# The RCAS Vector System

STEPHEN H. HUGHES

HIV Drug Resistance Program, National Cancer Institute-Frederick, Frederick, Maryland, USA

Retroviruses have long been studied in animal models and cell culture. Initially, the primary reason to study retroviruses was that certain retroviruses efficiently induced tumours in animals and transformed cells in culture. Experiments with one of these rapidly oncogenic retroviruses, Rous sarcoma virus (RSV), led to the discovery of cellular oncogenes, a discovery that provides the foundation for our current understanding of cancer. The discovery of pathogenic human retroviruses, in particular the identification about 25 years ago of human immunodeficiency virus (HIV)-1 as the causative agent for acquired immunodeficiency syndrome (AIDS), invigorated the study of retroviruses and, to a significant extent, changed the direction of the field. Before HIV was discovered, retroviral research focused primarily on cellular oncogenes and viral oncogenesis; now much more effort is expended on studies of viral replication, the host's immune response, and the development of effective antiviral therapies. A more comprehensive overview is given in Chapter 1 of Coffin, Hughes and Varmus "Retroviruses" (1997).

At approximately the same time that HIV-1 was first identified, the earliest efforts were made to convert retroviruses into efficient vectors that could be used to express genes either in cultured cells or in animals. Today, retroviral vectors are widely used as tools to study the properties of individual genes and proteins, to probe development, and to study the causes of cancer in animal models. Although the initial excitement based on the hope that retroviruses would be powerful tools for gene therapy in humans is still largely unfulfilled, there is still some hope that safe and reliable retroviral gene therapy vectors can be developed.

The purpose of this review is to describe the development and uses of a particular family of retroviral vectors, the RCAS vectors. The name RCAS stands for Replication Competent ALV LTR with a Splice acceptor. These vectors are different from most other retroviral vectors because they are replication-competent and because they derive from a parental virus of avian origin. This review does not provide a comprehensive overview of the literature, either for the RCAS vectors themselves or for the large number of alternative viral vector systems now available; however, it does describe some of the ways RCAS vectors are used in research and discusses some of the advantages (and disadvantages) of these vectors. Those who want additional information about the available RCAS vectors and their uses are urged to visit the RCAS website http://home.ncifcrf.gov/hivdrp/RCAS and the tva website http://rex.nci.nih.gov/RESEARCH/basic/varmus/tva-web/tva2.html.

## Historical perspective

Two specific events characterize the retroviral life cycle: 1) reverse transcription; the process that converts the single-stranded RNA genome found in virion into linear double-stranded DNA, 2) integration; the insertion of this linear DNA into the genome of the host (a simplified version of the retroviral life cycle is shown in Fig. 1). The fact that the retroviral life cycle involves reverse transcription and integration has important consequences for both retroviruses and their hosts.

Reverse transcription is a highly recombinogenic process. This means that retroviruses (and retroviral vectors) readily recombine with closely related retroviruses of either endogenous or exogenous origin. In addition, if there are duplications in the viral genome, reverse transcription results in frequent recombination between the duplicated segments, causing deletions. Deletions also arise less frequently between segments with no obvious homology.

The integration of viral DNA into the host genome is a mutagenic event; in most cases integration of a DNA copy of the viral genome does not obviously affect the infected cell. The integration of viral DNA occurs at many places in the host genome, and while it does not appear to be strictly random, viral DNA is inserted both in genes and in intragenic regions. However, even when integration disrupts a gene, this is usually acceptable in cells derived from an organism with a diploid genome. Rarely, viral DNA is inserted near a host gene that plays a critical role in the regulation of cell growth. Such genes are called cellular oncogenes, because altering their expression and/or biochemical properties can lead to unrestricted growth and cancer. The insertion of a retroviral genome (or provirus) can enhance the expression of a nearby oncogene, which can affect the growth of the cell. Specific integrations in or near oncogenes

Received May 28, 2004.

Corresponding author: Stephen H. Hughes, P.O. Box B, Building 539, Room 130A, Frederick, MD 21702-1201, USA. Tel.: (301) 846-1619; Fax: (301) 846-6966; e-mail: hughes@ncifcrf.gov.

Abbreviations: AIDS – acquired immunodeficiency syndrome, ALV – avian leukosis virus, ASLV – avian sarcoma-leukosis virus, CEF – chicken embryo fibroblast, DR – direct repeat, HIV – human immunodeficiency virus, MLV – murine leukaemia virus, RCAS – replication competent ALV LTR with a splice acceptor, RSV – Rous sarcoma virus.

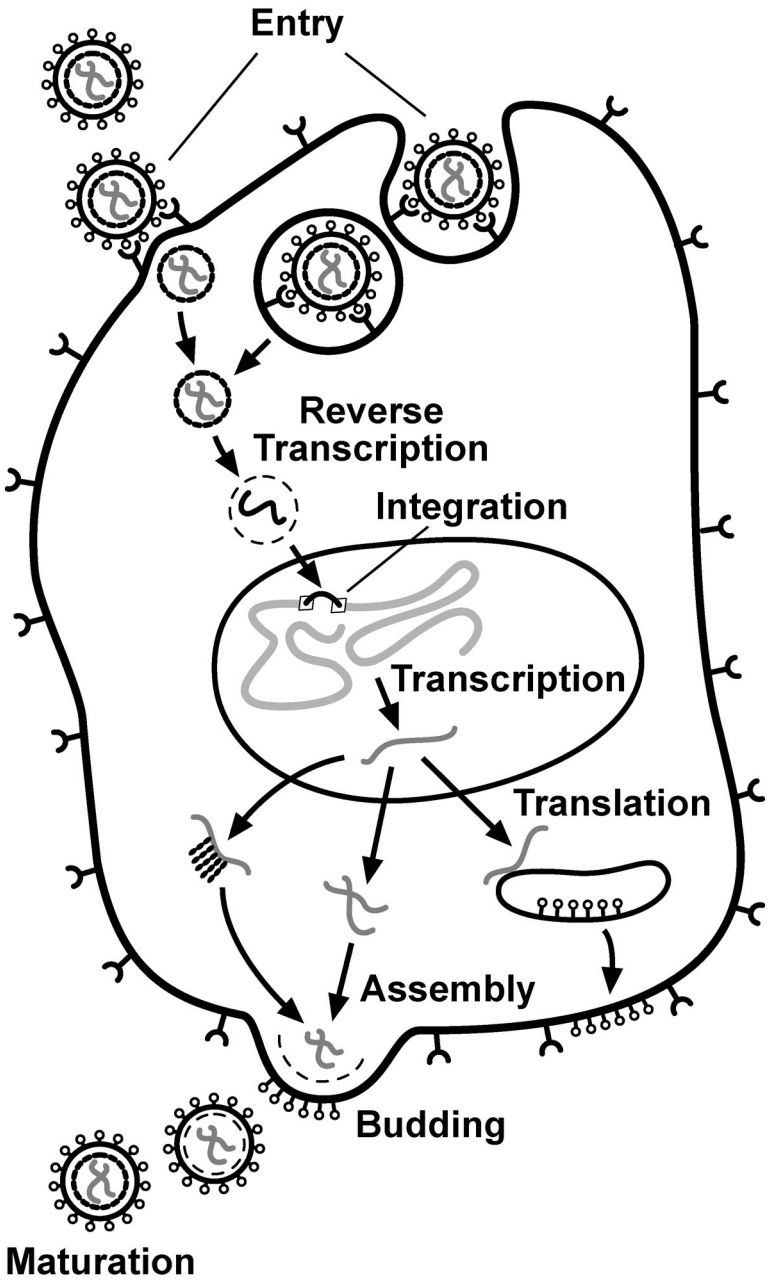


*Fig. 1.* The retroviral life cycle. The diagram shows, in simplified form, the processes involved in the retroviral life cycle. The life cycle proceeds from the top of the diagram to the bottom. In the first step, the viral envelope glycoprotein interacts with its cognate receptor on the surface of the cell. For most retroviruses the binding of the receptor leads directly to fusion of the viral and cellular membranes, which introduces the viral core into the cytoplasm of the cell (shown at the top left). However, for the ASLV viruses (and the RCAS vectors), the fusion of the viral entry requires both receptor binding and low pH. Entry of these retroviruses appears to occur in vesicles in the cytoplasm; it is the acidification of these vesicles that provides the low pH environment required for viral entry (shown at the top right). After the viral core is introduced into the cytoplasm, the diploid single-stranded RNA genome found in the virion is converted into double-stranded DNA (reverse transcription). The DNA form of the viral genome, with associated viral proteins, is then integrated into the host genome. In some cases (lentiviruses, ASLVs), this viral DNA/protein complex is able to transit the nuclear membrane; however, other retroviruses (MLV and its close relatives) cannot transit the nuclear membrane. For these viruses, successful infection requires the breakdown of the nuclear membrane that occurs during cell division. Once integrated, the retroviral genome is treated like a cellular gene and is transcribed into RNA by host DNA-dependent RNA polymerase. Both spliced and unspliced RNAs are produced; these are exported from the nucleus and used as mRNAs and, in the case of the unspliced RNA, as both mRNA and genomic RNA. Viral proteins are produced, by translation, as polyprotein precursors are brought together with viral RNA at the plasma membrane. There the nascent virions are assembled and budded from the cell. Newly budded virions still contain unprocessed viral proteins; the proteolytic processing of these polyproteins (maturation) converts the newly budded, immature virions into infectious viruses.

are rare when one considers a single infected cell; however, in an animal, the number of infected cells can be quite large and events that are rare in individual cells are common in the whole animal. This is one of the ways in which retroviral infections lead to tumour formation in animals. Unfortunately, proviral insertions near oncogenes have occurred in gene therapy protocols in which retroviral vectors were used to treat human patients; this is a significant problem that needs to be resolved if retroviral vectors are to be widely used for gene therapy (McCormack and Rabbitts, 2004).

The insertion of a provirus near a cellular gene can also lead to the acquisition of part, or all, of the cellular gene by the virus. Although there is an important exception that is directly relevant to the development of the RCAS vectors, in almost every case, the acquisition of the cellular gene by a replication-competent retrovirus was accompanied by the loss of one or more viral genes. Most oncogene-containing retroviruses are replication-defective; they cannot replicate unless the proteins encoded by the missing viral genes are supplied in *trans*, either from an endogenous virus carried by the host cell or by an accompanying replication-competent helper virus. The exception is RSV. RSV acquired the cellular oncogene *src* but did not lose any of the viral genes carried by the parental avian leukosis virus (ALV); in contrast to other oncogene-containing retroviruses, RSV is replication-competent.

The recognition that oncogene-containing retroviruses are naturally occurring vectors (viruses that carry and express genes derived from the host cell genome) provided the inspiration for the development of retroviral vectors in research laboratories. Retroviral vectors are designed to make it easy to insert (and express) foreign genes. The resulting vectors can be used to infect cells in culture or in animals; the vector facilitates the transfer of the gene of interest into the target cell. Because retroviral integration is efficient, retroviral vectors efficiently introduce the genes they carry into the genome of the host cell. Like the naturally occurring retroviral vectors on which they are based, most of the laboratory-derived retroviral vectors are replication-defective. However, most of the RCAS vectors, which are based on RSV, are replication-competent in avian cells. We believe that there are applications for which replication competence is advantageous: there is no need to provide helper sequences in *trans*. There is no problem with recombination between the helper sequences and the vector. There is no need to provide a selectable marker or select infected cells; a replication-competent vector spreads rapidly and in a short time will infect essentially all the cells in a culture dish. In general, because the virus spreads, the efficiency of the initial transfection is not particularly important. Moreover, if a replication-defective vector is required, there are replication-defective RCAS derivatives (in particular, the defective vector BBAN – which is discussed below). Although the RCAS vectors replicate efficiently in avian cells, these vectors are constitutively replication-defective in mammalian cells. This has implications for using these vectors in the mouse model; these are discussed later in this review.

Retroviruses infect cells through an interaction between the viral envelope glycoprotein and a cognate receptor on the surface of the cell (see Fig. 1). This interaction ultimately leads to the fusion of the viral membrane and the membrane of the target cell. This fusion event introduces the virion core into the cytoplasm of the target cell. Different retroviral envelopes recognize different receptors on the surface of host cells. The avian sarcoma-leukosis virus (ASLV) family, which includes the ALVs, RSV and the RCAS vectors, have a variety of envelope subtypes, designed by letters. Envelopes A–J have been described; A–E are the subgroups most commonly used in the laboratory. The cellular receptor for the A envelope is tva, a protein of unknown function which is, by both sequence and structure, related to an LDL receptor repeat (Bates et al., 1993). The B, D and E envelopes recognize aspects of a single receptor, tvb, which is a membrane protein of the fas ligand receptor family (Brojatsch et al., 1996; Adkins et al., 2000). The receptor for the C envelope has not yet been identified.

The RCAS vectors are propagated on cells of avian origin. Originally that meant preparing chicken embryo fibroblasts (CEF) from chicken embryos. CEFs, like other primary cells, have a limited lifespan in culture, usually 20–30 passages. Most chicken strains contain endogenous proviruses that are closely related to the ASLV virus and the RCAS vectors that derive from them. As has already been mentioned, this can lead to recombination between the vector and the endogenous viruses. To avoid this problem, a line of chickens was developed that has no endogenous proviruses closely related to the ASLVs; this line of chickens is called EV-0. Initially, there were no avian cell lines that were free of retroviruses that are closely related to the ASLVs that propagated the RCAS vectors efficiently. For example, the RCAS vectors grow relatively poorly on the quail cell line QT-6. The development of a permanent, non-transformed chicken fibroblastic cell line (DF-1) from an EV-0 embryo has resolved most of these problems; we now routinely use DF-1 cells to propagate the RCAS vectors (Himly et al., 1998; Schaefer-Klein et al., 1998).

ASLV viruses do not normally infect mammalian cells efficiently because none of the standard envelope subgroups effectively recognizes the cognate receptor on a mammalian cell. There are two simple solutions to this problem. One solution is to use envelope genes from viruses that can infect mammalian cells. We have prepared replication-competent RCAS vectors that use either the ecotropic or the amphotropic envelopes from the murine leukaemia virus (MLV) (Barsov and Hughes, 1996; Barsov et al., 2001). Alternatively, VSV-G

protein can be used to prepare BBAN viral stocks; however, these viruses are replication-defective. Attempts to make replication-competent VSV-G/RCAS derivatives have failed; presumably VSV-G is too toxic for the host cells to support the spread of a replication-competent retrovirus (unpublished observations). The other solution is to express an ASLV receptor (usually tva) in a mammalian cell (or a transgenic animal); this allows efficient infection with RCAS derivates that have the cognate envelope (usually subgroup A) (Federspiel et al., 1994; Fisher et al., 1999; Orsulic, 2002). This protocol has been used to generate strains of mice that can be infected in specific cells/tissues; this application is discussed below.

Although mammalian cells can be efficiently infected by RCAS vectors that carry MLV envelopes or by making use of cells/animals that express tva, the infected cells do not produce infectious virus. There is a problem with virion assembly; however, that problem can be overcome by using the Rev/RRE system from HIV-1 (Natsoulis et al., 1995). The viruses produced in this fashion are very poorly infectious, and we have not been able to find a way to get mammalian cells to efficiently propagate RCAS vectors. While this barrier does limit some of the experiments we would like to do, it also tells us that the RCAS vectors are relatively safe. Even our best efforts have failed to produce RCAS derivatives that replicate to any significant extent in mammalian cells.

## The RCAS system

Although the details are a bit more complicated, the fundamental idea behind the RCAS vectors is a simple one. A molecular clone of the DNA genome of the SR-A strain of RSV was modified to delete the *src* oncogene and a unique restriction site (*Cla*I) was inserted in place of the deleted *src* gene (Hughes and Kosik, 1984; Hughes et al., 1987). In RSV, the *src* gene is expressed from a spliced mRNA; the splice acceptor that is used to generate the *src* message is derived from the cellular *src* gene (see Fig. 2). The presence of the *src* splice acceptor in the RSV viruses provides strong support for the idea that the *src* oncogene was acquired when an ALV provirus integrated into the *src* gene and that the RSV was generated from an RNA arising from this integrated ALV provirus (Swanstrom et al., 1983). The alternative hypothesis, that c-*src* RNA was adventitiously packaged into an ALV virion and *src* was acquired by reverse transcription, does not explain the presence of a c-*src*-derived splice acceptor. In RSV, *src* is flanked by direct repeats approximately 100 bp long (these direct repeats are called DRs); the second copy of the DR was apparently acquired in the illegitimate recombination event that led to the acquisition of c-*src* by the retrovirus. However, the presence of the two DRs allows the *src* gene to be lost rapidly during viral replication because, when the RNA genome is copied into DNA, reverse transcriptase can jump from the 3' to the 5' DR. Not surprisingly, if another gene is inserted in place of *src* in a vector that contains both DRs, it too is lost rapidly. ALVs have one copy of the DR sequence, one copy of DR is both necessary and sufficient for viral replication (Sorge et al., 1983). The exact role of the DR is somewhat controversial; it appears to have a role in viral RNA transport and, perhaps indirectly, in packaging viral RNA into virions (Sorge et al., 1983; Ogert et al., 1996; Simpson et al., 1997; Aschoff et al, 1999). Both copies of DR that are present in the SR-A clone used to construct RCAS are fully functional; the RCAS vectors contain the downstream copy of DR, the upstream copy of DR was deleted to reduce the loss of inserted genes (see Fig. 2).

Although removing one copy of the DR substantially reduced the rate at which inserted sequences are lost from the RCAS vectors, inserted sequences still can be lost. This is probably the result of reverse transcriptase jumping between segments that have little or no homology. In general, viruses with smaller genomes (in the case of the RCAS vectors, those that have lost their inserted sequences) replicate more rapidly than those that retain an inserted sequence. For this reason, it is better to generate a fresh stock of an RCAS vector by transfecting cells with cloned viral DNA than to passage viral stocks from one culture to another. However, it should be remembered that it is viral passage (e.g., reverse transcription) that leads to the loss of inserted sequences; with the exceptions noted below, it is generally acceptable to passage cultured cells once they have been fully infected by an RCAS vector. Once integrated, proviral DNA is quite stable. However, in designing inserts for RCAS vectors there are several potential problems that should be avoided: 1) Sequences that interfere with the expression of the full-length viral genome. 2) Sequences that contain direct repeats or that duplicate other regions of the vector genome. 3) Sequences that are too large for the vector. 4) Sequences that, when expressed, are toxic for the host cell.

1) In general, sequences that lead to termination of the RNA transcript and/or polyadenylation should be removed from the insert. Successful propagation of an RCAS vector requires the efficient synthesis of full-length genomic RNA. Sequences that interfere with this process are strongly selected against; in practice, the viruses that are obtained will be deleted for sequences that interfere with synthesis of the full-length viral RNA.
2) Reverse transcriptase will jump between any direct repeats in the viral genome; the original DRs are simply an example of a larger problem. As such, good vector design will avoid creating vectors with direct repeats unless generating the deletion is the desired result.
3) Although the underlying mechanism is not well understood, there are limits on the size of a retroviral

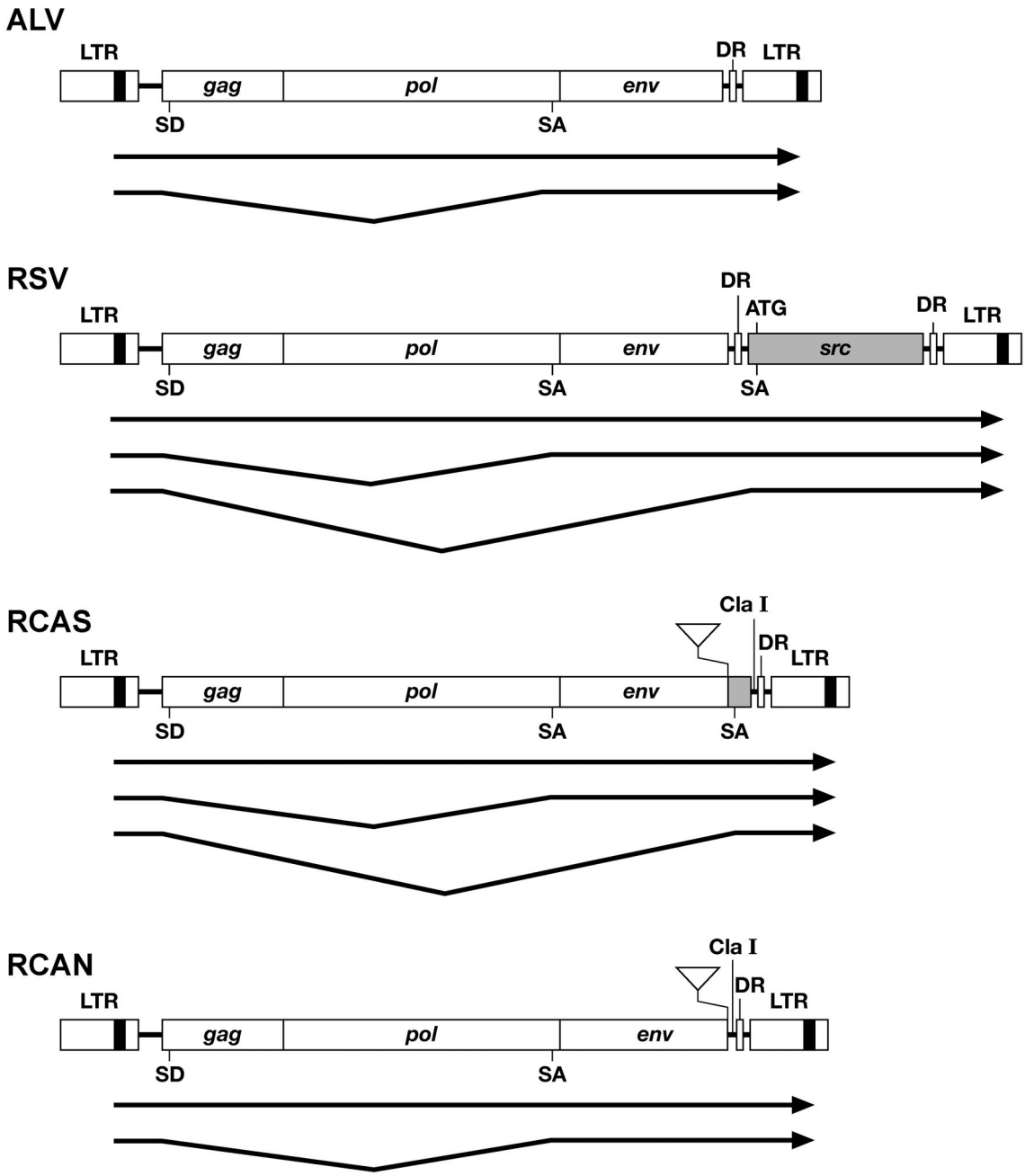


*Fig. 2.* A comparison of ALV, RSV, RCAS and RCAN. The diagrams (which are not to scale) show the organization of viral DNA genomes, the location of the genes (*gag*, *pol*, and *env*), the direct repeats (DR), the splice donor (SD) and splice acceptor (SA) sites. Under each diagram are drawings of the full-length and spliced RNAs produced from each viral DNA. The *src* gene of RSV carries a splice acceptor (SA) that leads to the production of a separate spliced *src* message. The *src* gene also contains its own initiator ATG. The upstream DR has been deleted from RCAS and RCAN (∇); *src* has also been deleted and replaced by a *Cla*I site. There is a small segment of the *src* gene in RCAS; this segment carries the *src* splice acceptor. This segment has been deleted from RCAN.

genome. In the ASLV viruses, the genome can be a bit larger than the genome of RSV. Since the *src* gene in RSV is about 2.0 kb, inserts into the RCAS vectors can be a bit larger than the *src* gene. We recommend, in the RCAS vectors, that inserts should be smaller than 2.5 kb. The size limit is not defined simply by the number of nucleotides in the insert (or the viral genome). In some (rare) cases, inserts larger than 2.5 kb have been successfully carried by RCAS vectors; likewise, in some cases, inserts smaller than 2.5 kb have not worked well. A simple, plausible explanation is that the key issue is not the length of the RNA genome, but the volume it occupies when optimally folded. All retroviruses have, in the virion, two copies of their RNA genomes, together with host tRNAs and some other small RNAs of host origin. In the virion the two copies of the viral genome are folded into a single, relatively stable, compact struc-

ture. Presumably, the size limit on the retroviral genome depends on the size of the folded dimeric RNA. Because some inserted sequences will, in the context of the rest of the RCAS genome, be able to fold more compactly than others, there are some differences in how long inserts can be. However, length does matter. No one has reported successfully using an insert in an RCAS vector longer than 3.0 kb; likewise there are no clear reports of size-related problems with inserts smaller than 2.0 kb.

4) Efficient viral replication requires the host cell to be in a healthy state. Cells that are made to express proteins that are toxic or interfere with cell growth will be overgrown by cells that don't express these proteins. As a consequence it is difficult to use the RCAS vectors to express genes that are toxic and/or interfere with cell growth. Attempts to use the RCAS vectors to express such proteins often result in the loss of the inserted gene; however, it is possible to do short-term experiments with proteins that interfere with cell growth if the initial transfection efficiency is high so that the entire culture is rapidly infected (Givol et al., 1998).

Reverse transcription produces a linear viral DNA that is subsequently integrated into the host genome. However, in the nucleus of infected cells, a portion of the linear viral DNA is circularized by host enzymes (Butler et al., 2002). This produces one-LTR circles (by homologous recombination within the LTRs) and two-LTR circles (by end-to-end joining of the linear viral DNA) (see Fig. 3). These circular forms have no direct role in the viral life cycle, but are extremely useful because they are easy to clone and provide information about mutations and the process of reverse transcription *in vivo*. In addition to the one-LTR and two-LTR circles, more complex circular forms arise by integration of the ends of the linear viral DNA into the body of the linear DNA (autointegrants). The SR-A recombinant DNA clone used to develop the RCAS clones was derived from a two-LTR circle. The first generation of vectors in the RCAS family (which were not called RCAS vectors but had complex names based on the deletions used to generate the vectors) was circularly permuted (Hughes and Kosik, 1984). A bacterial plasmid was inserted into the *env* gene of these vectors; before the vector DNA could be transfected into cells, the plasmid had to be removed and the viral DNA ligated to generate a functional copy of the viral genome. This was inconvenient and inefficient. Non-permuted versions were prepared, at the same time the vectors were given the name RCAS (Replication Competent with an ALV LTR and a Splice acceptor) (Hughes et al., 1987). Because the RCAS derivatives were prepared from a circular viral DNA, the plasmids used to generate the viral stocks have small ter-

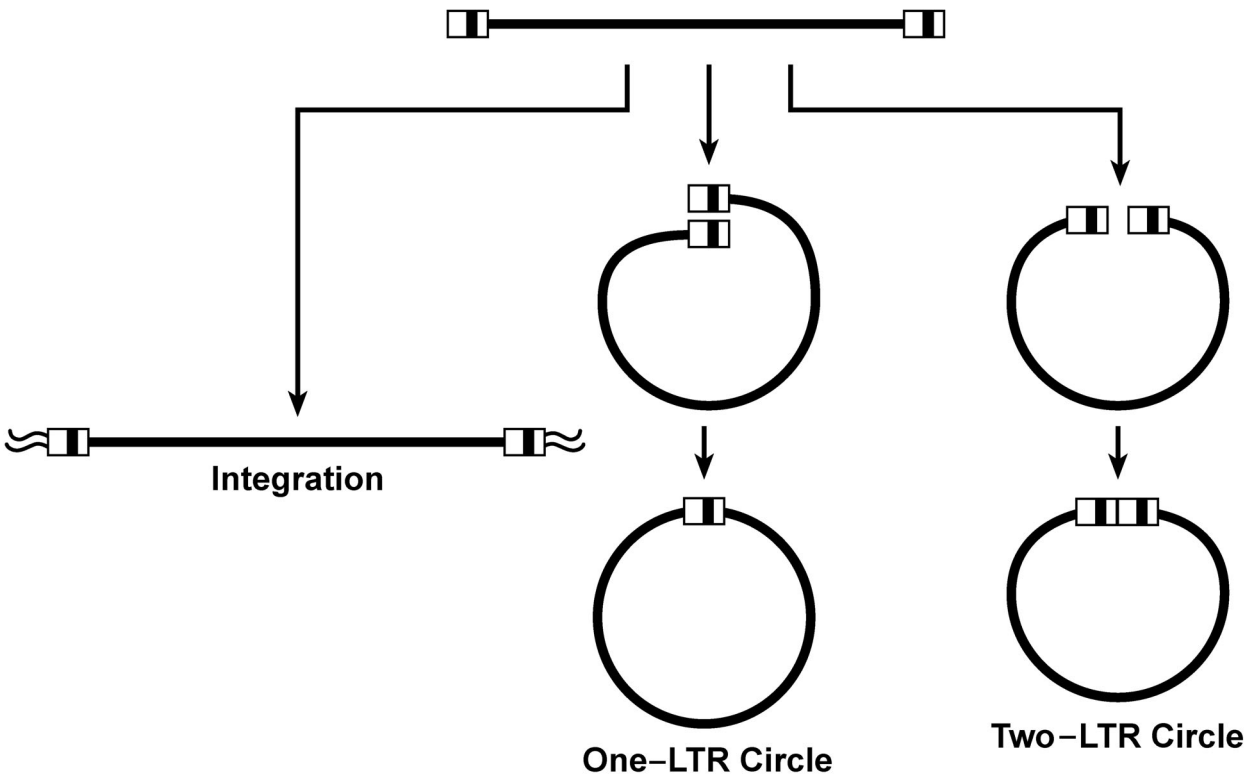


*Fig. 3.* Forms of viral DNA found in infected cells. The linear DNA shown at the top of the drawing is generated by reverse transcription. This linear DNA is transported into the nucleus, where it can be integrated into the host genome (left side of the drawing) or converted into one-LTR or two-LTR circles by host enzymes (right side of the diagram). The drawing is not to scale.

minal redundancies (there is, on both ends of RCAS viral genomes, a small amount of viral DNA sequence beyond the LTRs). For most purposes, this is irrelevant; however, for some types of analyses of the roles played by the viral sequences present in these redundant sequences (primarily the primer binding site or PBS and the polypurine tract or PPT), the redundancies do matter. As a consequence, we have prepared versions of the plasmids used to generate the RCAS vectors in which these terminal redundancies have been removed (described on the RCAS web site).

RCAS vectors have been prepared that contain several of the available ASLV *env* genes (A–E); these are designated RCAS(A), RCAS(B), etc. In an infected cell, expression of the envelope glycoprotein usually blocks the receptors on the surface of the cell. This is called receptor interference. One advantage of having RCAS vectors with several different envelope genes is that this makes it easy to infect a particular cell with two (or three) different vectors carrying different inserted genes. This procedure makes is possible to test the effects of two or more proteins (Givol et al., 1995, 1998).

In some cases, it is preferable to express an inserted gene from an internal promoter, from an internal ribosome entry segment (IRES), or not to express the inserted sequence at all. RCAN vectors (Replication Competent with an ALV LTR and No Splice acceptor) lack the *src* splice acceptor and, as a consequence, do not produce the smaller spliced message (see Fig. 2).

In the RCAS vectors, inserted genes are expressed from the LTR promoter; for this reason, the level of expression of an inserted gene depends on the level of expression of the LTR promoter. It is useful to be able to express inserted genes at various levels; we prepared RCAS derivatives in which the LTR promoter is expressed at different levels. Two elements in the vectors affect the level of expression: the enhancer in the LTR and the sequence of the *pol* gene. The ALV LTR contains a strong enhancer; there is little or no enhancer activity in the LTR of the corresponding endogenous avian retrovirus RAV-O. RCAS derivatives that contain the RAV-O LTR are called RCOS (Replication Competent with a RAV-O LTR and a Splice acceptor). The corresponding RCAN derivatives are the RCON vectors (Greenhouse et al., 1988).

The mechanism by which the *pol* sequences enhance the expression from the LTR is not well understood. It is clear that substituting the *pol* gene from the Bryan high-titre strain of RSV into the standard RCAS vectors (the resulting vectors are called RCASBP, which stands for RCAS Bryan Polymerase) increases the titre and the expression of the LTR promoter by about 5- to 10-fold (Nemeth et al., 1989; Petropoulos et al., 1991). The Bryan strain of RSV was selected for its rapid growth and the effect of the Bryan *pol* gene on viral replication suggests that it has enhancer-like properties. However, we were not able to locate a specific small enhancer element in the Bryan *pol* sequence (unpublished observations).

By combining the two *pol* genes and the two LTRs, four families of derivatives have been created: RCOS, RCOSBP, RCAS, RCASBP. These four sets of vectors differ in the level at which they express inserts (and the level at which they replicate) by about 5- to 10-fold. This makes it simple to express proteins at different levels. Alternatively (and particularly if tissue-specific expression is important), the level of expression of a protein can be controlled using an internal promoter in an RCAN or RCANBP vector. Although we have tested only a limited number of tissue-specific promoters, the presence of viral enhancers in RCASBP does not appear to perturb the tissue specificity of the internal promoter (Petropoulos et al., 1992).

The original RCAS vectors all have a *Cla*I site that is used to insert genes. In fact, there are actually two *Cla*I sites in the RCAS plasmids; however, the site that is not used to insert genes is subject to dam methylation and, as a consequence, plasmid DNA grown in standard *E. coli* strains (which contain an active dam methylase) is cleaved only at the desired site. There are now RCAS derivatives with multiple cloning sites and versions that allow the insertion of genes using the Gateway system (Loftus et al., 2001); however, if a DNA segment is inserted into original versions of the RCAS vectors, the segment must be converted into a *Cla*I fragment. We generated adaptor plasmids that can be used to facilitate

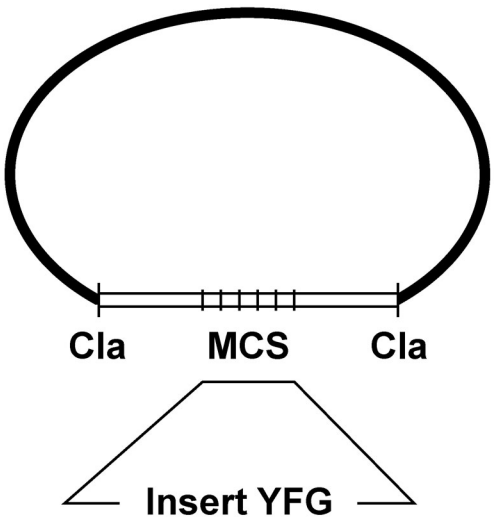


*Fig. 4.* Design of adaptor plasmids. The original RCAS vectors all have a *Cla*I site for inserting foreign genes (Fig. 2). This means that it is necessary to convert any DNA segment that will be inserted into the vectors into a *Cla*I fragment. Adaptor plasmids simplify this process. The plasmids all contain a multiple cloning site (MCS) flanked by *Cla*I sites. Your favorite gene (YFG) is inserted into the multiple cloning site and the desired segment recovered by *Cla*I digestion. Adaptor plasmids are available that contain additional elements including splice acceptors and initiator ATGs (see text and the RCAS web site).

the conversion of DNA segments into *Cla*I fragments for insertion into the RCAS vectors (Hughes et al., 1987). Basically, all the adaptor plasmids contain a multiple cloning site flanked by *Cla*I sites (see Fig. 4). The sequence of the region between the *Cla*I sites in these adaptors is designed not to interfere either with transcription or translation. Adaptor plasmids are available that supply, in addition to the multiple cloning site, an ATG. The Cla12Nco adaptor plasmid has an *Nco*I site (CCATGG) at the upstream end of the multiple cloning site; this *Nco*I/ATG is a favourable site for the initiation of translation when the Cla12Nco adaptor is used to insert a gene into an RCAS vector. The Cla12Nco *Nco*I site and multiple cloning sites were derived from pUC12N. pUC12N is an expression plasmid, and because it has the *Nco*I/ATG and multiple cloning sites that are present in Cla12Nco, it can be used to express exactly the same protein in *E. coli* that is expressed in avian (or mammalian) cells using the Cla12Nco adaptor in an RCAS vector. There is also an adaptor plasmid that carries not only an *Nco*I/ATG, but also a splice acceptor (SACla12Nco). More information about the adaptor plasmids is available on the RCAS web site.

Although one of the advantages of the RCAS vectors is that they are replication-competent, there are times when it is useful to have a replication-defective vector. One such replication-defective vector is BBAN (Boerkoel et al., 1993). The envelope gene has been deleted in BBAN; this allows the vector to accept inserts up to approximately 4.5 kb. BBAN can be complemented by ASLV envelope genes; however, it is often convenient to use VSV-G. BBAN complemented by VSV-G has a broad host range and there is no possibility that homologous recombination can reconstitute a replication-competent virus. As has already been discussed, it appears that, despite the fact that VSV-G efficiently complements BBAN in transient transfections, RCAS derivatives in which VSV-G has been substituted for the normal *env* gene do not replicate, presumably because VSV-G is too toxic (unpublished results).

## RCAS as a tool to study viral replication: RSVPs

One of the primary uses of RCAS is as a tool to study retroviral replication. An inserted marker gene can greatly simplify titering the virus [green fluorescent protein (GFP) is a convenient marker and the infected cells can be counted by looking in a microscope or by FACS]. In some experiments, it is important to recover either unintegrated or integrated viral DNA. A series of RCAS derivatives, the RSVPs (RCASBP Shuttle Vector Plasmids) have been prepared that make it simple and easy to recover either unintegrated or integrated viral DNA (Oh et al., 2002) (see Fig. 5). The RSVPs contain a selectable marker (either zeocin resistance or blastocidin resistance) that can be selected either in *E. coli* or higher eukaryotes. These vectors also contain a bacterial origin of replication so that circular DNAs derived from the RSVPs will replicate in *E. coli*. To simplify the recovery of viral DNA, the RSVPs also carry a lac operator (lacO) sequence. The presence of lacO makes it simple to enrich either unintegrated or integrated viral DNA. The Lac i protein binds selectively to lacO. DNAs bound to Lac i can be specifically eluted with IPTG. This provides a simple and rapid enrichment for RSVP DNA and makes it easy to molecularly clone either unintegrated or integrated RSVP DNA. The efficient recovery of retroviral DNA is valuable for experiments that monitor the mechanics of reverse transcription and/or viral mutation. For experiments that involve the analysis of the ends of viral DNA, it is convenient to use two-LTR circles; these circular DNAs preserve the ends of the parental linear DNAs from which they derive (see Fig. 3). However, integration can affect the pool of linear DNAs from which the circular forms are created. The circular forms

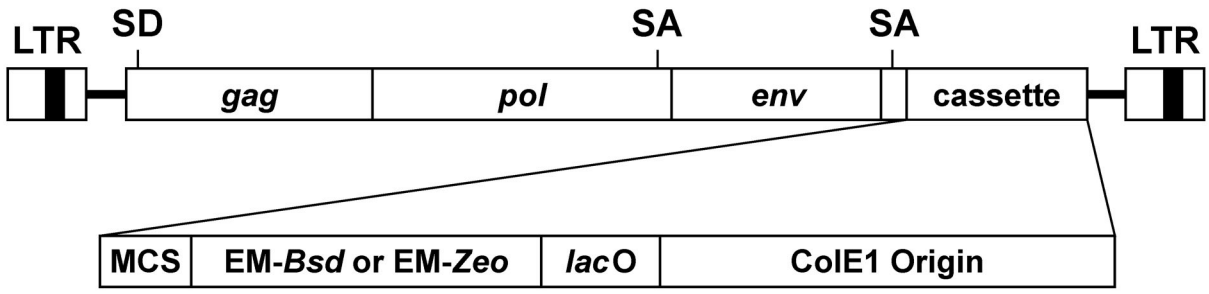


*Fig. 5.* RSVP vectors. An RSVP vector is shown at the top, an expanded view of the inserted cassette is shown at the bottom. The drawing is not to scale. The viral part of RSVP is RCASBP(A); the vector contains the *gag, pol* and *env* genes and has the standard splice donor (SD) and splice acceptor (SA). The cassette contains a multiple cloning site (MCS), and one of two selectable markers that work both in prokaryotic and eukaryotic cells: blastocidin resistance (*Bsd*) or zeocin resistance (*Zeo*). These are expressed as a spliced message when the vector is propagated as a retrovirus. When the vector is grown as a plasmid in *E. coli*, the *Bsd* and *Zeo* genes are expressed from a small prokaryotic promoter (EM). To allow the DNA to replicate as a plasmid, the cassette contains a ColE1 origin. The cassette also contains a lacO sequence, which facilitates recovery of the DNA from infected cells (see text).

of viral DNA derive from the portion of linear DNAs that do not integrate; integration preferentially removes linear DNAs with consensus ends from the pool of linear DNAs that gives rise to the circles. This can be an advantage for experiments designed to look at DNAs that have aberrant ends (these DNAs are enriched when linear DNAs with consensus ends are removed by the integration process). However, if it is important to quantitatively monitor the DNAs produced by reverse transcription, integration removes a substantial fraction of the linear DNAs with consensus ends. This potential problem can be overcome by making the vector integration-defective.

There are a number of reasons to study integrated proviruses. We need to better understand the process of integration in the infected cell and it will be important to better understand the rules that govern the choices different retroviruses make when they integrate their DNA. The available data shows that all retroviral DNAs can integrate into a large number of sites in the host genome; however, it appears that different retroviruses differ in their preferences for sites in the host genome (Schroder et al., 2002; Wu et al., 2003). The cassette that is used to make the RSVP DNA easy to recover can be introduced into the genomes of other retroviruses, which should facilitate experiments comparing the integration site preferences of different retroviruses.

## ASLV envelopes

The ability of retroviruses to infect target cells depends on a specific interaction between the envelope glycoprotein on the surface of the virus and the cognate receptor on the surface of the cell [see Coffin, Hughes and Varmus (1997) for a review]. As has already been discussed, there are a number of distinct envelope subgroups for the ASLVs and the RCAS vectors that derive from them. Not only are the ASLV envelopes important in directing the vectors to appropriate target cells; they have particular advantages for experiments designed to study the processes that lead to the fusion of viral and cellular membranes, which is a key step in viral infection. The flu HA protein is probably the best studied member of the large family of related viral proteins which carry out this fusion process. All these viral proteins are trimers. In their mature form, each of the monomers in this trimer is composed of two subunits (HA1 and HA2 for flu, SU and TM for retroviruses). When flu infects a target cell, HA1 binds to a sialic acid on the surface of the cell; this allows the virus to be imported into a vesicular compartment within the cell. If

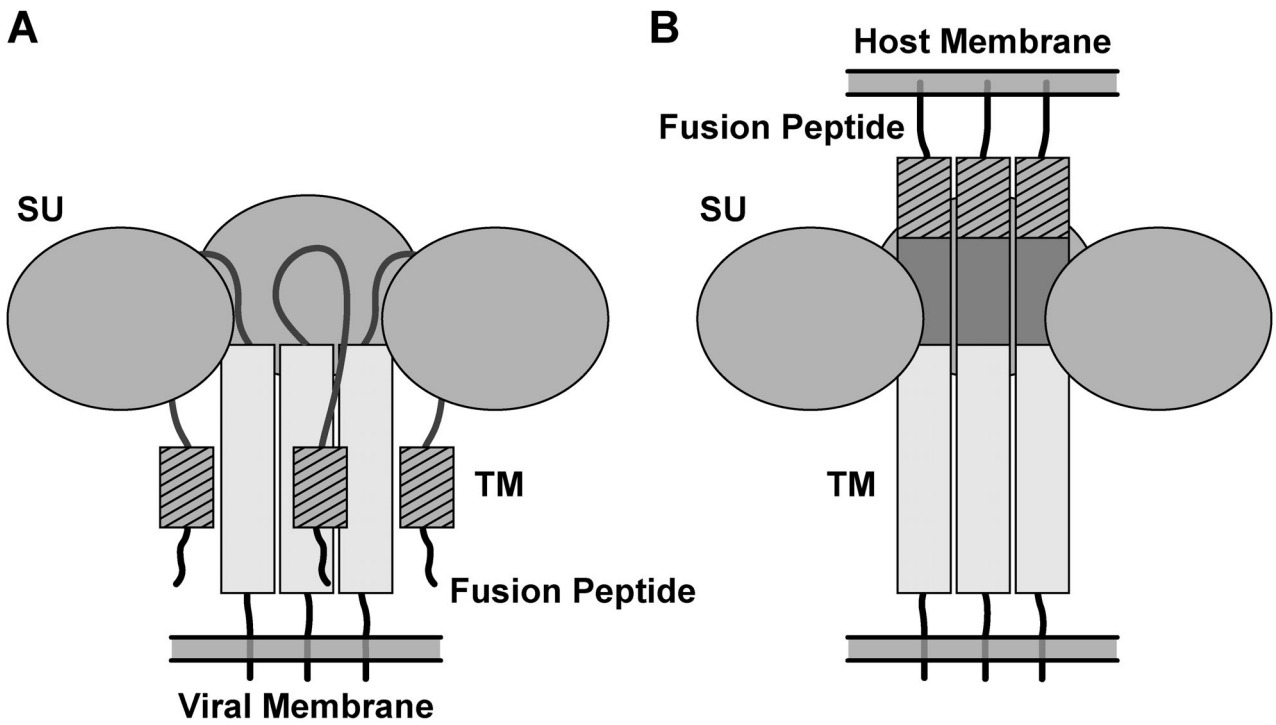


*Fig. 6.* Envelope rearrangements associated with viral infection. Panel A shows a diagram of the envelope glycoprotein of a retrovirus. The protein is a trimer composed of two sets of subunits, SU and TM. SU is the component that will interact with the receptor. TM is composed of two α-helical segments joined by a loop. TM anchors the protein in the viral membrane. Before SU interacts with a receptor, TM is a hairpin and the fusion peptide is directed toward the viral membrane. Panel B shows the first steps in the rearrangement of TM that leads to fusion of the viral and cellular membranes. In retroviral envelopes, this rearrangement depends on SU binding its receptor (not shown). This causes the loop in TM to convert to an α helix, which embeds the hydrophobic fusion peptide into the plasma membrane of the host cell. Subsequent rearrangements in TM bring the host membrane and viral membrane together, leading to membrane fusion and viral infection (see Fig. 7).

this vesicular compartment is acidified, the low pH triggers rearrangements of α helices within HA2. The first part of this rearrangement embeds the hydrophobic fusion peptide in HA2 into the membrane of the host vesicle; subsequent rearrangements bring the viral and host cell membranes together. Fusion of the membranes introduces the viral core into the cytoplasm of the cells, leading to infection.

For most retroviruses the overall process is similar, but the events that trigger the rearrangement of the HA2 component of a retroviral envelope (TM) are different (see Figs. 6 and 7). In the case of MLV envelope, pH plays no role in causing the rearrangements in TM, rather it is the binding of the SU component of the envelope to its cognate receptor that initiates the structural changes that result in fusion. This means that fusion takes place on the outer surface of the cell, rather than in an acidic vesicle (see Fig. 1). HIV-1 envelope interacts with two cellular proteins, a receptor (CD4) and a co-receptor (usually CXCR4 or CCR5). Both binding events cause changes in the structure of HIV-1 envelope; binding the receptor allows HIV-1 to efficiently bind the co-receptor. Binding the co-receptor leads to the rearrangements in TM that lead to fusion and infection. However, as is the case for MLV, HIV-1 envelope fusion does not require low pH and fusion occurs on the plasma membrane.

There is now good evidence that for the ASLV envelopes, both receptor binding and low pH are required for membrane fusion and viral entry (Mothes et al., 2000; Melikyan et al., 2004; Smith et al., 2004). In some ways, this places the ASLV envelopes somewhere between flu HA, for which the rearrangement requires low pH, but not receptor binding, and other retroviral envelopes, like MLV and HIV, which require receptor (and, in the case of HIV-1, co-receptor) binding, but not low pH.

The particular advantage of using the ASLV envelopes as tools to study the fusion process is that fusion depends on both receptor binding and low pH. Moreover, the tva and tvb receptors are cloned and portions of the receptors that are soluble and easy to generate can trigger the events in the avian envelope glycoproteins that lead to fusion. This allows the experimentalist to stop the reaction at intermediate steps in the fusion process, making it possible to investigate the structural and biochemical intermediates that underlie the fusion process.

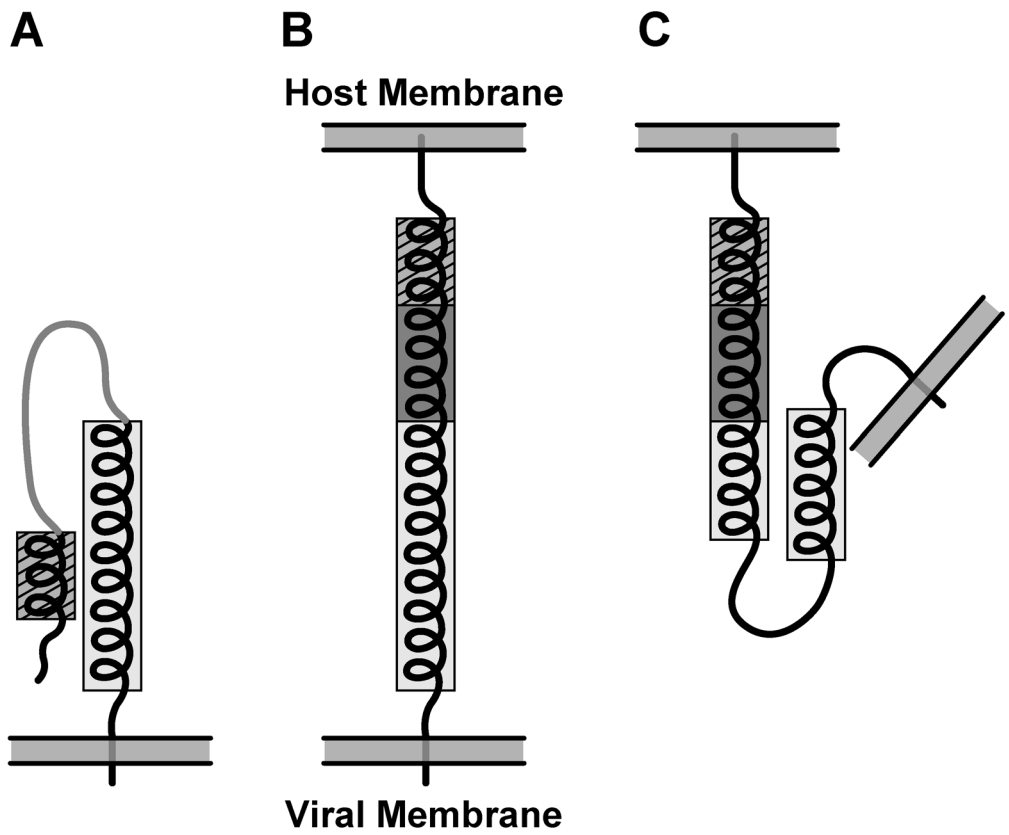


*Fig. 7.* Rearrangements in TM associated with viral infection. This diagram shows one TM; in the virus TM is trimeric (see Fig. 6). Initially, TM is composed of two α-helical components linked by a loop (Panel A). Interactions between the SU (not shown) and its receptor cause a rearrangement in TM, which embeds the fusion peptide in the host membrane (Panel B, see also Fig. 6). Subsequent rearrangements in TM (Panel C) bring the membranes close together, which leads to membrane fusion. For most retroviruses this secondary rearrangement in TM does not require any additional signal beyond binding the receptor (in the case of HIV-1 envelope, binding both the receptor and the co-receptor). However, it appears that the ASLV envelopes require a secondary signal (low pH) to carry out the rearrangement in TM shown in Panel C (see text).

## RCAS in animal models

Based on a survey of the literature, the most common use of the RCAS vectors is to study development in chicken embryos. Chicken embryos have several distinct advantages for developmental studies: 1) They are relatively large, even at early stages of development. 2) They develop outside the mother, and are more accessible to physical manipulation than mammalian embryos. 3) They are easy to obtain and relatively inexpensive.

Embryonic development involves an interplay of intra- and intercellular signals that cause the movement and differentiation of cells in the embryo. There are a variety of ways to perturb these signals; the RCAS vectors are a convenient way to overexpress proteins in developing avian embryos. The recent demonstration that RCAS vectors can be used to express RNAi opens up the possibility that the RCAS system will be used to block the expression of genes in the developing chicken embryo (Bromberg-White et al., 2004).

It has been exceptionally difficult to make transgenic birds, in part because it is difficult to stably deliver DNA to the newly fertilized egg. This is due, at least in part, to the fact that eggs must be fertilized in the hen's reproductive tract before the shell is added. By the time a fertilized chicken egg is laid, it contains thousands of cells. Injecting DNA into fertilized eggs after they have been laid does not yield transgenic birds at a significant frequency. Much better results have been obtained with retroviral vectors (Salter et al., 1987; Bosselman et al., 1989; Federspiel et al., 1991; Harvey et al., 2002; Mozdziak et al., 2003). Although more work needs to be done, retroviral vectors remain a promising way of creating transgenic birds. This is an important problem both because chickens are an important experimental model and because they are a major source of meat and eggs. It may be possible to use transgenic chickens to produce large amounts of valuable recombinant proteins; the egg is a convenient vehicle to collect such proteins (Harvey and Ivarie, 2003). It would also be useful to be able to introduce novel traits into chickens raised for meat or egg production. Being able to generate birds resistant to viral diseases (and other pathogens) would make it easier to maintain the birds in the extremely dense conditions found on a modern chicken farm.

The RCAS vectors have also been used in mice. Because mice do not express versions of the receptors that are efficiently recognized by viruses carrying any of the standard ASLV envelopes, mice are not normally infectable by the standard strains of ASLV, or by the RCAS vectors that derive from them. As has already been discussed, there are two solutions to this problem: 1) Substitute the normal ASLV envelope with that of a virus that can infect mammalian cells (we used the amphotropic and ecotropic envelope from MLV to generate derivatives that are replication-competent in the avian cells used to prepare the viral stocks and VSV-G to generate replication-defective derivatives). 2) Express a functional version of a receptor for an ASLV envelope in the mammalian cell (Federspiel et al., 1994). This procedure is particularly useful when used in transgenic mice because it is possible to place the receptor (commonly tva) under the control of a tissue-specific promoter. A number of different promoters have been used in combination with tva to generate mice that can be infected by RCAS vectors in specific cell/tissue types. This has led to the generation of a number of specific cancer models, and is being used to explore which particular oncogenes can cause tumours and/or metastasis in various cell types/specific tissues (Fisher et al., 1999; Orsulic, 2002) (a list of many of the available mouse strains is given on the tva web site).

## Infection of non-dividing cells

One potential issue in animal models, particularly in the mouse models, is the ability of the vector to infect non-dividing cells. MLV-based vectors do not infect non-dividing cells; apparently the residual virion core, which carries the reverse-transcribed viral DNA, cannot pass through the nuclear membrane. However, lentiviruses and lentiviral vectors do infect non-dividing cells with reduced, but acceptable, efficiency. Lentiviral vectors can be pseudotyped with ASLV envelopes and the resulting pseudotyped lentiviral vectors used in transgenic tva mice. However, it has recently been shown that ASLV viruses, and the RCAS vectors that derive from them, can also infect non-dividing cells (Hatziioannou and Goff, 2001; Katz et al., 2002; Greger et al., 2004). The published experiments were all done in cultured cells, and it is clear that, like lentiviruses, the ASLV viruses are less efficient when they infect non-dividing cells. However, the available data do not tell us exactly how much less efficient the ASLV viruses are in non-dividing cells. ASLV viruses grow to very high titres [unconcentrated RCAS-BP(A) stocks usually have titres greater than $10^7$ on avian cells] so a moderate loss of efficiency when these vectors are used to infect non-dividing cells could easily be tolerated. There are no published data on the ability of RCAS vectors to infect non-dividing cells in animals and no direct comparisons of the relative efficiencies of RCAS vectors and lentiviral vectors carrying the same promoter/marker cassette; a direct comparison would be quite useful.

## Summary/Reprise

There are a large number of RCAS vectors (and ancillary tools, including the adaptors, mammalian cell lines expressing receptors like tva, and tva transgenic mice). This system can be used for a number of different purposes including studying viral replication, gene function, development, and cancer. The recent demon-

stration that the vector system can be used to express RNAi suggests novel uses for the vectors and the recognition that RCAS vectors can be used to infect nondividing cells also suggests new applications. It is likely that, as we come to better understand retroviruses and their hosts, new uses for these vectors will be discovered. I certainly hope that this will be true.

*Acknowledgements*

I thank Hilda Marusiodis for help in preparing the manuscript, and Al Kane for help with the figures. This research was supported by the National Cancer Institute.

## References

Adkins, H. B., Brojatsch, J., Young, J. A. (2000) Identification and characterization of a shared TNFR-related receptor for subgroup B, D, and E avian leukosis viruses reveal cysteine residues required specifically for subgroup E viral entry. *J. Virol.* **74,** 3572-3578.

Aschoff, J. M., Foster, D., Coffin, J. M. (1999) Point mutations in the avian sarcoma/leukosis virus 3' untranslated region result in a packaging defect. *J. Virol.* **73,** 7421-7429.

Barsov, E. V., Hughes, S. H. (1996) Gene transfer into mammalian cells by a Rous sarcoma virus-based retroviral vector with the host range of the amphotropic murine leukemia virus. *J. Virol.* **70**, 3922-3929.

Barsov, E. V., Payne, W. S., Hughes, S. H. (2001) Adaptation of chimeric retroviruses in vitro and in vivo: Isolation of avian retroviral vectors with extended host range. *J. Virol.* **75,** 4973-4983.

Bates, P., Young, J. A., Varmus, H. E. (1993) A receptor for subgroup A Rous sarcoma virus is related to the low density lipoprotein receptor. *Cell* **74,** 1043-1051.

Boerkoel, C. F., Federspiel, M. J., Salter, D. W., Payne, W., Crittenden, L. B., Kung, H.-J, Hughes, S. H. (1993) A new defective retroviral vector system based on the Bryan strain of Rous sarcoma virus. *Virology* **195**, 669-679.

Bosselman, R. A., Hsu, R. Y., Boggs, T., Hu, S., Bruszewski, J., Ou, S., Kozar, L., Martin, F., Green, C., Jacobsen, F. et al. (1989) Germline transmission of exogenous genes in the chicken. *Science* **243,** 533-535.

Brojatsch, J., Naughton, J., Rolls, M. M., Zingler, K., Young, J. A. (1996) CAR1, a TNFR-related protein, is a cellular receptor for cytopathic avian leukosis-sarcoma viruses and mediates apoptosis. *Cell* **87,** 845-855.

Bromberg-White, J. L., Webb, C. P., Patacsil, V. S., Miranti, C. K., Williams, B. O., Holmen, S. L. (2004) Delivery of short hairpin RNA sequences by using a replication-competent avian retroviral vector. *J. Virol.* **78,** 4914-4916.

Butler, S. L., Johnson, E. P., Bushman, F. D. (2002) Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J. Virol.* **76,** 3739-3747.

Coffin, J. M., Hughes, S. H., Varmus, H. E. (1997) *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Federspiel, M. J., Crittenden, L. B., Provencher, L. P., Hughes, S. H. (1991) Experimentally introduced defective endogenous proviruses are highly expressed in chickens. *J. Virol.* **65,** 313-319.

Federspiel, M. J., Bates, P., Young, J. A. T., Varmus, H. E., Hughes, S. H. (1994) A system for tissue-specific gene targeting: Transgenic mice susceptible to subgroup A avian leukosis virus-based retroviral vectors. *Proc. Natl. Acad. Sci. USA* **91,** 11241-11245.

Fisher, G. H., Orsulic, S., Holland, E., Hively, W. P., Li, Y., Lewis, B. C., Williams, B. O., Varmus, H. E. (1999) Development of a flexible and specific gene delivery system for production of murine tumor models. *Oncogene* **18,** 5253-5260.

Givol, I., Givol., D., Hughes, S. H. (1998) Overexpression of p21waf1/cip1 arrests the growth of chicken embryo fibroblasts that overexpress E2F1. *Oncogene* **16,** 3115-3122.

Givol, I., Givol, D., Rulong, S., Resau, J., Tsarfaty, I., Hughes, S. H. (1995) Overexpression of human p21waf1/cip1 arrests the growth of chicken embryo fibroblasts transformed by individual oncogenes. *Oncogene* **11,** 2609-2618.

Greenhouse, J. J., Petropoulos, C. J., Crittenden, L. B., Hughes, S. H. (1988) Helper-independent retrovirus vectors with Rous-associated virus type O long terminal repeats. *J. Virol.* **62,** 4809-4812.

Greger, J. G., Katz, R. A., Taganov, K., Rall, G. F., Skalka, A. M. (2004) Transduction of terminally differentiated neurons by avian sarcoma virus. *J. Virol.* **78,** 4902-4906.

Harvey, A. J., Speksnijder, G., Baugh, L. R., Morris, J. A., Ivarie, R. (2002) Consistent production of transgenic chickens using replication-deficient retroviral vectors and high-throughput screening procedures. *Poult. Sci.* **81,** 202-212.

Harvey, A. J., Ivarie, R. (2003) Validating the hen as a bioreactor for the production of exogenous proteins in egg white. *Poult. Sci.* **82,** 927-930.

Hatziioannou, T., Goff, S. P. (2001) Infection of nondividing cells by Rous sarcoma virus. *J. Virol.* **75,** 9526-9531.

Himly, M., Foster, D. N., Bottoli, I., Iacovoni, J. S., Vogt, P. K. (1998) The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses. *Virology* **248,** 295-304.

Hughes, S., Kosik, E. (1984) Mutagenesis of the region between *env* and *src* of the SR-A strain of Rous sarcoma virus for the purpose of constructing helper-independent vectors. *Virology* **136,** 89-99.

Hughes, S. H., Greenhouse, J. J., Petropoulos, C. J., Sutrave, P. (1987) Adaptor plasmids simplify the insertion of foreign DNA into helper-independent retroviral vectors. *J. Virol.* **61,** 3004-3012.

Katz, R. A., Greger, J. G., Darby, K., Boimel, P., Rall, G. F., Skalka, A. M. (2002) Transduction of interphase cells by avian sarcoma virus. *J. Virol.* **76,** 5422-5434.

Loftus, S. K., Larson, D. M., Watkins-Chow, D., Church, D. M., Pavan, W. J. (2001) Generation of RCAS vectors useful for functional genomic analyses. *DNA Res.* **8,** 221-226.

McCormack, M. P., Rabbits, T. H. (2004) Activation of the T-cell oncogene *LMO2* after gene therapy for X-linked severe combined immunodeficiency. *New Engl. J. Med.* **350,** 913-922.

Melikyan, G. B., Barnard, R. J. O., Markosyan, R. M., Young, J. A. T., Cohen, F. S. (2004) Low pH is required for ASLV Env-induced hemifusion and fusion pore formation, but not for pore growth. *J. Virol.* **78**, 3753-3762.

Mothes, W., Boerger, A. L., Narayan, S., Cunningham, J. M., Young, J. A. (2000) Retroviral entry mediated by receptor

priming and low pH triggering of an envelope glcoprotein. *Cell* **103,** 679-689.

Mozdziak, P. E., Borwornpinyo, S., McCoy, D. W., Petitte, J. N. (2003) Development of transgenic chickens expressing bacterial beta-galactosidase. *Dev. Dyn.* **226,** 439-445.

Natsoulis, G., Seshaiah, P., Federspiel, M. F., Rein, A., Hughes, S. H., Boeke, J. D. (1995) Targeting of a nuclease to murine leukemia virus capsids inhibits viral multiplication. *Proc. Natl. Acad. Sci. USA* **92,** 364-368.

Nemeth, S. P., Fox, L. G., DeMarco, M., Brugge, J. S. (1989) Deletions within the amino-terminal half of the c-src gene product that alter the functional activity of the protein. *Mol. Cell. Biol.* **9,** 1109-1119.

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

Oh, J., Julias, J. G., Ferris, A. L., Hughes, S. H. (2002) Construction and characterization of a replication competent retroviral shuttle vector plasmid. *J. Virol.* **76,** 1762-1768.

Orsulic, S. (2002) An RCAS-TVA-based approach to designer mouse models. *Mamm. Genome* **13,** 543-547.

Petropoulos, C. J., Hughes, S. H. (1991) Replication-competent retroviral vectors for the transfer and expression of gene cassettes in avian cells. *J. Virol.* **65,** 3728-3737.

Petropoulos, C. J., Payne, W., Salter, D. W., Hughes, S. H. (1992) Appropriate *in vivo* expression of a muscle-specific promoter using avian retroviral vectors for gene transfer. *J. Virol.* **66,** 3391-3397.

Salter, D. W., Smith, E. J., Hughes, S. H., Wright, S. E., Crittenden, L. B. (1987) Transgenic chickens: insertion of retroviral genes into the chicken germ line. *Virology* **157,** 236-240.

Schaefer-Klein, J., Givol, I., Barsov, E. V., Whitcomb, J. M., VanBrocklin, M., Foster, D. N., Federspiel, M. J., Hughes, S. H. (1998) The EV-0 derived cell line DF-1 supports the efficient replication of avian leukosis-sarcoma viruses and vectors. *Virology* **248,** 305-311.

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

Simpson, S. B., Zhang, L., Craven, R. C., Stoltzfus, C. M. (1997) Rous sarcoma virus direct repeat cis elements exert effects at several points in the virus life cycle. *J. Virol.* **71,** 9150-9156.

Smith, J. G., Mothes, W., Blacklow, S. C., Cunningham, J. M. (2004) The mature avian leukosis virus subgroup A envelope glycoprotein is metastable, and refolding induced by the synergistic effects of receptor binding and low pH is coupled to infection. *J. Virol.* **78,** 1403-1410.

Sorge, J., Ricci, B., Hughes, S. H. (1983) Cis-acting RNA packaging locus in the 115-nucleotide direct repeat of Rous sarcoma virus. *J. Virol.* **48,** 667-675.

Swanstrom, R., Parker, R. C., Varmus, H. E., Bishop, J. M. (1983) Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. *Proc. Natl. Acad. Sci. USA* **80,** 2519-2523.

Wu, X., Li, Y., Crise, B., Burgess, S. M. (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300,** 1749-1751.