles released in the supernatants were quantified by measuring chemiluminescence in a Molecular Dynamics Storm 840 imaging system using ImageQuant TL software (Molecular Dynamics). Each experiment was repeated independently at least three times, and results are presented as the mean values for each sample (± standard errors). Env expression by enJS18Env-PPT was assessed by immunoprecipitation using a mouse monoclonal anti-FLAG antibody (Sigma) as already described (38).

Confocal microscopy. 293T and COS cells were plated onto two-well chambered glass slides (Lab-Tek; Nalge Nunc International) and transfected with the appropriate plasmids using Lipofectamine (Invitrogen) according to the manufacturer's instructions. Twenty-four h after transfection the cells were washed with phosphate-buffered saline (PBS) and fixed with 3% formaldehyde for 15 min. After fixation, cells were processed essentially as already described (25, 33). SP proteins tagged with the HA epitope were detected with a mouse monoclonal anti-HA (Abcam) antibody. Rabbit polyclonal antibody to fibrillarin (Abcam) was used as a marker for the nucleolus. Goat anti-mouse and anti-rabbit immunoglobulin G conjugated with Alexa Fluor 488 and Alexa Fluor 594, respectively (Molecular Probes), were used as secondary antibodies. Slides were mounted with medium containing DAPI (4′,6-diamidino-2-phenylindole; Vectashield, Vector Laboratories) and analyzed with a Leica TCS SP2 confocal microscope. Single sections from confocal optical sections along the *z* axis were analyzed.

Gag ELISA. For enzyme-linked immunosorbent assay (ELISA), 293T cells were transfected with the appropriate plasmids (1 μg) in the presence or absence of SP26-HA, SP18-HA, SP56-HA (or pRev, as a control), and HIV-1 Tat (0.2 μg). At 48 h posttransfection, cell supernatants were assessed for the presence of HIV-1 Gag proteins using a Murex HIV antigen monoclonal antibody (MAb) kit (Abbot Murex) according to the manufacturer's instructions. All experiments were repeated independently at least three times.

qPCR. Quantitative PCR (qPCR) assays were designed to estimate the dosage of target genes (enJS56A1, enJSRV-6, enJSRV-18, and *SOX9*) compared to that of the β-actin gene (used as reference gene) in genomic DNA of domestic and wild sheep. The genomic DNA tested was collected from various breeds (Dorset, Suffolk, Texel, Jacob, Red Maasai, Merino, Xalda, Rambouillet, Soay, Norway, and Finsheep) of domestic sheep (*Ovis aries*), wild sheep (*O. dalli*, *O. canadensis*, *O. ammon*, and *O. vignei*), Mediterranean mouflon (*O. orientalis musimon*), and members of the genera *Budorcas* (*B. taxicolor*) and *Pseudois* (*P. nayaur*). All DNA samples were obtained and used in a previous study (9). Standard curve efficiency was 99.2% for enJS56A1, 96.6% for enJSRV-6, 99% for enJSRV-18, 100% for β-actin, and 100% for *SOX9*. PCR assays were performed using a reverse primer complementary to the genomic 3′ flanking region for each pro-

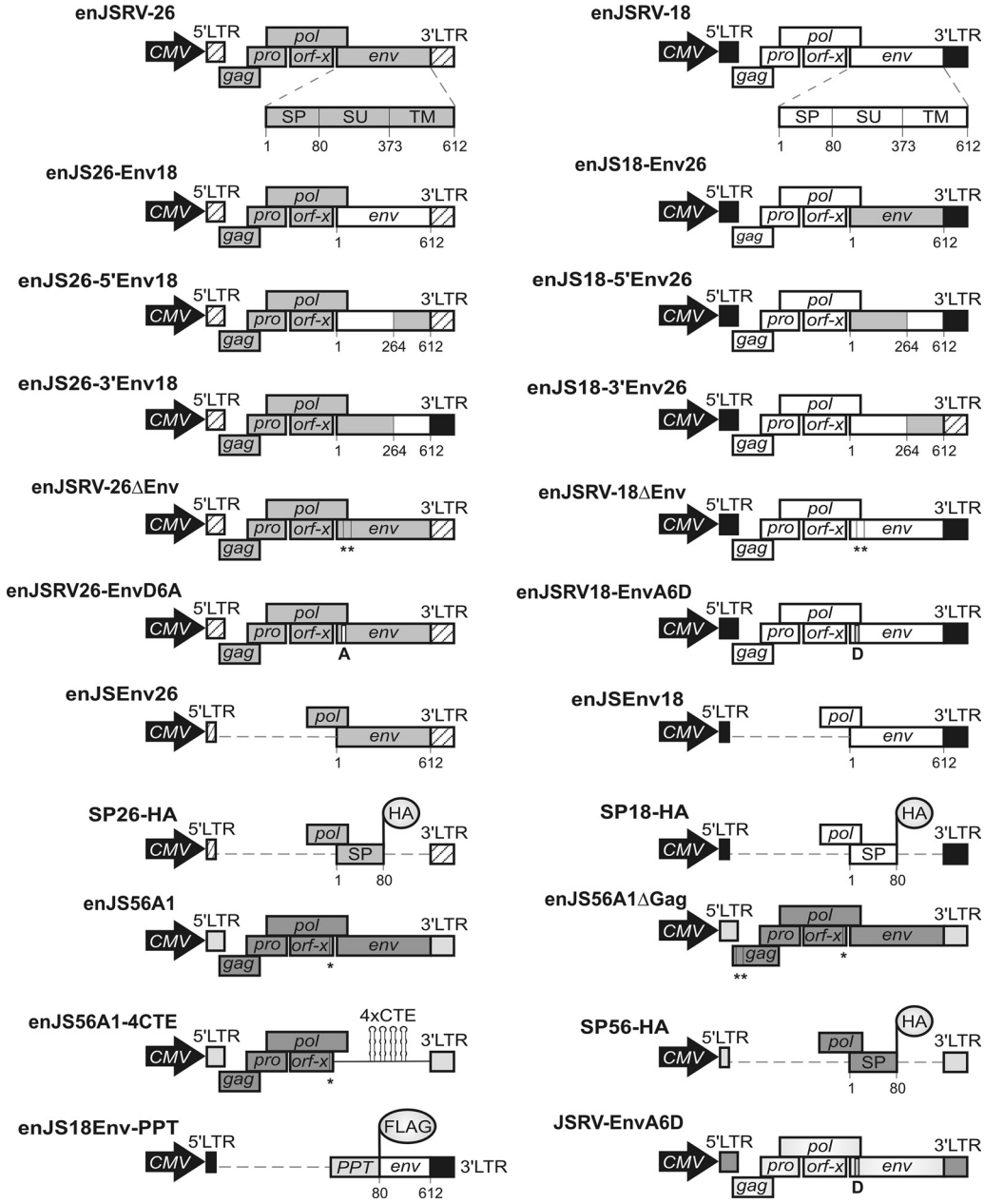


FIG. 1. Schematic representation of the plasmids employed in this study. All of the mutants/chimeras used in this study were derived from expression plasmids encoding the full-length enJSRV-26 and enJSRV-18 proviruses (top). Numbers indicate amino acid residues of Env. Premature termination codons are indicated with vertical lines and asterisks. LTR, long terminal repeat; CMV, cytomegalovirus.

virus (enJS56A1, 5′-GGA AGG ATC TGA AAC GTG GA-3′; enJSRV-6, 5′-CAG GGG AAT AAC TGG TGC TAC CT-3′; and enJSRV-18, 5′-CAA GTG CCA GAG CCC AGA GCC A-3′) and a forward primer designed in a conserved region in *env* (5′-ATA AAG AGA GGG GAG CTG CG-3′). Primers for β-actin (forward, 5′-ATC ATG TTT GAG ACC TTC AAC ACC CC-3′; reverse, 5′-CCA GGA AGG AAG GCT GGA AGA GAG C-3′) and *SOX9* (forward, 5′-CCT AGC TTT TCT TGC AGC C-3′; reverse, 5′-GCA TTC CCC AGA CAG ATT TC-3′) were designed on highly conserved regions of both genes. qPCR assays were carried out in triplicate in a total volume of 25 μl and performed in a Mx30005P (Stratagene) thermocycler, using the Brilliant II SYBR green QPCR low ROX master mix (Stratagene) and the Brilliant SYBR green QPCR core reagent kit (Stratagene) according to the manufacturer's instructions. The reaction mixture contained 20 ng of sheep genomic DNA. The reaction mixture was subjected to a denaturation step at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 30 s, primer annealing at the temperature appropriate for each primer (55°C to 60°C for 30 s), and elongation at 72°C for 30 s, ending with a melting-curve analysis to validate the specificity of the PCR products. Results were expressed as the ratio between the estimated number of molecules in the target and reference genes in each sample.

Genotyping enJS56A1 proviruses. The presence of the codon encoding an arginine or tryptophan residue in enJS56A1 (and enJS56A1-like proviruses) at position 21 in Gag was assessed by PCR. Genomic DNA samples collected from various breeds of domestic sheep (Texel, $n = 2$; Merino, $n = 1$), wild sheep (*O. dalli*, $n = 1$; *O. canadensis*, $n = 1$; *O. ammon*, $n = 1$), and the Mediterranean Mouflon (*Ovis orientalis musimon*, $n = 1$) were amplified by PCR using a forward primer complementary to the genomic 5′-flanking region of the enJS56A1 or enJSRV-20 provirus and a reverse primer complementary to their *gag* gene as previously described (2). PCR products then were cloned into a pCR4-TOPO

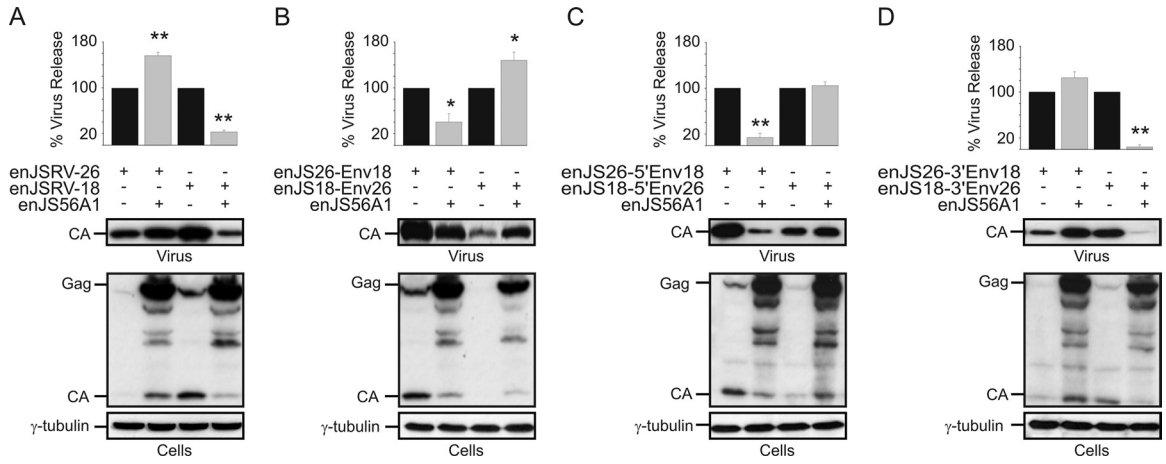


FIG. 2. enJSRV-26 *env* 5′ end contains the determinants necessary to escape JLR. Shown are Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the major capsid protein of JSRV (CA) or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by each chimera expressed in the absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P < 0.05$) or two ($P < 0.01$) asterisks. (A) Interference assays with wild-type enJSRV-26 and enJSRV-18. (B, C, and D) Interference assays using various enJSRV-26/enJSRV-18 chimeras as indicated in each panel. Only those chimeras containing the 5′ end of the *env* gene of enJSRV-26 are able to elude enJS56A1 restriction.

vector (Invitrogen), and 40 individual clones for each PCR product were completely sequenced.

FISH. Fluorescent *in situ* hybridization (FISH) analysis was carried out essentially as already described (9). Briefly, sheep peripheral blood cells were cultured at 37°C in RPMI 1640 medium (Invitrogen) supplemented with 15% fetal bovine serum, 1.5% concanavalin A (Sigma), and penicillin-streptomycin (Invitrogen). Cells then were synchronized with 300 μg/ml thymidine (Sigma) and, after 18 h, washed and resuspended in medium containing 15 μg/ml bromodeoxyuridine (Sigma) and 30 μg/ml Hoechst 33258 (Invitrogen). Cells then were incubated for 6 h at 37°C (the last hour in the presence of 0.1 μg/ml colcemid; Sigma) and then treated with a hypotonic solution and washed three times with methanol-acetic anhydride. Cell suspensions then were plated onto slides, incubated overnight at 50°C, and stained for 10 min with 25 μg/ml Hoechst 33258. Slides were further probed with biotin-labeled bacterial artificial chromosome (BAC) clones containing the appropriate provirus. Hybridization, chromosome staining, signal detection, and image processing were performed as already described (9) in at least 30 metaphases for each probe. Chromosome identification was carried out using the R-banding karyotype by adding fluorescein avidin DCS and biotinylated anti-avidin (Vectors Laboratories), as recommended by the manufacturer. Chromosome identification and band nomenclature followed the International System for Chromosome Nomenclature of Domestic Bovids (10).

RESULTS

The 5′ portion of the enJSRV-26 *env* is the main determinant of JLR escape. The first aim of this study was to identify the molecular determinants of enJSRV-26 that are necessary to escape JLR. At the nucleotide level, enJSRV-26 is 98% identical to enJSRV-18 along the entire genome. enJSRV-18 is another insertionally polymorphic provirus present in the sheep genome. Interestingly, enJSRV-26 and enJSRV-18 Gag and Env are 100 and 99.3% identical, respectively. However, enJSRV-26 and enJSRV-18 display different phenotypes in the presence of enJS56A1. enJSRV-18 is restricted by the transdominant enJS56A1 (like the exogenous JSRV), whereas enJSRV-26 escapes JLR (2) (Fig. 2A). Because the major differences between enJSRV-26 and enJSRV-18 are found in *env*, we reasoned that this region contains the main determinants for JLR escape. The enJSRV-26 and enJSRV-18 *env* genes differ for only 10 nucleotides, resulting in six synonymous and four nonsynonymous mutations.

To identify the determinants of JLR escape, we generated a series of chimeras between enJSRV-18 and enJSRV-26 (Fig. 1) and carried out interference assays with the transdominant enJS56A1 (Fig. 2). As shown in Fig. 2B, the release of enJS26-Env18 viral particles was restricted in the presence of enJS56A1, while enJS18-Env26 escaped JLR. Furthermore, chimeras enJS18-5′Env26 and enJS26-3′Env18 also were able to escape JLR, while enJS56A1 was able to inhibit enJS26-5′Env18 and enJS18-3′Env26 viral particle release (Fig. 2C and D). Collectively, these results indicated that the 5′ end of enJSRV-26 *env* contains the main determinants of JLR escape.

Amino acid residue D6 in the enJSRV-26 Env is the main determinant of JLR escape. We investigated whether the enJSRV-26 Env protein *per se* or *cis*-acting regions within the *env* gene were involved in JLR escape. We derived enJSRV-26 and enJSRV-18 mutants containing two premature termination codons in Env (enJSRV-26ΔEnv and enJSRV-18ΔEnv) to prevent its expression while maintaining the intact full-length viral genome. As expected, we found that enJSRV-26ΔEnv was able to escape JLR when the enJSRV-26 Env was provided in *trans*, whereas it was restricted in the presence of the enJSRV-18 Env (Fig. 3A). On the other hand, enJSRV18ΔEnv was restricted by enJS56A1 in the presence of the enJSRV-18 Env but escaped JLR when coexpressed with the enJSRV-26 Env (Fig. 3B). These data confirmed that the enJSRV-26 Env protein *per se* is necessary to escape JLR.

The N-terminal region of the retroviral Env includes the signal peptide. The SPs of enJSRV-26 and enJSRV-18 Env differ in only three amino acid residues (Fig. 3C). However, the SP of enJSRV-7 and enJSRV-15 (two enJSRV loci that, like enJSRV-18, are restricted by enJS56A1) differ by only a single

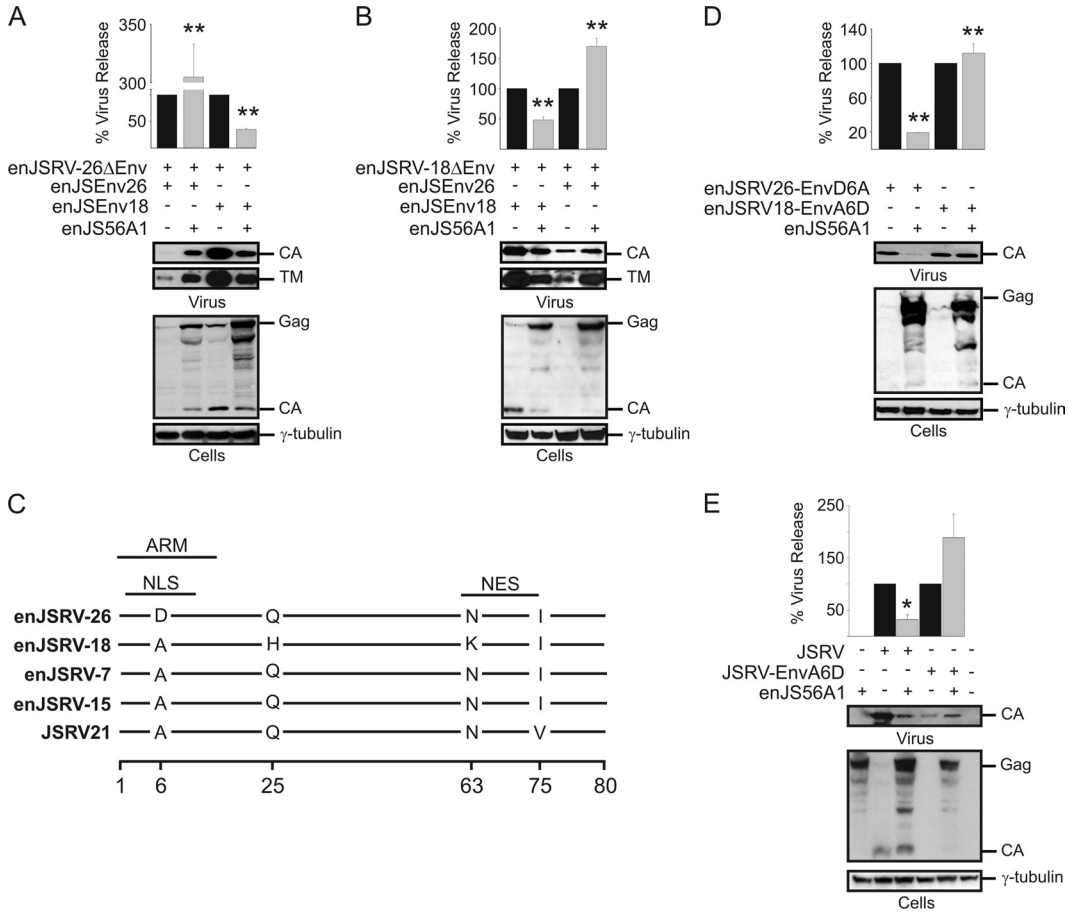


FIG. 3. Amino acid residue D6 in the enJSRV-26 Env plays a major role in JLR escape. (A, B, D, and E) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the major capsid protein (CA) of JSRV or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by each mutant expressed in the absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P < 0.05$) or two ($P < 0.01$) asterisks. Env expression was controlled by incubating membranes with antibodies against the transmembrane (TM) domain of JSRV. (C) Graphic representation of the alignment of the amino acid sequences of the signal peptides of four insertionally polymorphic enJSRV proviruses and the exogenous JSRV. Lines represent identical residues in the sequences, while letters indicate differences in the amino acid residues. Numbering corresponds to amino acid residues in Env. ARM, arginine-rich motif; NLS, nuclear localization signal; NES, nuclear export signal.

amino acid residue (residue 6) from the SP of enJSRV-26 (2). The alanine (A) residue in position 6 is well conserved in the SP of JSRV and all the insertionally polymorphic enJSRVs, with the exception of enJSRV-26, where it is replaced by an aspartic acid (D) residue (Fig. 3C). Thus, we hypothesized that the D6 residue plays a critical role in the ability of enJSRV-26 to escape JLR. To this end, we derived the full-length single mutants enJSRV26-EnvD6A and enJSRV18-EnvA6D. As shown in Fig. 3D, the D6A mutation conferred susceptibility to JLR to enJSRV-26, while the reciprocal mutation (A6D) allowed enJSRV-18 to escape enJS56A1-induced restriction. Similar results were obtained for the exogenous JSRV, where the A6D mutation allowed JSRV to escape JLR (Fig. 3E). These data show conclusively that a single amino acid substitution in the SP of the Env glycoprotein allows enJSRV-26 to escape JLR.

SPs of the enJSRV-26 and enJSRV-18 Env localize in different cellular compartments. We investigated whether the SP of the enJSRV-26 Env possessed a different biological activity from that of SPs of enJSRV-18 and enJS56A1. We and others have shown that the SPs of sheep betaretroviruses are Rev-like multifunctional proteins that localize in the nucleoli and favor full-length viral RNA nuclear export and enhance Gag synthesis and viral particle release (8, 17). We have shown above that the A6D substitution allows enJSRV-26 to escape JLR. This residue lays within a predicted nuclear localization signal (NLS) and an arginine-rich RNA binding motif (ARM) of the SP (Fig. 3C). Thus, we investigated whether the A6D substitution affected the intracellular localization of the enJSRV-26 SP.

By confocal microscopy, we observed that both the enJSRV-18 and enJSRV-26 SPs localize in the cytoplasm and in the nucleus of transfected cells (Fig. 4). The enJSRV-18 SP colocalized with nucleolar markers, such as fibrillarin, as we previously observed for the JSRV SP (8). On the other hand, the enJSRV-26 SP displayed a diffuse nuclear staining pattern

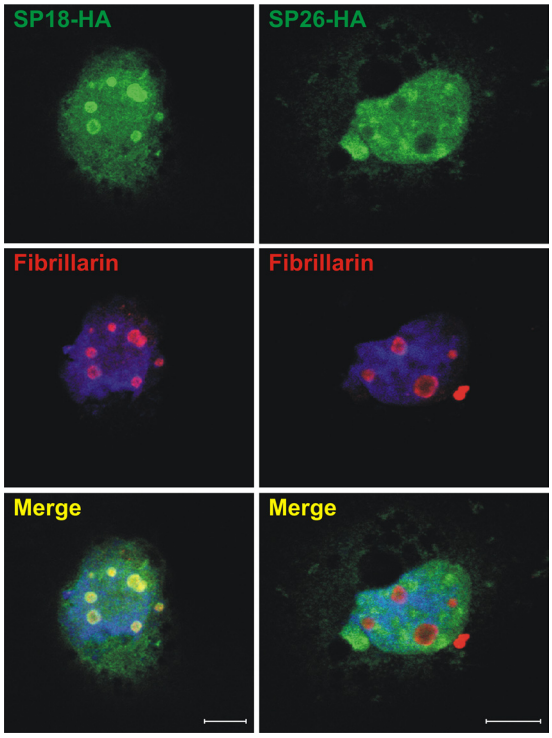


FIG. 4. Signal peptide (SP) of enJSRV-26 does not localize in the nucleoli. The intracellular localization of the signal peptide of enJSRV-18 and enJSRV-26 is shown. COS cells were transfected with expression plasmids for the SPs of enJSRV-18 or enJSRV-26 (tagged with the HA epitope), fixed 24 h posttransfection, and incubated with anti-HA (top) and fibrillarin (middle) antibodies. Nuclei are shown in blue. Both SPs display nuclear localization, but only the enJSRV-18 SP shows a strong colocalization with nucleolar markers such as fibrillarin. Bars correspond to 10 μm.

with no accumulation in the nucleoli. The relative number of cells expressing the enJSRV-18 SP with nucleolar localization was about 80-fold higher than those expressing the enJSRV-26 SP (data not shown). We also determined that the enJSRV-26 SP did not relocalize in the nucleoli in the presence of enJS56A1, indicating that its defect was not rescued by the functional SP of enJS56A1 (data not shown). We performed these assays in both COS and 293T cells, obtaining essentially the same results (data not shown).

The phenotype of enJSRV-26 can be attributed to a relative lack of function of its SP. We investigated whether the altered localization of the enJSRV-26 SP was correlated with its altered function and JLR escape. We cotransfected 293T cells with expression plasmids for the enJSRV-26ΔEnv mutant and the SPs of enJSRV-26 (pSP26-HA) or enJSRV-18 (pSP18-HA), and we used Western blotting to analyze the amount of virus produced in the presence or absence of enJS56A1 (Fig. 5A). We found that enJSRV-26ΔEnv was able to escape JLR if the enJSRV-26 SP was provided in *trans* (Fig. 5A, lanes 6 and 7), while it was impaired in the presence of the enJSRV-18 SP (Fig. 5A, lanes 8 and 9). enJSRV-26ΔEnv also was able to escape enJS56A1 restriction when expressed by itself (Fig. 5A, lanes 4 and 5). Accordingly, we also noticed that the release of viral particles from enJSRV26ΔEnv was much more strongly enhanced by the enJSRV-18 SP than by the enJSRV-26 SP (Fig. 5A, lanes 4, 6, and 8). Similar results were obtained with the enJSRV-18ΔEnv mutant (data not shown).

Note that in a previous study we showed that enJS56A1 also expresses a defective Env glycoprotein (2). Betaretroviruses assemble in the pericentriolar area, and their Env facilitate intracellular Gag trafficking and viral particle release (3, 33). Thus, interference assays involving enJSRV-26ΔEnv can be more difficult to interpret in the absence of functional Env. However, the data described above overall suggest that the A6D mutation in the enJSRV-26 Env impairs at least some of the functions played by its SP, resulting in JLR escape. To test this hypothesis, we used a Rev-RRE-dependent HIV-1 Gag-Pol expression vector (pNLgagSty330; termed pHIV1-RRE in this study) (16, 21) and derived constructs where the HIV-1 RRE was replaced with the signal peptide responsive elements (SPREs) of enJS56A1 (pHIV-SPRE56), enJSRV-18 (pHIV-SPRE18), and enJSRV-26 (pHIV-SPRE26). These constructs where cotransfected with either pSP18-HA, pSP26-HA, or pSP56-HA (encoding the enJSRV-18, enJSRV-26, and enJS56A1 SPs, respectively), and the release of HIV particles in the supernatants was measured by ELISA (Fig. 5B). We observed that in the presence or absence of the enJSRV-26 SP, the levels of Gag in the supernatants of cells transfected with pHIV-SPRE56, pHIV-SPRE18, or pHIV-SPRE26 did not change, while they increased substantially in the presence of the SP of either enJSRV-18 or enJS56A1 (Fig. 5B).

Collectively, the data presented so far suggested that the ability of enJSRV-26 to escape JLR was due to the relative lack of function of its SP. To experimentally prove this point, we derived a mutant of the enJSRV-18 Env expression plasmid in which the SP was replaced with the heterologous SP of the human preprotrypsin protein (penJS18Env-PPT). Interestingly, we found that enJSRV-18ΔEnv escaped JLR when enJS18Env-PPT was provided in *trans* (Fig. 5C). These data suggest that JLR escape is due to a lack of function, rather than a gain of function, of the SP of the enJSRV-26 Env.

enJS56A1 Env plays a key role in enJSRV-26 JLR escape. The data obtained so far suggested that the SP of the transdominant enJS56A1 is able to enhance Gag expression as well as enJSRV-26 viral particle release. Thus, our next aim was to assess whether the expression of the enJS56A1 Env SP was a requirement for JLR escape. To this end, we generated an enJS56A1 mutant (enJS56A1-4CTE) in which the Env glycoprotein (including its SP) was replaced by four repeats of the M-PMV CTE. Gag expression of enJS56A1-4CTE therefore is SP independent, but it relies on the M-PMV CTE, which functions in *cis* (6). As expected, we found that enJSRV-26 viral particle release was blocked by enJS56A1-4CTE, indicating that the expression of the enJS56A1 Env is necessary for enJSRV-26 to escape JLR (Fig. 6A, lanes 1, 6, and 7).

Taken together, the results obtained suggested that the enJS56A1 and enJSRV-26 SPRE compete for the only functional SP within the cell (i.e., the enJS56A1 SP). Consequently, we hypothesized that by overexpressing enJS56A1, the intracellular levels of enJS56A1 SP would favor the synthesis of the transdominant Gag to levels sufficient to block enJSRV-26. To test this point, we cotransfected 293T cells with different ratios of the enJSRV-26 and enJS56A1 expression plasmids and assessed viral particle release by Western blotting (Fig. 6B). We found that enJSRV-26 was able to escape JLR when the ratio

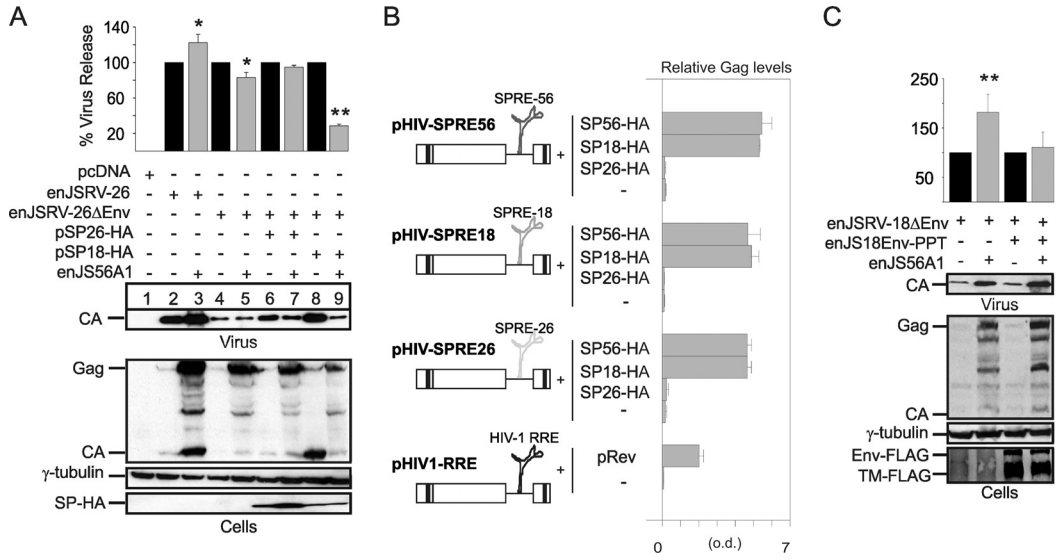


FIG. 5. Signal peptide of enJSRV-26 does not function as a posttranscriptional regulator of viral gene expression and is essential to escape JLR. (A and C) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA), the HA or FLAG epitopes (to detect SPs or Env, respectively), or γ-tubulin as a loading control. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by enJSRV-26 or related mutants in the absence of enJS56A1 (black bars) were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P < 0.05$) or two ($P < 0.01$) asterisks. (B) HIV Gag ELISAs were performed on supernatants of 293T cells transfected with HIV-1 Gag-Pol expression plasmids (pHIV-SPRE26, pHIV-SPRE18, and pHIV-SPRE56) in the presence or absence of expression plasmids for the enJSRV-26, enJSRV-18, and enJS56A1 SPs (SP26-HA, SP18-HA, and SP56-HA). Controls included supernatants of cells transfected with the HIV-1 Gag-Pol expression plasmid containing HIV-1 RRE (pHIV1-RRE) in the presence or absence of the HIV-1 Rev (pRev).

of the transfected expression plasmids was 1 to 1 (Fig. 6B, lane 4), while it was blocked when the ratio was 1 to 2 (Fig. 6B, lane 5), suggesting that enJS56A1 can inherently block enJSRV-26 viral particle release. To confirm these data, we performed cotransfection assays with expression plasmids for enJSRV-26 and either wild-type enJS56A1 or an enJS56A1 mutant deleted of *gag* (enJS56A1ΔGag) (Fig. 6C). As expected, we found that the levels of enJSRV-26 viral particles were higher in the

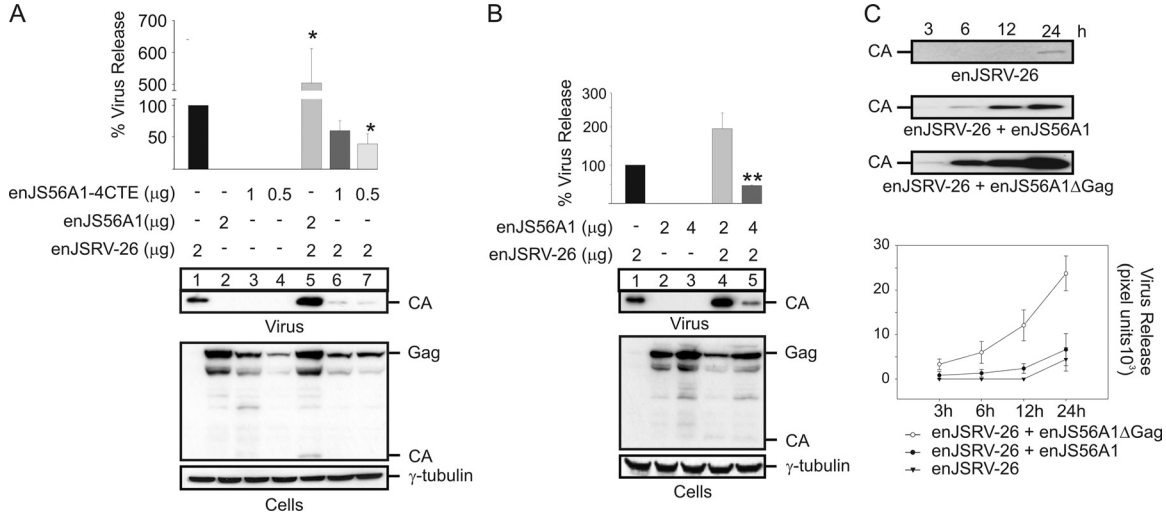


FIG. 6. enJS56A1 Env expression is critical to escape JLR. (A and B) Western blots of concentrated supernatants (virus) and cell extracts (cells) of 293T cells transfected with the plasmids indicated in each panel. Membranes were incubated with antibodies against the JSRV major capsid protein (CA) and γ-tubulin as a loading control. Note that enJS56A1-4CTE lacks the viral Env and is able to block enJSRV-26. Levels of CA associated with viral particles released in the supernatants were quantified by chemifluorescence using ImageQuant TL software (Molecular Dynamics). Graphs represent data obtained from three independent experiments. The values obtained by enJSRV-26 were arbitrarily set as 100%. Error bars indicate standard errors; statistically significant differences are indicated with one ($P < 0.05$) or two ($P < 0.01$) asterisks. (C) Western blot analysis of concentrated supernatants of 293T cells transfected with the indicated plasmids at 3, 6, 12, and 24 h posttransfection. The bottom panel represents the quantification of blots by chemifluorescence as described for panels A and B. Values are expressed as arbitrary pixel units derived from three independent experiments.

presence of enJS56A1ΔGag than with the full-length enJS56A1. These data suggest that enJS56A1 is intrinsically able to interfere with enJSRV-26, and that JLR is directly related to the relative ratio between transdominant and functional Gag. By using interference assays, we also established that the enJSRV-26 SP does not possess any dominant-negative role and does not affect, by itself, enJS56A1 Gag synthesis (data not shown).

enJS56A1 is amplified within the genome of domestic sheep. It is difficult to correlate the data obtained in *in vitro* experiments with events that in nature lead to the selection of transdominant proviruses and viruses escaping JLR. Our results suggest that the relative ratio between defective transdominant and functional Gag could determine the efficiency of JLR. The sheep genome has an overwhelming majority of enJSRV loci with functional Gag. However, in a previous study, we noticed that the BAC clones containing enJS56A1 were overrepresented in the sheep genomic BAC library used to clone the known enJSRV loci (2). Indeed, 22% of BAC positive for enJSRV sequences contained the enJS56A1 provirus, whereas the expected frequency was 3.7% (2). The overrepresentation of particular clones could be due to artifacts related to the construction, and the screening of the library but also could be due to the amplification of the genomic region containing enJS56A1. Here, using locus-specific primers, we determined the relative gene dosage of the transdominant enJS56A1 in genomic DNA of wild sheep within the genus *Ovis* (*O. dalli*, $n = 1$; *O. canadensis*, $n = 2$; *O. ammon*, $n = 4$; *O. vignei*, $n = 1$), the Mediterranean Mouflon (*Ovis orientalis musimon*, $n = 4$), and different breeds of domestic sheep (Dorset, $n = 1$; Suffolk, $n = 1$; Texel, $n = 10$; Jacob, $n = 2$; Red Maasai, $n = 2$; Merino, $n = 3$; Xalda, $n = 2$; Rambouillet, $n = 1$; Soay, $n = 3$; Norway, $n = 2$; Finsheep, $n = 3$). DNA from animals within the genera *Budorcas* (*B. taxicolor*, $n = 2$) and *Pseudois* (*P. nayaur*, $n = 1$) was used as additional negative controls. In addition, we estimated the relative gene dosage of enJSRV-6 (fixed in the *Ovis* genus), the insertionally polymorphic enJSRV-18 (present in most but not all domestic sheep), and the ovine *SOX9* gene as additional controls. We expressed the data as the ratio between the estimated number of molecules of target genes and the β-*Actin* gene, which was used as a reference gene (Fig. 7A).

As expected, the dosage of the *SOX9* gene did not change significantly across domestic and wild sheep. The modern enJSRV-18 provirus also showed low variations between the samples tested, whereas enJSRV-6 (which integrated before the divergence of the genera *Ovis* and *Capra*) displayed modest variations in some wild and domestic sheep (Fig. 7A). In contrast, the transdominant enJS56A1 revealed major differences between the samples analyzed, with clear indications of genomic amplification in some domestic sheep breeds (Fig. 7A). FISH analysis on metaphase chromosomes derived from domestic sheep (Fig. 7B) showed that at least one of the copies of the transdominant enJS56A1 provirus (enJSRV-20) maps exactly to the same chromosomal location as enJS56A1 in chromosome 6 (6q13).

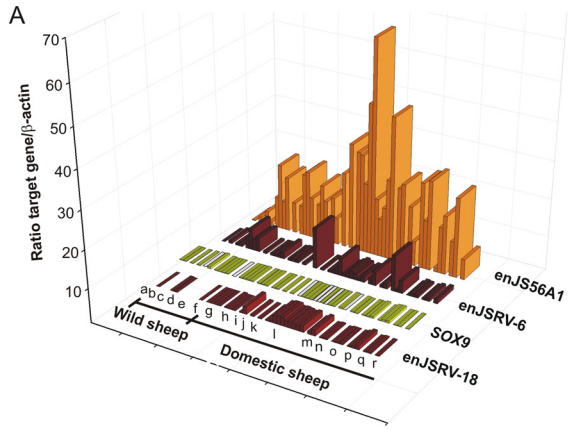
In a previous study, we speculated that enJSRV-20 arose by a process of recombination/gene conversion with enJS56A1 (2). enJSRV-20 and enJS56A1 can be distinguished by minor nucleotide sequence differences (23 nucleotides along the entire genome) and by the 5′ flanking region (the *env* sequences of an enJSRV provirus for enJSRV-20), but they share identical 3′-flanking regions. However, this study suggests that genome amplification has driven the amplification of enJS56A1-like proviruses with identical genomic flanking regions in any given animal.

Previously, we also showed that the W21 amino acid residue in Gag of enJS56A1 (and enJSRV-20) became fixed during sheep domestication. Considering that the present study showed the presence of multiple copies of enJS56A1, we sought to determine the relative frequency of the codon for the Gag R/W21 residue in the proviruses harbored in the genome of representative wild and domestic sheep. To this end, we amplified the 5′ *gag* region of enJS56A1-like proviruses (including enJSRV-20) from genomic DNA of wild sheep (*O. dalli*, $n = 1$; *O. canadensis*, $n = 1$; *O. ammon*, $n = 1$), the Mediterranean Mouflon (*O. orientalis musimon*, $n = 1$), and two different breeds of domestic sheep (Texel, $n = 2$; Merino, $n = 1$). We then cloned the PCR products obtained and sequenced at least 40 individual clones for each sample. We detected the codon corresponding to the Gag R21 residue in 100% of the PCR clones derived from the amplification of DNA collected from *O. dalli* and *O. canadensis*, which are the species phylogenetically more distant from the domestic sheep among those analyzed in this study, and for which we had no evidence of genomic amplification (Fig. 7C) (2). In species phylogenetically closer to domestic sheep (*O. ammon* and *O. orientalis*), 73 to 90% of the PCR clones sequenced contained the codon corresponding to W21 in Gag. On the other hand, 100% of the clones amplified from domestic sheep (Texel and Merino breeds) contained the codon corresponding to W21 in Gag. These data confirm that enJS56A1-like transdominant proviruses became fixed in the host genome around sheep domestication.

DISCUSSION

In this study, we revealed the molecular mechanisms underlying the ability of the recently integrated enJSRV-26 provirus to elude the late restriction induced by the defective and transdominant enJS56A1. We demonstrated that a single point mutation in the SP of the enJSRV-26 Env allows this virus to escape enJS56A1. Signal peptides mediate the targeting and translocation of membrane and secretory proteins to the endoplasmic reticulum (20, 41). Generally, they are 15 to 25 amino acid residues long and contain the cleavage site for the cellular signal peptidase which, in turn, releases the signal peptide from the rest of the protein (40). Signal peptides usually are degraded by the signal peptide peptidase. However, in some cases signal peptide sequences display other important biological functions. For instance, it has been shown that the mouse mammary tumor virus (MMTV) possesses an SP that is identical for both the Rem regulatory protein (another HIV Rev-like protein) and the envelope glycoprotein (7, 22–24). The MMTV SP requires processing by the cellular signal peptidase and retrotranslocation for nuclear function (7).

The signal peptides of JSRV and related enJSRVs are unusually long (80 amino acid residues) compared to those of



C

		enJS56A1	enJSRV-20
Ovis dalli (c)		100% R	-
Ovis canadensis (d)		100% R	100% R
Ovis ammon (e)		90% W; 10% R	-
Ovis orientalis (g)		73% W; 27% R	100% W
	(l)	100% W	100% W
Ovis aries	(l)	100% W	100% W
	(v)	100% W	100% W

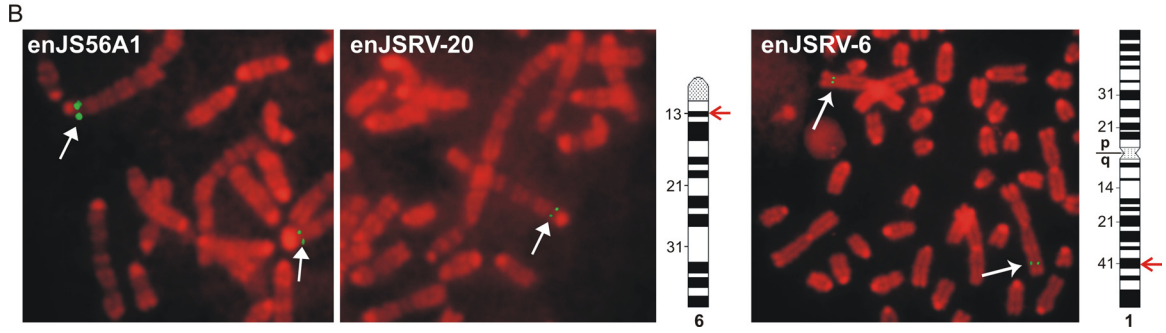


FIG. 7. enJS56A1 is amplified in the domestic sheep genome. (A) Graph representing qPCR used to estimate the gene dosage, in wild and domestic sheep, of the transdominant proviruses enJSRV-6, enJSRV-18, and *SOX9*. Each bar represents an individual animal, and each letter represents a different species/breed. Gray boxes indicate assays that could not be performed due to the exhaustion of sample DNA. The absence of a bar indicates no amplification of the corresponding enJSRV locus. Samples tested included genomic DNA collected from *B. taxicolor* (a), *P. nayaur* (b), *O. dalli* (c), *O. canadensis* (d), *O. ammon* (e), *O. vignei* (f), *O. orientalis* (g), and various breeds of domestic sheep (*O. aries*), such as Soay (h), Norway (i), Dorset (j), Suffolk (k), Texel (l), Jacob (m), Red Maasai (n), Finsheep (o), Merino (p), Xalda (q), and Rambouillet (r). (B) Fluorescent *in situ* hybridization of metaphase R-banded chromosomes derived from a Merino sheep (mixed breed). Fluorescent probes were derived from BAC clones containing the enJSRV-20, enJS56A1, or enJSRV-6 proviruses as described in Materials and Methods. The green fluorescent signals, indicated by arrows, are specific for the two transdominant proviruses (both located on chromosome 6 at band 6q13) and the enJSRV-6 locus (situated on chromosome 1 at band 1q41). Ideograms of *Ovis aries* chromosomes with R-banding patterns also are shown. (C) Relative frequency of the wild-type arginine (R) or the transdominant tryptophan (W) Gag residue at position 21 of the enJS56A1-like proviruses. The 5′ region of *gag* of enJS56A1 and enJSRV-20 was amplified by PCR from genomic DNA collected from the species indicated in the panel. Note that letters in parentheses (c, d, e, etc.) refer to the code used for panel A. PCR products were cloned into the pCR4-TOPO vector (Invitrogen), and 40 individual clones for each PCR product were sequenced to determine the relative presence of an arginine or trypthophan residue at position 21 in Gag.

other retroviruses. We and others have shown that the JSRV SP is a multifunctional protein that favors full-length viral RNA nuclear export and enhances Gag synthesis and viral particle release (8, 17). Conversely, in this study, we demonstrated that enJSRV-26 SP function is impaired due to a single point mutation. Interestingly, the A6D substitution lies within two predicted regions of the enJSRV-26 SP, the NLS and the ARM domains, which are important for SP intracellular localization and function (8, 17). Our results showed that the enJSRV-26 SP is excluded from the nucleolus, though it still can enter the nucleus, where it may passively diffuse given the small size of this protein. Interestingly, the staining pattern of the enJSRV-26 SP observed by confocal microscopy resembled what we previously obtained with a JSRV SP mutant deleted of its NLS (8). The HIV-1 Rev is targeted to the nucleolus, and this localization is crucial for mRNA trafficking and viral replication (19). Indeed, our data have shown that, unlike JSRV/enJSRV SPs, the enJSRV-26 SP function is impaired, and this defect allows this virus to elude enJS56A1 restriction. Our study supports a model in which enJS56A1 and enJSRV-26 SPREs compete for the only functional SP (i.e., the enJS56A1 SP), resulting in an increased synthesis of enJSRV-26 and reduced levels of enJS56A1 Gag proteins (Fig. 8). Using a Rev-RRE-dependent HIV-1 Gag-Pol vector, we have shown data suggesting that SPREs of enJS56A1 and enJSRV-26 respond equally well to the SP of enJS56A1. Indeed, we demonstrated that JLR escape depends on the stoichiometry between enJS56A1 and enJSRV-26 Gag, which in turn is regulated by the SPs of these proviruses.

By studying the evolutionary history of the enJSRV proviruses, we found that enJS56A1 possessed the wild-type R residue at position 21 in Gag when it first entered the host genome. Only subsequently did the transdominant enJS56A1 genotype harboring W21 appear in the host genome, and it became fixed around the time of sheep domestication. With domestication, a relatively large number of animals suddenly

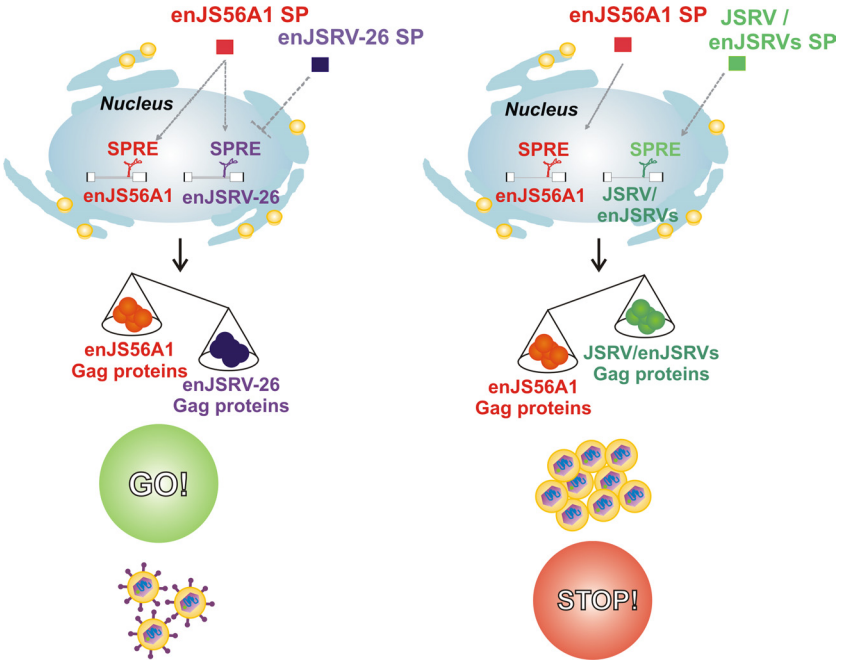


FIG. 8. Model of enJSRV-26 JLR escape. The ability of enJSRV-26 to elude JLR restriction is dependent on the impaired function of its SP. Consequently, the signal peptide responsive element (SPRE) of enJSRV-26 competes with the SPRE of enJS56A1 for the functional SP of the latter, resulting in the reduced expression of the transdominant Gag.

were kept in restricted spaces, and this likely facilitated the spread of infectious agents more easily than before. Under these circumstances, it is feasible to hypothesize that sheep with transdominant proviruses had a selective advantage.

In our previous study, we speculated that a second transdominant provirus, enJSRV-20, arose by processes of recombination and/or gene conversion with enJS56A1. The results obtained in this study suggest that the chromosomal location containing enJS56A1 has been amplified several times, especially in some breeds of domestic sheep. Thus, sheep domestication has contributed to the selection and amplification of transdominant proviruses.

The interplay between host and pathogen is a dynamic process. The host has evolved sophisticated mechanisms to block infection by pathogens, which in turn have developed countermeasures to escape host defenses. The endogenous betaretroviruses of sheep represent a unique model to study virus-host coevolution during long evolutionary periods. Sheep, like koalas, harbor several copies of intact and insertionally polymorphic endogenous proviruses (2, 27, 36). The presence of enJSRV-26, in particular, a provirus that we estimated integrated in its host within the last 200 years and escapes restriction by enJS56A1, suggests that betaretroviruses still are invading the sheep genome. The presence of multiple copies of transdominant enJSRV proviruses in modern sheep breeds therefore may be another mechanism adopted by the host to counteract retrovirus infection.

ACKNOWLEDGMENTS

This study was funded by the Wellcome Trust and in part by NIH grant HD052745, by the BBSRC, and by a Strategic Research Developmental Grant by the Scottish Funding Council.

We thank Pietro Parma (Università di Milano), Andrew Shaw, and members of our laboratories for useful suggestions. We also thank Barbara Felber, Eric Hunter, Mauro Giacca, and Jim DeMartini for providing useful reagents.

REFERENCES

1. **Arnaud, F., et al.** 2010. Interplay between ovine bone marrow stromal cell antigen 2/tetherin and endogenous retroviruses. J. Virol. **84:**4415–4425.
2. **Arnaud, F., et al.** 2007. A paradigm for virus-host coevolution: sequential counter-adaptations between endogenous and exogenous retroviruses. PLoS Pathog. **3:**e170.
3. **Arnaud, F., P. R. Murcia, and M. Palmarini.** 2007. Mechanisms of late restriction induced by an endogenous retrovirus. J. Virol. **81:**11441–11451.
4. **Arnaud, F., M. Varela, T. E. Spencer, and M. Palmarini.** 2008. Coevolution of endogenous betaretroviruses of sheep and their host. Cell Mol. Life Sci. **65:**3422–3432.
5. **Blond, J. L., et al.** 2000. An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J. Virol. **74:**3321–3329.
6. **Bray, M., et al.** 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. Proc. Natl. Acad. Sci. U. S. A. **91:**1256–1260.
7. **Byun, H., et al.** 2010. Retroviral Rem protein requires processing by signal peptidase and retrotranslocation for nuclear function. Proc. Natl. Acad. Sci. U. S. A. **107:**12287–12292.
8. **Caporale, M., et al.** 2009. The signal peptide of a simple retrovirus envelope functions as a posttranscriptional regulator of viral gene expression. J. Virol. **83:**4591–4604.
9. **Chessa, B., et al.** 2009. Revealing the history of sheep domestication using retrovirus integrations. Science **324:**532–536.
10. **Cribiu, E. P.** 2001. International system for chromosome nomenclature of domestic bovids. Cytogenet. Cell Genet. **92:**283–299.
11. **Dunlap, K. A., M. Palmarini, and T. E. Spencer.** 2006. Ovine endogenous betaretroviruses (enJSRVs) and placental morphogenesis. Placenta (Suppl. A)**:**S135–S140.
12. **Dunlap, K. A., et al.** 2006. Endogenous retroviruses regulate peri-implantation conceptus growth and differentiation. Proc. Natl. Acad. Sci. U. S. A. **103:**14390–14395.
13. **Dupressoir, A., et al.** 2005. Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. Proc. Natl. Acad. Sci. U. S. A. **102:**725–730.
14. **Dupressoir, A., et al.** 2009. Syncytin-A knockout mice demonstrate the crit-

ical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. Proc. Natl. Acad. Sci. U. S. A. **106:**12127–12132.
15. **Eiden, M. V.** 2008. Endogenous retroviruses–aiding and abetting genomic plasticity. Cell Mol. Life Sci. **65:**3325–3328.
16. **Felber, B. K., M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland, and G. N. Pavlakis.** 1989. Rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. Proc. Natl. Acad. Sci. U. S. A. **86:**1495–1499.
17. **Hofacre, A., T. Nitta, and H. Fan.** 2009. Jaagsiekte sheep retrovirus encodes a regulatory factor, Rej, required for synthesis of Gag protein. J. Virol. **83:**12483–12498.
18. **Jern, P., and J. M. Coffin.** 2008. Effects of retroviruses on host genome function. Annu. Rev. Genet. **42:**709–732.
19. **Kubota, S., R. Furuta, M. Maki, and M. Hatanaka.** 1992. Inhibition of human immunodeficiency virus type 1 Rev function by a Rev mutant which interferes with nuclear/nucleolar localization of Rev. J. Virol. **66:**2510–2513.
20. **Martoglio, B., R. Graf, and B. Dobberstein.** 1997. Signal peptide fragments of preprolactin and HIV-1 p-gp160 interact with calmodulin. EMBO J. **16:**6636–6645.
21. **Mermer, B., B. K. Felber, M. Campbell, and G. N. Pavlakis.** 1990. Identification of trans-dominant HIV-1 rev protein mutants by direct transfer of bacterially produced proteins into human cells. Nucleic Acids Res. **18:**2037–2044.
22. **Mertz, J. A., A. B. Chadee, H. Byun, R. Russell, and J. P. Dudley.** 2009. Mapping of the functional boundaries and secondary structure of the mouse mammary tumor virus Rem-responsive element. J. Biol. Chem. **284:**25642–25652.
23. **Mertz, J. A., M. M. Lozano, and J. P. Dudley.** 2009. Rev and Rex proteins of human complex retroviruses function with the MMTV Rem-responsive element. Retrovirology **6:**10.
24. **Mertz, J. A., M. S. Simper, M. M. Lozano, S. M. Payne, and J. P. Dudley.** 2005. Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. J. Virol. **79:**14737–14747.
25. **Mura, M., et al.** 2004. Late viral interference induced by transdominant Gag of an endogenous retrovirus. Proc. Natl. Acad. Sci. U. S. A. **101:**11117–11122.
26. **Murcia, P. R., F. Arnaud, and M. Palmarini.** 2007. The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of the JSRV Gag. J. Virol. **81:**1762–1772.
27. **Oliveira, N. M., K. B. Farrell, and M. V. Eiden.** 2006. In vitro characterization of a koala retrovirus. J. Virol. **80:**3104–3107.
28. **Palmarini, M., et al.** 2000. Molecular cloning and functional analysis of three type D endogenous retroviruses of sheep reveal a different cell tropism from that of the highly related exogenous Jaagsiekte sheep retrovirus. J. Virol. **74:**8065–8076.
29. **Palmarini, M., M. Mura, and T. E. Spencer.** 2004. Endogenous betaretroviruses of sheep: teaching new lessons in retroviral interference and adaptation. J. Gen. Virol. **85:**1–13.
30. **Palmarini, M., J. M. Sharp, M. De las Heras, and H. Fan.** 1999. Jaagsiekte sheep retrovirus is necessary and sufficient to induce a contagious lung cancer in sheep. J. Virol. **73:**6964–6972.
31. **Palmarini, M., J. M. Sharp, C. Lee, and H. Fan.** 1999. In vitro infection of ovine cell lines by Jaagsiekte sheep retrovirus (JSRV). J. Virol. **73:**10070–10078.
32. **Rhee, S. S., H. X. Hui, and E. Hunter.** 1990. Preassembled capsids of type D retroviruses contain a signal sufficient for targeting specifically to the plasma membrane. J. Virol. **64:**3844–3852.
33. **Sfakianos, J. N., and E. Hunter.** 2003. M-PMV capsid transport is mediated by Env/Gag interactions at the pericentriolar recycling endosome. Traffic **4:**671–680.
34. **Spencer, T. E., G. A. Johnson, F. W. Bazer, R. C. Burghardt, and M. Palmarini.** 2007. Pregnancy recognition and conceptus implantation in domestic ruminants: roles of progesterone, interferons and endogenous retroviruses. Reprod. Fertil. Dev. **19:**65–78.
35. **Spencer, T. E., M. Mura, C. A. Gray, P. J. Griebel, and M. Palmarini.** 2003. Receptor usage and fetal expression of ovine endogenous betaretroviruses: implications for coevolution of endogenous and exogenous retroviruses. J. Virol. **77:**749–753.
36. **Tarlinton, R. E., J. Meers, and P. R. Young.** 2006. Retroviral invasion of the koala genome. Nature **442:**79–81.
37. **Vardabasso, C., L. Manganaro, M. Lusic, A. Marcello, and M. Giacca.** 2008. The histone chaperone protein nucleosome assembly protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription. Retrovirology **5:**8.
38. **Varela, M., Y. H. Chow, C. Sturkie, P. Murcia, and M. Palmarini.** 2006. Association of RON tyrosine kinase with the Jaagsiekte sheep retrovirus envelope glycoprotein. Virology **350:**347–357.
39. **Varela, M., T. E. Spencer, M. Palmarini, and F. Arnaud.** 2009. Friendly viruses: the special relationship between sheep and retroviruses. Proceedings of the New York Academy of Sciences **1178:**157–172.
40. **von Heijne, G.** 1990. Protein targeting signals. Curr. Opin. Cell Biol. **2:**604–608.
41. **Walter, P., and A. E. Johnson.** 1994. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. Annu. Rev. Cell Biol. **10:**87–119.
42. **Weiss, R. A.** 1995. Retrovirus receptors. Cell **82:**531–533.