Emergence of vertebrate retroviruses and envelope capture

Felix J. Kim, Jean-Luc Battini, Nicolas Manel, and Marc Sitbon*

Institut de Génétique Moléculaire de Montpellier (IGMM), CNRS-UMR5535, IFR122, F-34293 Montpellier, cedex 5, France

Received 31 July 2003; returned to author for revision 10 September 2003; accepted 14 September 2003

Abstract

Retroviruses are members of the superfamily of retroelements, mobile genetic elements that transpose via an RNA intermediate. However, retroviruses are distinct from other retroelements in that their “transposition” is not confined to single cells but extends to neighboring cells and organisms. As such, the “transposition” of these elements is defined as infection. It appears that a key step in the conversion of a retrotransposon into a retrovirus is the modular acquisition or capture of an envelope glycoprotein (Env) which facilitates dissemination from its initial host cell. Here we present several examples of retroviruses for which envelope capture has been identified. Indeed, capture may explain the notable conservation of env sequences among otherwise phylogenetically distant retroviruses. In a recent example, sequence homologies reported between the env of the phylogenetically distant murine leukemia viruses (MLV) and human T cell leukemia viruses (HTLV) argue in favor of an env capture by the latter. Env acquisition can provide new adaptive properties to replication-competent viruses in addition to altering their host range. Also, the captured env can alter the spectrum of physiological affects of infection in new host cells and organisms. The elucidation of such envelope exchanges and properties thereof should contribute significantly to the clarification of retroviral phylogeny, insight into retroviral pathogenesis, and to the discovery of new retroviruses.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Vertebrate retroviruses; Human T cell leukemia viruses; Envelope capture

Retrotransposons and retroviruses

Genomic retroelements

Nearly half of the human genome can be recognized as derived from transposable genetic elements, most of which are retrotransposable (Lander et al., 2001; Ostertag and Kazazian, 2001; Smit, 1999). Retrotransposition is characterized by a step in which the transposon-derived RNA is reverse-transcribed into DNA before integration into the cell genome (for review see Links 1 and 2). Among the retroelements, viral retrotransposons comprise long terminal repeat (LTR) sequences at the 5’ and 3’ extremities surrounding gag-pol coding sequences (Fig. 1). Endogenous retroviruses (ERV) are viral retrotransposons which are present in the vertebrate host germ line as proviruses and, as such, are part of the host cell genomic patrimony (Link 2). A significant percentage of ERV is the vestige of ancient infections of the germ line by exogenous infectious retroviruses (Benit et al., 2001; Pavlicek et al., 2002). It has been hypothesized that replication-competent retroviruses have evolved from transposable genetic elements (Temin, 1980). The human genome contains numerous endogenous retroviruses, or HERV, distributed among at least 26 multigenic families, some comprising several hundred elements (Lower et al., 1996; Tristem, 2000) and occupying 5–8% of the genome (Lower et al., 1996; Ostertag and Kazazian, 2001; Smit, 1999). Although the majority of ERV sequences are found in the generally nontranscribed heterochromatic regions of the host genome, their direct influence in viral evolution (see below) and in gene expression has been well established (Robins and Samuelson, 1992; Samuelson et al., 1996).

The envelope brings out the retrovirus

Infectious retroviruses harbor a virus-encoded envelope glycoprotein, designated Env (Fig. 2, and see below), which is inserted into the viral envelope formed by the host cell plasma membrane. The viral Env confers properties that are essential for viral entry and the dissemination of retroviral particles (Hunter and Swanstrom, 1990). The initial steps of the retroviral infection cycle comprise assembly, budding, and release of infectious viral particles from the cell (see...
Link 3). The Env does not appear to be essential for particle assembly, as virus-like particles (VLP) can form in the absence of Env. Furthermore, extracellular budding of env-less VLP is observed in the context of artificially overexpressed Gag, suggesting that the env is not necessary for viral egress (Garoff et al., 1998). However, in naturally expressing cells, VLP accumulation driven by Gag occurs in the cytosol or via budding into the endoplasmic reticulum or the nucleus (Fig. 3A, and see Assembly and Maturation chapter of Link 4). Adding env to this process can considerably alter the topology of particle expression, for example, by favoring assembly at the plasma membrane or by directing polarized, basolateral assembly of viral particles in epithelial cells (Garoff et al., 1998; Owens et al., 1991) (Fig. 3B). Moreover, we recently showed that the envelope glycoprotein of murine leukemia virus (MLV) redirects Gag and viral RNA trafficking from lysosomes to the plasma membrane and extracellular release via endosomal vesicles (Basyuk et al., 2003). This influence of env on vesicle routing may constitute one of the initial selective advantages leading to the formation of env-containing retroviruses. Thus, the infectious properties of env-containing retroviruses may have developed through the cumulative acquisition of at least three env-dependent properties: (i) the ability to exit cells via endosomal membrane trafficking; (ii) a protective affect conferred by the cellular lipid bilayer to viral capsids that are released to extracellular environments; and (iii) the ability to disseminate to neighboring cells and organisms (Fig. 3). Splicing-dependent expression of env mRNA could have subsequently evolved from the large panel of alternative and cryptic splice sites that are present in the gag-pol sequences of simple retroviruses (Dejardin et al., 2000).

The colinear insertion of an env gene into the genome of retrotransposons thus appears to be a key step in the emergence of replication-competent infectious retroviruses (see Link 5). env capture by retrotransposons recapitulates the evolution of a sporadic trans-complementation system (Figs. 3A–B) into one of a constitutive cis-complementation (Fig. 3C). Nevertheless, viral Env’s conserve their capacity for trans-complementation, as illustrated by the phenomenon

Fig. 1. Schematic representation of retrotransposons and retroviruses. (A) Nonviral (i.e., non-LTR) retrotransposons, such as LINEs, containing a gag-like open reading frame (orf) and pol sequences comprising endonuclease, reverse transcriptase, and in some cases RnaseH encoding sequences; (B) viral retrotransposons (i.e., flanking, 5' and 3'LTR) such as Ty (of Saccharomyces cerevisiae), Copia (of D. melanogaster), and certain endogenous retroviral sequences (see text); (C) retrovirus harboring the three reading frames gag, pol, and env.

Fig. 2. Schematic representation of a virion and a retroviral Env. (Left) Env trimers anchored on the surface of a virion. (Right) Schematic representation of a murine leukemia virus (MLV) Env monomer composed of a receptor-binding-domain-containing SU and a transmembrane fusion protein TM.
of viral particle pseudotyping, in which the Env from one retrovirus is used by heterologous viruses during coinfection (Fig. 3D) (Sitbon et al., 1985). Pseudotyping following activation of endogenous retroviral envelope sequences may lead to extended spreading of replication-competent retroviruses or defective retroviral vectors (An et al., 2001; Bonham et al., 1997).

The evolution of retroviruses through env capture by retrotransposons does not preclude the coexisting reverse evolutive process, that is, the extinction of Env expression in infectious retroviruses. Such Env extinction has likely contributed to the stable insertion of exogenous retroviruses in the germ line of the infected organisms. While extinction by open reading frame mutations is commonly observed for HERV (Griffiths, 2001), the excision of an entire env gene remains more difficult to ascertain.

Env, the viral envelope glycoprotein

The Env of vertebrate retroviruses is encoded by a third reading frame of the retroviral genome, the env gene situated 3' of the viral gag and pol genes (Fig. 1). env mRNA is the product of the spliced viral genomic RNA (Fig. 3C) and its accumulation in the cytoplasm is dependent on cellular factors and virus-encoded proteins in the case of complex retroviruses (Cullen, 1991).

Despite their diversity, all retroviral Env-mediated viral entry comprises two principle steps: (i) binding to a cell surface receptor(s), preceding viral entry, and (ii) env-induced fusion of viral and cellular membranes (Weissenhorn et al., 1999). For vertebrate retroviruses, these two functions are performed by the entirely extracellular Env surface component (SU) and the membrane-anchored transmembrane component (TM), respectively. The env SU of vertebrate retroviruses contains receptor-binding determinants that affect viral tropism (Gallagher et al., 1995; Hunter and Swanstrom, 1990) and is one of the most variable regions of the viral genome, suggesting that it is the target of positive selective pressures (Pancino et al., 1994). The TM structure is highly conserved among the retroviruses (Fig. 2) (Benit et al., 2001; Kobe et al., 1999). Despite its characteristic variability, the SU harbors a modular organization as well as certain determinants that appear to be conserved among some highly divergent retroviruses (see below).

In addition to their role in viral infection, several Env domains influence other aspects of host physiology such as cellular membrane structure (Kozak et al., 2002), immune response (Benit et al., 2001; Wyatt and Sodroski, 1998), cell signaling (Rai et al., 2001), cell metabolism (unpublished observations), and proliferation (Kinet et al., 2002).
env capture

The emergence of infectious invertebrate retroviruses evidences the de novo acquisition or “capture” of env genes by retrotransposons. One of the best-documented examples of such capture is the Gypsy retrotransposon of Drosophila melanogaster (Pelisson et al., 2002; Terzian et al., 2001). Gypsy is in fact an infectious insect retrovirus (genus erantivirus) partly due to the capture of an env-like gene from baculovirus (double-stranded DNA insect virus) (Malik et al., 2000; Pearson and Rohrmann, 2002; Song et al., 1994). Further evidence of env capture by invertebrate retroviruses has been suggested. For example, the env of the nematode (Caenorhabditis elegans) retrovirus Cer shares homologies with a phleboviral fusion protein; and Tas, a retrovirus of Ascaris lumbricoides, appears to have captured the gB glycoprotein of an ancestral herpesvirus (Malik et al., 2000).

The capture of env-like genes as described above suggests that gene acquisition may occur via intragenomic recombination events after integration of retrotransposons into the genome of these large dsDNA viruses (Pearson and Rohrmann, 2002). Evidence in support of this capture mechanism is provided by the TED retrotransposon of the lepidopteran Trichoplusia ni, whose env gene is homologous to the baculovirus F gene, combined with the observation that integrated forms of TED have been found in the genome of the baculovirus ACNV (Pearson and Rohrmann, 2002). A similar mechanism may also explain the capture of herpesvirus gB sequences by an ancestral retroelement of Tas (Malik et al., 2000). Capture may not require the integration of retroelements into the genome of infecting viruses. However, according to the above examples, infection by large dsDNA viruses such as baculovirus or herpesvirus appears to provide conditions favorable to the sporadic capture of env by host retroelements—and the emergence of new infectious retroviruses.

env capture by vertebrate retroviruses

The first and most thoroughly described example of the de novo capture of a genetic element by an infectious retrovirus is the acquisition of the oncogene src by Rous sarcoma virus (Stehelin et al., 1976). However, de novo capture of an env gene by an initially env-less vertebrate viral retrotransposon remains to be demonstrated. In contrast to the scenario of env capture by invertebrate retrotransposons described above, vertebrate env capture has been described only in terms of the acquisition of heterologous env sequences among different endogenous or exogenous retroviruses. This is evidenced by the presence of chimeric retroviral sequences in a broad range of animals, including chickens, mice, cats, sheep, goats, monkeys, and humans. A recent example of this phenomenon is the emergence of a highly infectious subgroup J avian retrovirus which is at least partly due to the acquisition of a functional env, closely related to an endogenous sequence found in chickens (Denesvre et al., 2003).

Cat and mouse retroviruses

Infection of both mice and cats may lead to the emergence of new retroviruses through recombination between the incoming exogenous viruses and endogenous viral sequences. The mouse mink cell focus-forming (MCF) viruses, also designated polytropic MLV, were the first retroviruses identified as resulting from such de novo recombinations following MLV infection (Fischinger et al., 1975; Hartley and Rowe, 1976). Subsequent to infection by an ecotropic MLV, acquired mcf env sequences broaden the cellular tropism of the parental ecotropic MLV through the acquisition of heterologous receptor-binding determinants in the env SU (Fan, 1997). Generation of new env-recombinant viruses also occurs upon infection of cats with feline leukemia viruses (FeLV). Thus, infection with the FeLV-A subtype can result in the generation of FeLV-B subtypes with markedly modified tropism (Stewart et al., 1986).

The acquisition of heterologous viral sequences, however, is not limited to env; and it is important to note that recombination among or with mcf sequences involves other regions of the MLV genome as well (Evans and Cloyd, 1985). Irrespective of the region of the viral genome, it is generally accepted that recombination between endogenous mcf and exogenous MLV sequences occurs after coencapsidation of heterologous genomic RNA into virus particles (Fig. 3D) (Katz and Skalka, 1990). The production and characteristics of the resulting recombinant viruses depend upon the subspecies or strain of mouse and the type of exogenous infectious virus inoculated (Chesebro et al., 1983; Lavignon et al., 1997). It is also noteworthy that the initial dissemination of these recombinant viruses depends largely on pseudotyping with env from the infecting, exogenous MLV (Sitbon et al., 1985) (Fig. 3D).

Retroviral chimerism and predatory chains

Interspecies retroviral env capture is exemplified by the remarkable series of acquisitions that led to the formation of the endogenous feline retrovirus RD114 (Fig. 4). Moreover, emergence of this chimeric virus evokes a scenario of infection during predatory exchanges among feline and primate species. As presented in Fig. 4, at least four distinct retroviruses are implicated in the emergence of RD114: baboon endogenous retrovirus (BaEV), Papio cynocephalus endogenous virus (PeEV), simian endogenous retrovirus (SERV), and Felis catus endogenous virus (FeCV). BaEV is a replication-competent endogenous retrovirus (Bennie niste et al., 1974) found in several simian species including mandrills, baboons, mangabeys, and African green monkeys (van der Kuy et al., 1995). BaEV is a chimeric retrovirus composed of type C gag and pol sequences.
originating from another endogenous baboon retrovirus, PcEV, and the env of a type D simian endogenous retrovirus, SERV (van der Kuyl et al., 1999). The emergence of RD114 is thus explained by at least two env acquisition events: the capture of the SERV env by PcEV, to produce BaEV; and the capture of the BaEV env by FcEV, that resulted in the emergence of RD114 (van der Kuyl et al., 1999) (Fig. 4).

The type C gibbon ape leukemia virus (GaLV) may reveal another cascade of env captures resulting from murine–feline–primate predatory relationships. According to its genomic organization and sequence, GaLV is highly related to MLV, and it has been suggested that the origin of GaLV may be traced to an ancient infection of monkeys by a xenotropic murine retrovirus (Wolgamot et al., 1998 and references therein). However, the Env SU of GaLV binds the same cell surface receptor as the FeLV-B feline leukemia virus (Takeuchi et al., 1992). Based on these data, one possible scenario for the emergence of GaLV could involve a successive transmission of viral sequences. A murine retrovirus exhibiting an interspecies host range emerged (possibly an MLV-like virus generated through MCF sequence acquisition as described above) and infected feline predators, resulting in the emergence of infectious feline retroviruses such as FeLV-A (see above). Subsequently, env recombinant retroviruses, such as FeLV-B, may have then been generated and disseminated from their feline host to predatory primates (or through the inverse predatory relationship) to form GaLV.

In addition to scenarios implicating infection after blood exchange during predation, examples of chimeric retroviruses involving animals with no apparent predatory relationship have also been described. For example, the ovine or caprine retrovirus Jaagsiekte retrovirus (JSRV), the etiologic agent of ovine pulmonary adenocarcinoma, is a type D virus with an env derived from a type B virus (York et al., 1992). Related type D or B chimeric viral sequences have been described not only in sheep and goats, but also in other ungulates, including the vast majority of the genus ovis and genus capra, as well as in domestic cattle, suggesting an ancient env capture and interspecies dissemination of chimeric viruses (Hecht et al., 1996).

Highly divergent human and murine retroviruses share homologous env.

Recent comparative analysis of the env of the human T-cell leukemia or lymphoma virus (HTLV) and the murine leukemia virus (MLV) demonstrated that these phylogenetically highly divergent viruses (Tristem, 2000) share remarkable homologies of modular organization and motifs within the env SU that are usually particularly variable even among viruses of the same species. Thus, we showed that the amino acid sequence LLTLVQ in the env SU of HTLV-1
and LLNLVQ of Friend-MLV delimits homologous and interchangeable functional domains (Fig. 5 and Kim et al., 2000, and submitted for publication).

Such SU close homology and motif conservations within the genomes of otherwise highly divergent retroviruses suggest a common ancestral origin of the HTLV and MLV env and are compatible with a scenario of env capture by HTLV (Kim et al., 2000, 2003, and submitted for publication). The routes of interspecies transmission and the acquisition of new sequences by ancestors to these viruses may be deduced by precise phylogenetic analyses of env sequences of HTLV, its simian homologues STLV, and other simian, feline, and murine C-type retroviruses. Such analyses and the identification of yet to be identified HTLV-related endogenous sequences, env or other, in human genomes may help elucidate an apparently ubiquitous phenomenon of recombinatory sequence acquisition by retroviruses.

The origin(s) of env?

The acquisition and recombinatorial exchange of heterologous env is well documented. However, the origin of the env itself remains enigmatic. The general structure of the SU receptor binding domain associates it with the immunoglobulin superfamily (Fass et al., 1997), and the membrane fusion capacity of env TM naturally implies a common origin with cellular fusion proteins and membrane proteins involved in membrane trafficking (Poumbourios et al., 1999; Weissenhorn et al., 1999). As such, one possible scenario suggests that env genes could have formed de novo by stochastic recombination of distinct protein coding sequences (Lerat and Capy, 1999; Lerat et al., 1999), possibly via mechanisms such as exon shuffling (Boeke and Pickeral, 1999; Pavlicek et al., 2002). The remarkable conservation of TM among highly divergent viruses, in contrast to the SU, suggests the possibility of modular acquisition of SU and TM as distinct domains (Benit et al., 2001; Kobe et al., 1999). Thus ancestral TM-harboring highly conserved functional domains such as the fusion peptide, immunodominant and immunosuppressive peptides (Gallaher et al., 1995; Pancino et al., 1994), and the highly conserved coiled-coiled extracellular domain (Kobe et al., 1999) may have acquired SU of various origins. This scenario is supported by experiments demonstrating that functional, truncated receptor-binding domains of the SU can be expressed independent of other env domains (Battini et al., 1995; Kim et al., submitted for publication; Manel et al., 2002), and that env function can be reconstituted by complementing defective env with these independently expressed, soluble SU domains (Barnett and Cunningham, 2001; Lavillette et al., 2000).

In addition to the role of Env in driving viral infection, endogenous env sequences can also exert cellular functions (Blond et al., 2000; Britt et al., 1984; Frendo et al., 2003; Kattstrom et al., 1989; Mi et al., 2000) or confer either resistance or increased susceptibility to infection and pathogenesis (Anderson et al., 2000; Chesebro et al., 1983; Ikeda and Sugimura, 1989). A major focus of current research is to determine causal relationships between endogenous retroviral sequences and idiopathic pathologies including certain

---

**Fig. 5.** Organization and sequence homologies of the SU of MLV and HTLV env (Kim et al., 2000, and submitted for publication). Key homologous regions are indicated for prototype polytropic (P-MLV), xenotropic (X-MLV), amphotropic (A-MLV), Friend ecotropic (F-MLV) murine leukemia virus, and HTLV-1 and -2. Variable amino acid residues among the indicated blocks of homology are represented by dashed lines. The proline residues (P) of the “proline rich region” are shown and the other amino acid residues within this domain are represented by dashed lines.
leukemias, and neurodegenerative and autoimmune diseases (Lower, 1999; Perron and Seigneurin, 1999; Portis, 2002; Power, 2001).

Conclusion

Retroviruses and retroelements in general constitute a continuously evolving catalog of sequences driving the production of new viruses. Retroviral envelope “capture” appears to be a recurrent phenomenon in the generation of new infectious retroviral species. At least three general mechanisms can be at the origin of such a new insertion in a retroviral genome: (i) recombination after coencapsidation of heterologous retroviral genomes (Katz and Skalka, 1990), or cellular mRNA as it is generally believed for oncogene “capture”; (ii) intragenomic exon shuffling of retroelements (Boeke and Pickeral, 1999; Pavlicek et al., 2002); (iii) recombination after integration of retrotransposons or retroviruses into the genome of large dsDNA viruses such as herpes or baculovirus (Isoft et al., 1992; Malik et al., 2000; Pearson and Rohmann, 2002).

These recombination events illustrate the permanent co-evolution of retrotransposons and retroviruses with their host organisms. As a key element in the dissemination of retroelements to new hosts, env capture and the consequent dissemination of new retroviruses may be considered a deleterious event as generally perceived; however, it may also be considered a fundamental component in the evolution of genomes.

Acknowledgments

We thank E. Bertrand, L. Bénit, A. Bucheton, C. Denesvre, A. Pelisson, P. Sonigo, N. Taylor, and the members of our laboratory for stimulating discussion regarding the manuscript. This work was supported by grants from the Fondation de France, AFM, Association pour la Recherche contre le Cancer (ARC), Agence Nationale pour la Recherche contre le SIDA (ANRS), and the Fondation Philippe. FJK was supported by an award from the Fondation Philippe and successive fellowships from ARC, Fondation de France, and AFM; NM was supported by a graduate student fellowship from the MRT; JLB and MS are supported by INSERM.

References


dogenous retrovirus families by phylogenetic screening of the human genome mapping project database. J. Virol. 74 (8), 3715–3730.
van der Kuyl, A.C., Dekker, J.T., Goudsmit, J., 1995. Full-length proviruses of baboon endogenous virus (BaEV) and dispersed BaEV reverse transcriptase retroelements in the genome of baboon species. J. Virol. 69 (9), 5917–5924.
van der Kuyl, A.C., Dekker, J.T., Goudsmit, J., 1999. Discovery of a new endogenous type C retrovirus (FeEV) in cats: evidence for RD-114 being an FeEV(Gag-Pol)/baboon endogenous virus BaEV(env) recombinant. J. Virol. 73 (10), 7994–8002.

References Internet links