Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements

Marie Dewannieux, Francis Harper, Aurélien Richaud, Claire Letzelter, David Ribet, Gérard Pierron and Thierry Heidmann

Genome Res. published online Oct 31, 2006; Access the most recent version at doi:10.1101/gr.5565706

Supplementary data

"Supplemental Research Data"
http://www.genome.org/cgi/content/full/gr.5565706/DC1

P<P

Published online October 31, 2006 in advance of the print journal.

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here

Notes

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to Genome Research go to:
http://www.genome.org/subscriptions/

© 2006 Cold Spring Harbor Laboratory Press
Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements

Marie Dewannieux,1,3 Francis Harper,2,4 Aurélien Richaud,1,4 Claire Letzelter,1 David Ribet,1 Gérard Pierron,2 and Thierry Heidmann1,5

1Unité des Rétrovirus Endogènes et Éléments Rétroïdes des Eucaryotes Supérieurs, UMR 8122 CNRS, Institut Gustave Roussy, 94805 Villejuif Cedex, France; 2Laboratoire de Réplication de l’ADN et Ultrastructure du Noyau, UPR1983 Institut André Lwoff, 94801 Villejuif Cedex, France

Human Endogenous Retroviruses are expected to be the remnants of ancestral infections of primates by active retroviruses that have thereafter been transmitted in a Mendelian fashion. Here, we derived in silico the sequence of the putative ancestral “progenitor” element of one of the most recently amplified family—the HERV-K family—and constructed it. This element, Phoenix, produces viral particles that disclose all of the structural and functional properties of a bona-fide retrovirus, can infect mammalian, including human, cells, and integrate with the exact signature of the presently found endogenous HERV-K progeny. We also show that this element amplifies via an extracellular pathway involving reinfection, at variance with the non-LTR-retrotransposons (LINEs SINEs) or LTR-retrotransposons, thus recapitulating ex vivo the molecular events responsible for its dissemination in the host genomes. We also show that in vitro recombinations among present-day human HERV-K loci can similarly generate functional HERV-K elements, indicating that human cells still have the potential to produce infectious retroviruses.

[Supplemental material is available online at www.genome.org.]

Nearly 8% of the human genome is composed of sequences of retroviral origin. Most of them are degenerate, either due to recombination between the two provirus LTRs or to mutations interrupting the retroviral ORFs. The env gene seems to be best conserved, with 18 intact genes retaining a full coding capacity in the human genome (Benit et al. 2001; de Parseval et al. 2003; Villesen et al. 2004), possibly because of its potential role in human physiology. The HERV-K(HML2) family of endogenous retroviruses is an exception to this general rule, since some copies still contain complete ORFs for the other retroviral genes (Löwer et al. 1996; for review, see Bannert and Kurth 2004). This family includes the most recently amplified endogenous retroviruses, most of which have integrated into the genome <5 million years ago, with a few insertions showing polymorphism within the human population (Steinhuber et al. 1995; Medstrand and Mager 1998; Barbulescu et al. 1999; Turner et al. 2001; Hughes and Coffin 2004; Belshaw et al. 2005). Some of these recently integrated proviruses are responsible for the synthesis of retroviral particles that can be observed in teratocarcinoma and melanoma-derived cell lines (Boller et al. 1993; Löwer et al. 1993; Bieda et al. 2001; Muster et al. 2003; Buscher et al. 2005), and possibly in human placenta (Kalter et al. 1973; Dirksen and Levy 1977; Wilkinson et al. 1994). Because of this “activity,” the HERV-K(HML2) family has been the subject of numerous studies in the past years, with the description of alleles with nearly intact proviruses and complete coding capacity (Mayer et al. 1999; Reus et al. 2001; Turner et al. 2001). Despite these efforts, no functional provirus able to produce infectious particles has yet been described. Even if several loci containing complete ORFs have been identified, and in spite of the availability of the complete sequence of the human genome, the search for a functional provirus is hampered by the significant polymorphism that can be detected at each HERV-K(HML2) locus within the human population. To overcome this limitation, we used a reverse strategy and generated a consensus HERV-K(HML2) provirus, thus “resuscitating” Phoenix, the likely progenitor of the last, humanspecific HERV-K(HML2) amplification burst. This allowed a definite characterization of the structure and life cycle of an “ancestral” retroviral element and recapitulation of the molecular events responsible for its amplification.

Results

Phoenix, the ancestral HERV-K(HML2) retrovirus

To construct a consensus HERV-K(HML2) provirus, we assembled all of the complete copies of the 9.4-kb proviruses that are human specific (excluding those with the 292-nt deletion at the beginning of the env gene) and aligned their nucleotide sequence to generate the consensus in silico, taking for each position the most frequent nucleotide. The resulting provirus sequence contains, as expected, ORFs for all of the HERV-K(HML2)-encoded proteins (Gag, Pro, Pol, Env, and the accessory Rec protein), with gag, pro, and pol separated by ~1 frameshifts. Noteworthily, this consensus provirus is distinct from each of the sequences used to generate it, with at least 20 amino acid changes on the overall sequences (Fig. 1).

Based on this in silico reconstruction, we then generated a
molecular clone corresponding to the consensus DNA proviral sequence that we named *Phoenix*, using the related K108 and K109 proviruses (which we had previously cloned from a commercial human BAC library) as a backbone, and a commercial kit to introduce single nucleotide mutations at each position required to match the consensus. As the HERV-K(HML2) LTRs are not functional in every cell line (Ruda et al. 2004; Lavie et al. 2005), we also replaced the U3 part of the 5′ LTR by the CMV promoter, with its start site positioned so as to conserve the expected nucleotide sequence of the native retroviral transcript. We also introduced a cloning site downstream of the env gene into the noncoding U3 region of the 3′ LTR to possibly tag (with the neo gene) the retroviral transcript without altering its coding capacity.

*Phoenix codes for bona-fide retroviral particles*

In a first assay, we introduced the *Phoenix* expression vector into human 293T cells, and looked for retroviral particles by transmission electron microscopy. As illustrated in Figure 2, A and B, this resulted in the synthesis of particles budding from the plasma membrane (not observed with a control vector), as classically observed with γ- (type-C) or δ- (HTLV) retroviruses, or even lentiviruses (HIV). No preassembled particles were detected within the cytoplasm, although HERV-K(HML2) elements are more closely related to β- (type-B/D) retroviruses based on their pol gene (but this is consistent with viral particle morphology and site of assembly being essentially dependent on the gag rather than the pol gene). In the extracellular space, two classes of particles could be observed, either with a hollow interior surrounded by a dense ring of stain corresponding to immature particles or with a condensed hexagonal core corresponding most probably to the mature form of the virus (Fig. 2C), in which the polyproteins are cleaved into the final products. Consistently, a provirus with a mutation introduced into the active site of the protease (Pro) also produces particles, but all of them disclose the immature morphology (Fig. 2F). The particles are surrounded by prominent spikes (Fig. 2D) that can be marked in immunoelectron microscopy experiments with an antibody directed against the SU subunit of the HERV-K(HML2) envelope protein (Fig. 2E). As expected, these spikes are absent when the provirus contains a premature Stop codon in the env gene (Fig. 2G). Altogether, these features suggest that *Phoenix* can direct the synthesis of complete retroviral particles.

As a confirmation, we performed immunoblot analyses of the 293T cells transfected with the proviral vector. As illustrated in Figure 3A, *Phoenix*, as well as its pro mutant, promotes the synthesis of a Gag precursor of ∼80 kDa in cell lysates, as well as some cleaved products (not observed with the pro mutant). Cleavage products could also be observed in the supernatant of...
Figure 2. Electron microscopy of the viral-like particles generated by the Phoenix provirus. Human 293T cells were transfected with an expression vector for Phoenix (A–E), or mutants (F–G), and observed 48 h post-transfection. (A) Low magnification of particles assembled at the cell membrane. (B) Representative image of particles budding from the plasma membrane. (C) High magnification of two particles, one of which (bottom) discloses a mature (M) morphology with a condensed core, while the other appears to be still immature (IM) with two dark peripheral rings surrounding an electron-lucent core. (D) High magnification of a particle with prominent spikes, corresponding to the Env protein. (E) Image of a particle after labeling with an antibody specific for the HERV-K envelope protein and a secondary antibody linked to gold beads, obtained by immuno-electron microscopy. Quantification of the labeling on 11 independent fields demonstrates association of the gold beads with the viral particles: 307 ± 121 gold beads/µm² for the viral particles, versus 4.9 ± 3.2 and 1.1 ± 1.5 gold beads/µm² for the cytoplasm and particle-free extracellular space, respectively (P < 0.001 between viral particles and any of the two other compartments, Student’s t-test). (F) Image of representative particles obtained after transfection with an expression vector for the Phoenix pro mutant. All of them disclosed an immature morphology (41 of 41 identified “free” particles, i.e., no more in the budding process, for the pro mutant, vs. 15 of 37 for Phoenix WT). (G) High magnification of a particle obtained after transfection with an expression vector for the Phoenix env mutant. The membrane surrounding the particle is clearly detectable, without any spike. Scale bars: (A): 200 nm, (B–G):100 nm.

The transfected cells, except for the pro mutant that only expresses the Gag precursor. Similarly, using an antibody directed against the SU component of the HERV-K(HML2) envelope protein, we detected, in the supernatants, a protein of a size (55 kDa) consistent with that of the processed SU subunit, but we were unable to detect the Env protein in cell lysates, most probably because of a too-low expression level. Notwithstanding, expression of the envelope protein could be observed by immunofluorescence analysis using HeLa and 293T cells transiently transfected with Phoenix (Fig. 3B). In these cells, it can further be observed that Gag and Env colocalize to some common subcellular domains including, as expected, the plasma membrane.

Phoenix is an infectious retrovirus

To get insight into the functionality of this ancestral virus, we then looked for a reverse-transcriptase activity in the supernatant of the transfected cells. Using the previously described sensitive product-enhanced reverse transcriptase (PERT) assay (Pyra et al. 1994; see Methods), an RT activity could indeed be detected in the supernatants of cells transfected with Phoenix, not observed with a provirus in which we had introduced a mutation within the catalytic site of the RT domain (Fig. 3C). To further study the infectivity potential of Phoenix-derived particles, we then infected a panel of mammalian cell lines with the supernatant of 293T cells transfected with a neo-marked Phoenix provirus, and subjected them to G418 selection. G418§ clones were obtained for several target cells including hamster BHK21 cells, feline G355.5 cells, and, noteworthy, the human SH-SYSY and 293 cells, indicating that Phoenix is fully functional and infectious (Fig. 4A). In the same assay, the human HeLa cells and the murine WOP cells seemed to be resistant to infection, most probably due to the absence of the appropriate receptor for the Env protein at the cell surface. We then ensured that the G418§ clones that were obtained are the result of bona-fide infection. Their occurrence actually depends on the presence of intact gag, pro, pol, and env genes, since the inactivation of any of the encoded proteins renders Phoenix noninfectious without decreasing the amount of viral particles produced, except for the gag mutant, as assessed by Western Blot (Fig. 4A; data not shown). In these experiments, the apparent titer of Phoenix is quite low even for the cell lines permissive for infection. A quantitative assay by real-time PCR on the genomic DNA from BHK21 cells infected with particles generated by Phoenix, marked, or not, by the neo gene (see Methods), indicated that the presence of neo within the provirus contributes only to a limited extent to this low titer. A possible explanation is that infection with Phoenix occurs more efficiently via cell–cell interaction, as observed for HTLV and in some cases HIV (see Discussion), consistent with some images obtained by electron microscopy where budding particles appear to be directly captured by a recipient cell (Fig. 4B).

Next, we characterized further the G418§ clones that we isolated and determined the provirus insertion sites using an inverse PCR strategy. Identification of the target loci of these new insertions was further confirmed by direct PCR reactions between the identified flanking sequences and the neo gene of the marked proviruses (see Methods). In all cases, we found completeLTRs with short target-site duplications (TSDs) bordering the newly inserted proviruses (Fig. 5), thus confirming that the infection process is canonical. Rather surprisingly, the TSDs were of a variable length (one of 5 bp and two of 6 bp), which was unexpected since well-characterized retroviral integrases yield TSDs of a fixed length (for review, see Brown 1997). However, our data are consistent with the structures observed for the HERV-K(HML2) endogenous copies of the human genome, since, even if the most common length of the endogenous HERV-K proviruses TSD is 6 bp, we could identify, both in the literature (Barbulescu et al. 1999; Turner et al. 2001) and databases (our unpublished observation), a few elements (three of 15 full-length proviruses) bordered by perfect TSD of only 5 bp (see the endogenous K115
provirus in the lower part of Fig. 5 for an illustration). Finally, we mapped the three Phoenix insertion sites using the UCSC Genome Browser and found that integration occurred within (for two of three) or close to (<20 kb, for one of three) genes. Although the number of target sites that we characterized is low, and thus cannot reveal subtle insertion site preferences, the identified loci are consistent with the preferences observed for classical retroviruses that favor insertion in transcribed regions (for review, see Bushman 2003). Interestingly, the characterization of the insertion sites of 11 (nearly) full-length HERV-K(HML2) endogenous elements gave a rather different result, with five of them found far from known human genes (>100 kb on each side), four being in the vicinity (<20 kb) of genes, and only two inserted within genes (M. Dewannieux, unpubl.). This discrepancy is probably due to the strong counter-selection that should operate in vivo against insertions within genes, which are most probably deleterious for the proper expression of the targeted genes.

**Figure 3.** Characterization of Phoenix. (A) Immunoblot analysis of the Phoenix products present in the lysates of transiently transfected 293T cells, or in the concentrated supernatant. Cells were transfected with (1) a control plasmid (pCMV-β), (2) an expression vector for Phoenix, or (3) an expression vector for a pro mutant. The membranes were hybridized with an antibody directed against the Gag protein (left) or the Env protein (right). (B) Immunofluorescence study of 293T cells transfected with Phoenix. Gag is marked in red, Env in green, and the nuclei are stained in blue. Note the colocalization of the two proteins, especially at the membrane that separates the two adjacent marked cells. (C) Detection of RT activity in the concentrated supernatant of 293T cells transfected with Phoenix and a series of proviral and control vectors. Concentrated virions were used as a source of RT activity to reverse transcribe synthetic MS2 RNA. Presence of cDNA was revealed via a PCR amplification using appropriate primers. (1) pCMV-β; (2) Phoenix; (3) Phoenix mutated in RT; (4) chimera K109-K115; (5) chimera K109-K113; (6) chimera K109-K108; (7) MLV core proteins.

**HERV-K(HML2) provirus amplification via reinfection**

Next, to get an insight into the possible physiological mechanism responsible for the amplification of the HERV-K(HML2) family of endogenous retroviruses, we analyzed the relative efficiency of intracellular retrotransposition versus reinfection by Phoenix. To do so, we marked Phoenix with a specific neo indicator gene (neoTNEF), in which neo becomes active only after the marked element has been transcribed, reverse-transcribed, and integrated (Esnault et al. 2002). The marked element was introduced by transfection into various cell lines, permissive or not, for infec-

**Figure 4.** Infectious properties of Phoenix encoded particles. (A) Infectivity of the HERV-K retroviral particles. The apparent viral titer observed with the neo-marked HERV-K(HML2) provirus is indicated for a panel of target cell lines (the quantitative results for the 293 cell line is not shown since their restricted adherence impaired the proper quantification of individual G418R clones), together with that obtained with a mutant for the RT domain tested in the same conditions (no clone observed); data are the results from three to eight independent experiments, with the standard deviation indicated. The other mutants (gag, pro, and env) gave the same, negative, results when assayed, with no clone observed in three independent experiments (difference measured between Phoenix WT and mutants significant with \( P < 0.001 \) for G355.5 and BHK21 cells and \( P < 0.01 \) for SH-SYSY cells, Student’s t-test). The rationale of the assay is schematized at top. (B) Infection of cells by Phoenix particles. On the left, two mature particles appear to interact with the cell membrane of the closest cell, with a thickening of the membrane at the exact point of interaction with the particle. In the middle and on the right, two images suggestive of a cell-to-cell infection are presented, with a particle still budding from its progenitor cell already in contact with the neighboring cell.
A functional HERV-K human endogenous retrovirus

**Figure 5.** Structure of three de novo integration sites of *Phoenix* provirus, and comparison with the structure of natural HERV-K proviruses. Insertion sites were performed using individual clones from human SH-SY5Y cells after infection and G418 selection. A provirus insertion and the corresponding empty site are schematized at the top. The sequences of three characterized *Phoenix* de novo insertions are shown below, with the flanking DNA in uppercase and the proviral sequences in lowercase; target-site duplications of 5/6 bp (TSD, yellow) are found in all cases, associated with full-length LTRs. For comparison, the corresponding structures of two resident HERV-K proviruses with a polymorphic insertion are presented in the lower part, one with a 5-bp TSD and one with a 6-bp TSD.

**Figure 6.** Assay for *Phoenix* "retrotransposition." (Top) Rationale of the assay. *Phoenix* was marked with the neoR indicator gene for retrotransposition (Esnault et al. 2002), in which neo becomes active only after a complete retrotransposition cycle (i.e., transcription, reverse transcription, and integration), and the marked element was introduced by transfection into SH-SYSY human cells permissive for *Phoenix* infection. The cells were then amplified and subjected to G418 selection for 2 wk; the number of G418R clones (several of which were assayed by PCR to demonstrate splicing out of the intron within the neo indicator gene) yielded the frequency of retrotransposition of the marked elements. (Bottom) "Retrotransposition" is dependent on the presence of a functional env gene, either present within the native *Phoenix* provirus (*Phoenix* WT) or added in trans to an env-mutated *Phoenix* provirus (*Phoenix Stop Env*) by cotransfection of the cells with an env expression vector (CMV Env). The results presented here correspond to three independent experiments, performed with 5–10 x 10⁶ cells, and are given as relative transposition frequencies as compared to neoR-marked *Phoenix* (whose transposition frequency is 2 x 10⁻⁵ clone/seeded cell under these conditions).
pseudotype experiments, at variance with the K108 env gene) (Dewannieux et al. 2005). Therefore we generated a “triple chimera” by replacing the env region in the above construct with that from K108 and assayed it for infection. With this element, specific integration of the provirus (i.e., not observed with Phoenix mutated in the RT domain used as a control) could be detected in the BHK21 cells by quantitative PCR (see Methods) at a level only five- to 10-fold lower than that observed with Phoenix (Fig. 7). This indicates that, as observed in AKR mice where a multistep recombination process reproducibly generates new infectious retroviruses (Stoye et al. 1991; Boeke and Stoye 1997), some individuals could reconstitute functional HERV-K(HML2) proviruses either via a three-fragment recombination event, or more likely via a single recombination event between 2 HERV-K copies (e.g., K109 and K115) with complementation in trans by an intact envelope protein (e.g., K108). Recombination events have indeed occurred frequently between HERV-K(HML2) proviruses (Hughes and Coffin 2005). In this way, or even using functional (but still uncharacterized) alleles that might be present in some individuals, the HERV-K(HML2) family could repeat its life cycle and regenerate functional elements that could, in turn, be the initiators of amplification events to come.

Discussion

In the present investigation, the reconstruction of an active HERV-K(HML2) provirus gives access de facto to the first identified human retrovirus of endogenous origin that can consequently be characterized at both the structural and functional level. Actually, following “endogenization,” which transformed its retroviral status into that of a host genomic locus and thus dramatically reduced its evolutionary rate, this element has been frozen in time in its ancestral state, which we have recovered here. The essential attributes of Phoenix are those of a type-C retrovirus, with assembly and budding at the cell membrane, release in the cell supernatant of particles disclosing protruding spikes that can be unambiguously identified as envelope proteins, morphological maturation of the core structure upon viral protease processing, and finally, infectivity of the particles for a range of target cells. Furthermore, analysis of the integrated proviruses associated with infection discloses the specific structural features of retroviral integration, with precise opening of the target sites, insertion of complete proviral structures with reconstituted LTR, and target-site duplications of a few base pairs. Of note, the target-site duplications are not of a fixed length, but vary from 5 to 6 bp. Although this is not common for present-day infectious retroviruses, where the target site duplications are of fixed lengths (for review, see Brown 1997), this peculiarity is also observed for the endogenous proviral HERV-K loci, as illustrated in Figure 5. Possibly, this unusual feature might reflect the “age” of Phoenix and be a specific characteristic of its ancestral integrase.

Another particular trait of Phoenix is its low infectious titer. One possible interpretation could be that this results from its sensitivity to one or more of the multiple cellular defense mechanisms that have evolved to counteract the amplification of retroelements, thus guarding their hosts against the potentially harmful consequences of retrotransposition. The replication of Phoenix could be limited by RNA interference-mediated gene silencing, as well as by the action of a variety of host gene-encoded restriction factors, such as members of the AID/APOBEC cytosine deaminase family or of the TRIM protein family (for review, see Goff 2004; Harris and Liddament 2004). Another possible interpretation could be that the still not-identified receptor for this ancestral retrovirus has diverged since its entry into the primate genome and is no more adapted to Phoenix, and could even have been positively selected for this divergence for the sake of the host. However, a more likely interpretation, consistent with the frequently observed images shown in Figure 4B, is that its mode of transmission is more closely related to that of retroviruses such as human HTLV, where virions are poorly infectious and spread essentially via cell–cell contacts, most probably at the level of dedicated “virolological synapses” (for review, see Jolly and Sattentau 2004). Similar structures might exist for Phoenix, also consistent with its relative phylogenetic proximity to delta- (HTLV) and lentiviruses (HIV) viruses (Triboulet 2000; Benet et al. 2001), and should actually be particularly well suited to optimize cell reinfecction and proviral amplification. As a matter of fact, despite its low infectious titer, another important and related outcome of the present investigation is the demonstration that past amplification of HERV-K has most probably involved an extracellular pathway and cell reinfecion: Indeed, using an appropriate retrotransposition assay, we demonstrate that amplification requires the presence of a functional env gene, and is only observed with cells that are permissive for infection. Accordingly, the mode of amplification of Phoenix is clearly distinct from that prevailing for most retroelements naturally present in eukaryote genomes, which amplify via a strictly intracellular pathway. These elements include the highly reiterated LINE and SINE sequences, which are non-LTR retrotransposons that replicate also as human HTLV, where virions are poorly infectious and spread essentially via cell–cell contacts, most probably at the level of dedicated “virolological synapses” (for review, see Jolly and Sattentau 2004). Similar structures might exist for Phoenix, also consistent with its relative phylogenetic proximity to delta- (HTLV) and lentiviruses (HIV) viruses (Triboulet 2000; Benet et al. 2001), and should actually be particularly well suited to optimize cell reinfecction and proviral amplification. As a matter of fact, despite its low infectious titer, another important and related outcome of the present investigation is the demonstration that past amplification of HERV-K has most probably involved an extracellular pathway and cell reinfecion: Indeed, using an appropriate retrotransposition assay, we demonstrate that amplification requires the presence of a functional env gene, and is only observed with cells that are permissive for infection. Accordingly, the mode of amplification of Phoenix is clearly distinct from that prevailing for most retroelements naturally present in eukaryote genomes, which amplify via a strictly intracellular pathway. These elements include the highly reiterated LINE and SINE sequences, which are non-LTR retrotransposons that replicate also by a copy-and-paste mechanism but are not infectious, and the classical LTR retrotransposons, such as the yeast Ty1 and the Drosophila copia elements, which are closely related to retroviruses, of which they are most probably the evolutionary precursors but do not possess an env gene and transpose intracellularly (for review, see Boeke and Stoye 1997; Kazazian Jr. 2004; Dewannieux and Heidmann 2005). In this respect, the HERV-K family of elements has a distinct and specific mode of amplification, which very likely also applies to most of the so-called “endogenous retroviruses,” and Phoenix, being actually the first infectious
HERV available to date, can therefore be considered as a model element to characterize the life cycle of these invaders of all mammalian genomes.

Finally, the present active cloned element provides a potent tool to appraise the role of elements of the HERV-K family in a range of diseases where related particles and viral proteins have been detected. These include essentially, but not exclusively, human tumors such as germ-line tumors and melanomas, where such elements have been suggested to play a role (Boller et al. 1993; Löwer et al. 1993, 1996; Bieda et al. 2001; Muster et al. 2003; Buscher et al. 2005) (for review, see Bannert and Kurth 2004). Taking into account their retroviral origin, firmly established by the present investigation, it is indeed legitimate to hypothesize that HERV-K elements have conserved some of the pathological potency of infectious retroviruses, which could arise upon their transcriptional activation. Clearly, most present-day infectious retroviruses are pathogenic and mediate such effects via both insertional mutagenesis and the activity of the viral proteins they encode. The present study clearly shows that Pho- nix is a complex retrovirus that can insert into new genomic locations, and that most probably, human cells can generate active retroviruses by recombination between genomic loci. As shown in the Results section, in vitro recombination between cloned HERV-K copies indeed generate a functional element (and active when further placed under the control of the strong CMV promoter) and it is likely, taking into account the polymorphism of such elements in the human population, that some alleles might still be functional and even become fully active in some definite circumstances. Search for such alleles, possibly among individuals with specific tumors where HERV-K elements are ex- pressed, could perhaps help in identifying new tumors that might still be functional and even become fully active in some definite circumstances. Search for such alleles, possibly among individuals with specific tumors where HERV-K elements are expressed and viral particles detected, should now be performed using Phoenix as a functional “reference” element together with the presently devised infectivity assay.

Altogether, the present results have allowed the structural and functional characterization of the oldest active retrovirus isolated to date and the unraveling of the molecular scenario responsible for its successful amplification in primate genomes. They also provide tools for elucidating the role of its “genomic offspring” in human diseases, in particular in some tumors that have primarily been qualified as of genetic origin, but could ult- imately turn out to have a bona-fide retroviral basis. In this respect, this study also suggests that more appropriate safety con- ditions should be considered for the manipulation of such tissues and tumor cell lines.

Methods

In silico engineering and in vitro reconstruction of the HERV-K(HML2) consensus provirus

To generate the consensus HERV-K(HML2) provirus sequence, we took the nucleotide sequences of the nine full-length, human-specific, 9.4-kb class proviruses that can be found in sequence databases and publications and aligned them using ClustalW to generate the consensus sequence (see Supplemental Fig. 1; Supplemental Table 1). To obtain a backbone for the Phoenix provirus expression vector, we subcloned the K108 and K109 proviruses by digesting commercial BAC clones RP11-152J15 (BACPAC Resources), and generated chimeric con- structs to get closer to the consensus sequence. Fragments from these chimeric constructs were placed in PALTER-1 plasmid (Promega) and subjected to successive rounds of site-directed muta- generation using the Altered Sites II in vitro Mutagenesis System (Promega) and appropriate primers. When all expected muta- tions had been introduced, the Phoenix provirus was reassembled in the pBluescript plasmid and we checked the whole construct by sequencing. To generate the CMV expression vector, the CMV promoter region was isolated from pCMV6 (Clontech, EcoRl- SacI fragment) and used it to replace the U3 part of the 5’ LTR, using PCR, so as to keep the transcription start site unchanged.

Point mutations within the protease and RT motifs (from LV/DYGA to LVaAa and from YIDD to alaa, respectively) were introduced using the Altered Sites II in vitro Mutagenesis System (Promega) and appropriate primers. Inactivation of the gag gene was achieved via an in-phase deletion between the Siol and PsbAl unique sites, and the env gene was disrupted by inserting a small synthetic linker containing Stop codons in all six frames at the unique Stul site (located at the end of the SU subunit, position 7784 in the sequence provided in Supplemental Fig. 1).

This series of expression vectors was marked with the SV neo resistance gene or SV neoTM retrotransposition indicator gene by inserting fragments taken from pSvneo* and pSvneoTM plas- mids (Esnault et al. 2002), respectively, at position 8662 of the consensus sequence (position according to the sequence provided in Supplemental Fig. 1). This provirus was replaced by the corresponding sequences of K113 and K115 proviruses (from the unique SexAI site to the end) or K108-I (from the SacI site to the end). For the triple chimera K109-K115-K108, the 3’ domain of K109-K115 was replaced by the corresponding part of the K108-I provirus (from the Xbal site to the end).

Production of antibodies: Western blot analysis and immunofluorescence studies

We developed a polyclonal anti-HERV-K Gag antiserum by im- munizing rabbits with a mix of purified recombinant proteins produced in bacteria, covering the complete HERV-K113 Gag protein. The resulting sera were checked by Immunoblot analyses and affinity-purified before use (Agrobio). The monoclonal antibody directed against the SU subunit of the HERV-K envelope protein has already been described (Dewannieux et al. 2005). We also raised a rabbit antiserum specific for the HERV-K envelope protein using a recombinant protein produced in bacteria, corre- sponding to aa 223–437 of the K109 Env protein.

For analysis of retroviral protein expression and cleavage, 293T cells transiently transfected with the provirus expression vectors (using Lipofectamine and +Reagent; Invitrogen) were lysed in Laemmli buffer 48 h post-transfection and lysate proteins were checked by Immunoblot analyses and affinity-purified before use (Agrobio). The monoclonal antibody directed against the SU subunit of the HERV-K envelope protein has already been described (Dewannieux et al. 2005). We also raised a rabbit antiserum specific for the HERV-K envelope protein using a recombinant protein produced in bacteria, corre- sponding to aa 223–437 of the K109 Env protein.

For analysis of retroviral protein expression and cleavage, 293T cells transiently transfected with the provirus expression vectors (using Lipofectamine and +Reagent; Invitrogen) were lysed in Laemmli buffer 48 h post-transfection and lysate proteins were checked by Immunoblot analyses and affinity-purified before use (Agrobio). The monoclonal antibody directed against the SU subunit of the HERV-K envelope protein has already been described (Dewannieux et al. 2005). We also raised a rabbit antiserum specific for the HERV-K envelope protein using a recombinant protein produced in bacteria, corre- sponding to aa 223–437 of the K109 Env protein.

For analysis of retroviral protein expression and cleavage, 293T cells transiently transfected with the provirus expression vectors (using Lipofectamine and +Reagent; Invitrogen) were lysed in Laemmli buffer 48 h post-transfection and lysate proteins were checked by Immunoblot analyses and affinity-purified before use (Agrobio). The monoclonal antibody directed against the SU subunit of the HERV-K envelope protein has already been described (Dewannieux et al. 2005). We also raised a rabbit antiserum specific for the HERV-K envelope protein using a recombinant protein produced in bacteria, corre- sponding to aa 223–437 of the K109 Env protein.

For analysis of retroviral protein expression and cleavage, 293T cells transiently transfected with the provirus expression vectors (using Lipofectamine and +Reagent; Invitrogen) were lysed in Laemmli buffer 48 h post-transfection and lysate proteins were checked by Immunoblot analyses and affinity-purified before use (Agrobio). The monoclonal antibody directed against the SU subunit of the HERV-K envelope protein has already been described (Dewannieux et al. 2005). We also raised a rabbit antiserum specific for the HERV-K envelope protein using a recombinant protein produced in bacteria, corre- sponding to aa 223–437 of the K109 Env protein.

For analysis of retroviral protein expression and cleavage, 293T cells transiently transfected with the provirus expression vectors (using Lipofectamine and +Reagent; Invitrogen) were lysed in Laemmli buffer 48 h post-transfection and lysate proteins were checked by Immunoblot analyses and affinity-purified before use (Agrobio). The monoclonal antibody directed against the SU subunit of the HERV-K envelope protein has already been described (Dewannieux et al. 2005). We also raised a rabbit antiserum specific for the HERV-K envelope protein using a recombinant protein produced in bacteria, corre- sponding to aa 223–437 of the K109 Env protein.
polyclonal anti-Gag serum and a mouse monoclonal anti-Env antibody. Alexa 546-conjugated goat anti-rabbit IgG and Alexa 488-conjugated goat anti-mouse IgG were used as secondary antibodies (Molecular Probes); nuclei were stained with Dapi and Topro3 (Molecular Probes).

Electron microscopy

For morphological studies, cells were fixed in phosphate buffer (pH 7.2), 3% glutaraldehyde for 1 h, and post-fixed in 0.1 M cacodylate buffer, 1% osmium tetroxide for 2 h. After being rinsed for 5 min in water and 15 min in 0.1 M cacodylate buffer, cells were transferred to 0.2 M cacodylate buffer for 30 min. Cells were washed in 30% methanol for 10 min, stained in 2% uranyl acetate in 0.1 M cacodylate buffer, 30% methanol for 1 h, and washed in 30% methanol. Cells were then dehydrated through a graded ethanol-propylene oxide series and embedded in Epon 812. Ultrathin sections were collected on formvar-carbon-coated gold grids (200-mesh), stained with 4% uranyl acetate and lead citrate, and examined with a Zeiss 902 microscope at 80 kV.

For immuno-electron microscopy experiments, transfected cells were fixed for 1 h at 4°C with either 4% formaldehyde or 1.6% glutaraldehyde in 0.1 M phosphate buffer (pH 7.3). During fixation, the cells were scraped from the plates and centrifuged. Pellets were dehydrated in increasing concentrations of methanol and embedded in Lowicryl K4M at low temperature. Polymerization was carried out for 5 d at −30°C under long-wavelength UV light. Ultrathin sections of Lowicryl-embedded material were collected on formvar-carbon-coated gold grids (200-mesh) and used for labeling with a rabbit antiserum specific for the SU subunit of the HERV-K envelope protein, followed by incubation with a goat anti-rabbit IgG conjugated to gold particles, 10 nm in diameter. Sections were contrasted as above. The specificity of the association between the gold beads and the viral particles was performed by counting, in 11 independent fields, the number of gold particles/µm² within three “compartments” (viral particles, cytoplasm, and particle-free extracellular space).

RT assay

Assays for the presence of RT activity in cell supernatants were performed using transiently transfected 293T cells. At day 2 post-transfection, the medium was harvested, filtered through 0.45-µm pore-sized membranes, and added to target cell lines (SH-SY5Y, BHK21, G355.5, HeLa, and WOP) seeded in 6-well plates the day prior to transfection. Transfection, the medium was harvested, filtered through 0.45-µm pore-sized membranes, and particulate material concentrated by ultracentrifugation on a 20% sucrose cushion. The virions were lysed in Buffer A (25 mM Tris, 0.25 mM EDTA, 50 mM KCl, 5 mM DTT, 0.025% Triton X-100, 50% glycerol) and used as a reverse-transcriptase source for reverse transcription of commercial MS2 RNA (adapted from the PERT assay described by Pyra et al. 1994). Briefly, 300 ng of MS2 RNA were annealed to 16 pmol of RT-1 primer and incubated with viral lysate corresponding to 1–2 mL of cell supernatant. cDNA synthesis was then assayed using 1/25 of the RT reaction as a template for a 25 cycle PCR reaction conducted with RT-1 and RT-2 primers (Pyra et al. 1994). In these assays, expression vectors for the MLV and HIV core proteins were used as positive controls and CMV-β (Clontech) as a negative control.

Infection and retrotransposition assays

For infection assays, 293T cells were used as producer cells, following transient transfection with the consensus provirus by the Calcium Phosphate method (Stratagene MBS commercial kit). Target cell lines (SH-SYSY, BHK21, G355.5, HeLa, and WOP) were seeded the day before infection in 6-well plates. The day of infection, the supernatants of 293T cells were harvested, filtered through 0.45-µm pore-sized membranes, and added to target cells (1 mL per well, with 4 µg/mL polybrene). To increase the sensitivity of the assay, the target cells were then subjected to “spinoculation” (centrifugation at 1200g for 2.5 h) and incubated overnight at 37°C before medium change. For detection of infection events, target cells were split into 10-cm plates (~5 × 10⁶ cells per plate) at day 3 post infection and subjected to G418 selection 24 h later. For infection assays using quantitative PCR, 10⁵ BHK21 cells were seeded in 6-well plates and used as target cells. At day 2 post-infection, the cells were washed with DNaseI to remove residual DNA and split, and genomic DNA was recovered at days 5–6 post-infection. DNA purification was performed via two successive phenol/chloroform extractions followed by ethanol precipitation. Integration events were detected with the Taqman technology using a 5′-FAM MGB probe (6FAM-ATAGTGATGGTTACCTGAG-MGBNFQ, Applied Biosystems) and primers specific for the HERV-K(HML2) family (5′-GCCTTTCCCCGCCCATTATT-3′ and 5′-GCAGGGCAGGGATCATATA-3′), and 500 ng of genomic DNA as a template. The infectious titers were estimated by comparison with a standard obtained by mixing genomic DNAs extracted from naive BHK21 cells and from a clone containing a single inserted provirus in varying proportions. In this assay, the titer of Phoenix WT was in the 10⁶ infectious particles/milliliter range. BHK21 cells were chosen as the target cell line, since their genomic DNA generates only very limited background with the HERV-K-specific primers and probe used in this assay; genomic DNA of BHK21 cells infected with Phoenix particles deficient for the RT was used as a negative control to rule out contamination by residual DNA that could originate from the transfected plasmid or genomic DNA of the producer cells.

For retrotransposition assays, 7.5 × 10⁵ SH-SYSY cells were seeded in 60-mm dishes the day prior to transfection. Transfection was performed using 8 µL of Plus Reagent (Invitrogen), 12 µL of Lipofectamine (Invitrogen), and 3 µg of DNA (2.5 µg of Phoexin-derived expression vector and 0.5 µg of env expression plasmid or control plasmid). Cells were grown for 1 wk before plating on 10-cm plates and subsequent selection with G418. The results presented here correspond to three independent experiments performed with 5–10 × 10⁶ cells and are given as relative transposition frequencies as compared with neo- or-marked Phoenix (whose absolute transposition frequency is 2 × 10⁻⁵ clone/seeded cell under these conditions).

Characterization of HERV-K(HML2) de novo insertion sites

Integration sites were analyzed by inverse PCR using individual G418° clones essentially as described (Dewannieux et al. 2003). Briefly, 5 µg of DNA from each clone was digested with PvuII and either (1) DraI, (2) SspI, or (3) with a mix of XbaI, AvrII, and SphI, or (4) with Apol alone. After enzyme removal, DNA was diluted and self-ligated overnight. It was ethanol-precipitated, dissolved in 20 µL of Tris 10 mM, and subjected to one (or in some cases two) round of PCR with sets of divergent primers located within the neo gene and designed for the amplification of the 5′ or 3′ end and flanking sequences of the newly inserted proviruses. The amplified DNAs were gel-purified, cloned, and sequenced, thus yielding the target-site sequences. The identification of these insertion sites was further confirmed by PCR reactions allowing direct amplification of both left and right junctions between the newly inserted proviruses and the identified loci, using for each reaction a primer specific for the identified insertion locus and an internal primer located in the neo gene. In each case, we checked that the same reactions performed on the naïve, noninfected target cells did not yield any specific amplification. Primer sequences are available on request.
Safety precautions
All manipulations involving the reconstructed HERV-K were carried out in our lab according to the rules established by the “Commission de Génie Génétique” from the “Ministère délégué à l’Enseignement supérieur et à la Recherche” French authority that regulates handling of genetically modified organisms in all research institutions in France.

Albeit the HERV-K virus has a very low infectivity and does not sustain multiple-cycle infection, at least in all the cells tested, \textit{Phoénix} is a retrovirus, and as such, is a priori eligible to BL3 conditions for manipulation. Accordingly, the material will only be sent to other labs in appropriate sealed containers in the form of small amounts of plasmid DNA that will require it to be amplified before use as a transfection vector to produce viral particles. At the present time and as a precautionary principle, it will only be distributed under a material transfer agreement specifying the commitment of the recipient labs to carry out every experiment using the material under BL3 conditions and accompanied by a duly signed authorization form from the Biosafety Committee responsible for genetic manipulations in their country of origin.

Acknowledgments
We thank A. Jalil for assistance with confocal microscopy, E. Pichard for technical assistance, and C. Laviolle for critical reading of the manuscript. This work was supported by the CNRS and by grants from the Ligue Nationale contre Le Cancer (Equipe Labellisée).

References


Received May 28, 2006; accepted in revised form August 30, 2006.