Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 334 (2005) 203–213

VIROLOGY

www.elsevier.com/locate/yviro

The distribution of *pol* containing human endogenous retroviruses in non-human primates

Alex D. Greenwood[a,b,*], Anna Stengel[a], Volker Erfle[a,b], Wolfgang Seifarth[c], Christine Leib-Mösch[a,c]

[a]*GSF-National Research Center for Environment and Health, Institute of Molecular Virology, D-85764 Neuherberg, Germany*
[b]*Technical University of Munich, Institute of Virology, D-81675 Munich, Germany*
[c]*Medical Clinic III, Faculty of Clinical Medicine Mannheim, University of Heidelberg, D-68305 Mannheim, Germany*

Received 11 November 2004; returned to author for revision 30 December 2004; accepted 31 January 2005

Abstract

Few human endogenous retroviruses (HERVs) have been extensively studied in non-human primates. Such investigations have demonstrated that several element classes are primate unique, contain members with important biological function, are conserved in specific primate lineages, and have in some cases expanded in copy number. We have examined multiple sub-families of all major groups of HERVs using a DNA microarray based on the reverse transcriptase (RT) domain of the viral polymerase gene (*pol*). The microarray was used to investigate the distribution of HERVs in non-human primates with particular focus on the differences between New World monkeys (NWMs) and other anthropoids. This is the first study examining most HERV families in multiple non-human primate DNAs using a uniform and sensitive method and suggests that major differences exist between primate groups. The results indicate that a major invasion and expansion of *pol* containing HERVs occurred after the platyrrhine (NWM) lineage separated from the catarrhines (Old World Monkeys and apes).

Keywords: Primates; HERV; Endogenous retrovirus; New World monkey; Old World monkey; DNA microarray; Quantitative PCR; Evolution

Introduction

HERVs are a major component of the human genome. They represent the successful colonization and expansion of retroelements in the germline of the species in which they are found. They can be divided into three broad classes each with multiple sub-families. HERV research has recently identified several elements in primates that may play a significant functional role in development. For example, syncytin-1 and syncytin-2 are both HERV envelope genes that are conserved among non-human primates (Blaise et al., 2003). In cell culture, they exhibit cell fusiogenicity, an activity which is demonstrated in most primates studied including NWMs in the case of syncytin-2. Both genes may be essential for syncytium formation in the placenta.

The lineage leading to the NWM or Platyrrhini separated from the lineage leading to the Catarrhini (Old World Monkeys, OWMs, and hominoid primates) 33 to 57 Mya (Glazko and Nei, 2003; Takahata and Satta, 1997). In some cases, retroelement evolution, that is, LINEs and SINEs, is similar among non-human primates (Boissinot et al., 2004). However, the genomic composition of HERVs differs markedly among major primate lineages. HERV-K, HERV-L, and HERV-H family distribution in non-human primates have been investigated for specific subgroups (Bénit et al., 1999; Mager and Freeman, 1995; Mayer et al., 1998; Reus et al., 2001). With the exception of HERV-L, NWMs appear to either lack or have low copy numbers of investigated HERVs. Expansion is largely restricted to catarrhine primates. Human-specific HERV integrations have also been observed indicating that the process of

* Corresponding author. GSF-National Research Center for Environment and Health, Institute of Molecular Virology, Building 35, Room 2028, Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany.
E-mail address: greenwood@gsf.de (A.D. Greenwood).

0042-6822/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2005.01.045

HERV expansion is still in progress among apes including humans (Barbulescu et al., 1999, 2001; Buzdin et al., 2003; Lebedev et al., 2000; Medstrand and Mager, 1998; Turner et al., 2001).

Most studies of HERV distribution in primates have used different methods, different primate species, and different HERV genes in determining presence or absence of specific element types. DNA microarrays provide a means to analyze multiple sequences rapidly and with high specificity using a single method. Recently, a microarray system using multiplex PCR with primers based on the most conserved portion of HERV subfamily RT and a HERV-specific DNA microarray spotted with portions of the RT gene of representatives of each HERV family and subfamily was developed (Seifarth et al., 2003). Using this system, we have screened multiple NWM, OWM, and hominoid DNAs for the presence or absence of specific *pol* containing HERV types. This is the largest sampling to date of non-human primates in an ERV study. The results obtained were reproducible and for specific elements confirmed using quantitative PCR and sequence analysis. It appears that the expansion of HERVs in OWM and hominoids is not restricted to specific types of elements. The general results demonstrate that NWM show a lower abundance of all HERV groups with the particular exception of class III spuma-like retroviruses (HERV-L). The data suggest that a major genome-wide evolutionary event occurred in HERV activity after the platyrrhine and catarrhine lineages separated.

Results

Class I HERVs

Fig. 1 summarizes the results of seven hybridization experiments for the class I HERVs. Members of the HERV-I family (HERV-IP, Seq 65) were detected in hominoids and OWMs and Seq 65 was detected in all groups studied though in only one NWM sample (Fig. 1). The overall copy number for HERV-I in humans is estimated at 250 copies per genome (Mager and Medstrand, 2003). For HERV-IP it is estimated to be 35 copies (Seifarth et al., 2000). In said study, a hybridization signal was detected in *Aotus* with a 700-bp RT hybridization probe. It is possible that HERV-IP has suffered deletions in NWMs and thus the smaller hybridization probe spotted on the microarray in this study does not detect it. This is not the case for Seq 65 as it was detectable in a NWM sample. Though clearly not an extremely high copy sequence in non-human primates given the weak signal detected and previous investigations of the HERV-IP subgroup (Seifarth et al., 2000), the results are consistent for this subclass of HERVs, homologs of which have been detected in such divergent phyla as fish, reptiles, and birds (Martin et al., 1997).

HERV-T (HERVS71) is known to be distributed in apes and OWMs (Haltmeier et al., 1995). This distribution was observed in this study as well. However, the element was not detected in *Cercopithecus aethiops* or *Presbytis cristata* suggesting the distribution is not continuous in OWMs. The element was not detected in any NWM tested. HERV-T is more closely related to mammalian γ-retroviruses such as MLV, GaLV, and FeLV than to other class I HERVs (Tristem, 2000; Werner et al., 1990) suggesting that HERV-T may have been acquired by cross-species transmission, that is, horizontal transfer during the diversification of the OWM lineage.

HERV-FRD (ERV-FRD, HERV-Z) were detected in Catarrhini. HERV-FRD *env* genes are known to be present in NWMs (Blaise et al., 2003, 2004). The *env* gene is also known to be highly conserved and in some cases retains active fusogenic properties. It suggests that the copy number is higher in OWMs and apes than in NWMs and/or that there is differential conservation of the *env* gene which could have an important function (Blaise et al., 2004), whereas the *pol* gene could be expendable.

HERV-W demonstrated a profile consistent with results of other studies. It has been shown previously that this element subclass is distributed throughout apes and OWMs but is absent from NWMs (Kim et al., 1999). An identical result is observed in this study showing that all apes and OWMs tested were strongly positive for HERV-W *pol* sequences and all NWMs were negative.

ERV9 distribution was found in all primates tested though there was among subclass variation. There are approximately 300 copies per genome of ERV9 related elements (Mager and Medstrand, 2003). Seq 63, ERV-9, and Seq 59 were detected in all apes and OWMs tested with the exception of *Gorilla gorilla* which did not show the presence of Seq 63. All NWMs were negative. Seq 60 had the widest distribution appearing in all apes except *G. gorilla*, all OWMs except *Mandrillus sphinx*, and in two NWMs (*Saguinus oedipus* and *Saguinus fuscicollis*). This result is consistent with the detection of ERV9 in OWMs in other studies (Widegren et al., 1996) though NWMs were never empirically tested.

Members of the HERV-ERI superfamily (ERV3, E 4-1, and Seq 32) were detected in catarrhines exclusively. This is generally consistent with previous information for OWMs (Herve et al., 2004; Shih et al., 1991). ERV3, however, was not detected in OWMs in contrast to previous findings. As the degenerate primer mix does not contain specific primers for ERV3, this may represent the favoring of related elements over ERV3 with the result that the method did not detect the element in OWMs although it is present.

It has been suggested that the HERV-ERI superfamily predates the split between Platyrrhini and Catarrhini as determined from sequence divergence estimates (Shih et al., 1991). However, none of the elements were detected in any NWM tested. This suggests that the NWM lineage lost these elements, suffered *pol* deletions, they remained at very low

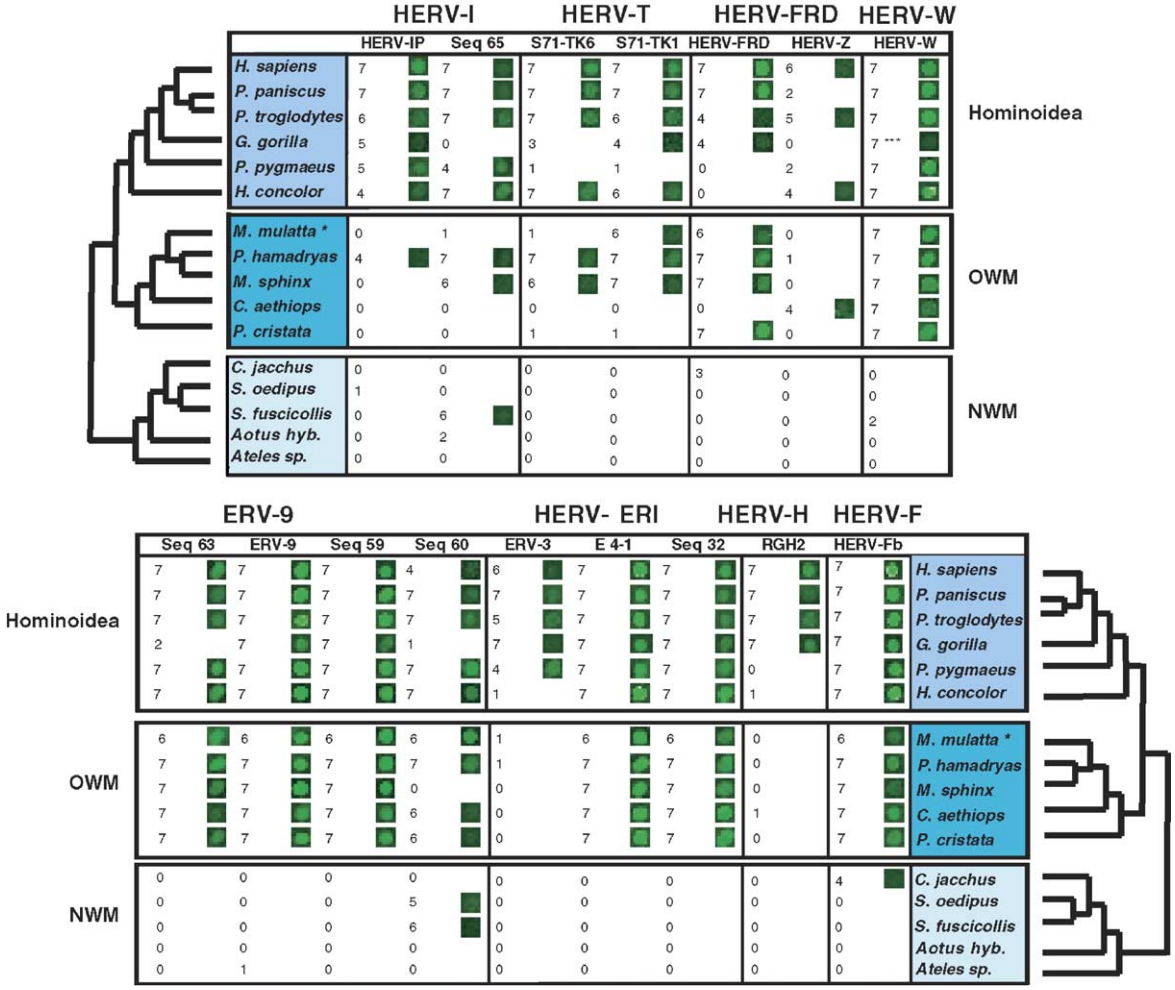

Fig. 1. Class I gammaretrovirus-like HERVs. The phylogenetic association of the different primates is indicated next to the names of each species. The broader taxonomic classification of NWM, OWM, and Hominoidea are indicated by color and are separated into different boxes. HERV subclasses are separated by a black vertical line and where applicable, HERV subgroups are indicated (for subgroup designations, see Seifarth et al., 2003 nomenclature used). The number of times a signal was detected for a given species is indicated. In some cases (1 to 3 times for a specific HERV) a signal was detected but was usually extremely faint and likely represents an artifact, that is, mishybridization due to sequence divergence. If not reproducible in 50% or more of the experiments, a specific HERV was scored negative to remain as conservative as possible. A representative image of a positive signal is provided for all elements determined to be present for each primate sample. "*" indicates that in the case of *M. mulatta*, only 6 repetitions were scored as one spotted grid was not detected. "***" indicates that a background signal was detected for the element and that the signal was clearly above the background for the species indicated.

copy number, or have rapidly diverged beyond the ability to recognize them with our methodology. Alternatively, the estimated age could be in error and the family is in fact younger. In any case, the evolutionary history of this element differs between NWMs and other non-human primates.

HERV-H (RGH2) was only detected in hominoids in this study. This contrasts with what is known about their distribution (Mager and Freeman, 1995). This subclass of element is known to occur in NWMs and catarrhines. However, in humans, only approximately 50–100 copies of about 1000 HERV-H elements are intact (Goodchild et al., 1993). The vast majority of HERV-H elements (ca 900) lack the *pol* sequences used as capture probes for the microarray. The remainder are partially deleted. NWMs were shown to have much lower copy numbers of the full-length elements with the only genomic amplification event occurring just prior to the separation of OWMs from apes (Mager and Freeman, 1995). The high occurrence of deletions in this subclass suggests that an accurate determination of the distribution of this class will not be feasible with our microarray strategy. Consistent with this finding, PCR amplification with the HERV-H primers designed for the multiplex PCR alone from catarrhines yielded a strong product. Weaker bands of a range of different sizes were amplified from *Ateles* sp., *S. oedipus*, and *S. fuscicollis* though not in *Aotus* hyb. suggesting non-specific amplification occurred. Sequencing of the products revealed a HERV-H like sequence. In addition, several larger fragments with homology to LINE elements were also detected (data not shown). This was not observed in any OWM or hominoid ape tested. The results emphasize that the evolu-

tionary history of this group of elements differs markedly between NMWs and other primates.

HERV-F (HERV-Fb) is known to be present in apes, OWMs and NWMs (Widegren et al., 1996). We found that the HERV-Fb subfamily (Tristem, 2000) is distributed widely among the primates tested. HERV-Fb was detected in all apes, OWMs and in the NWM *Callithrix jacchus*. Our results indicate HERV-Fb is present in NWMs though not in all species. This may represent sequence divergence or copy number changes among different lineages of NWMs.

Class II HERVs

The results for the class II betaretrovirus-like HERVs (HERV-K superfamily) are shown in Fig. 2. HML-1 has not been previously investigated for their distribution among non-human primates. HML-1 was identified in all the apes except for *Hylobates concolor* and *G. gorilla*. It was also identified in one OWM specimen. This suggests that the element entered the primate lineage after the NWMs split from the catarrhines. However, it does not appear to have been greatly amplified in these lineages or suffers from deletions in the *pol* gene as in the case of HERV-H.

HML-2 is known to occur in apes and OWMs. LTR divergence data suggest the group is 28 million years old and thus far younger than the platyrrhine/catarrhine split (Reus et al., 2001; Steinhuber et al., 1995). Consistent with this analysis, our data demonstrate that K(HP1) and K(D1.2) are found in all apes and OWMs with the exception of *C. aethiops* which lacked a signal for K(HP1). None of the NWMs tested yielded a signal for either HML-2 subgroup.

HML-3 (Seq 26, Seq 34, HML-3, SLE666, Seq 43) was detected in all apes and OWMs with two exceptions, the absence of Seq 34 from *Macaca mulatta* and Seq 43 from *M. sphinx*. No NWM yielded a signal for any subgroup of HML-3. The age of this subclass has been estimated at 36 million years from LTR sequence divergence (Mayer and Meese, 2002). This is younger than the estimated divergence date for NWM and catarrhines and thus our results confirm the prediction empirically.

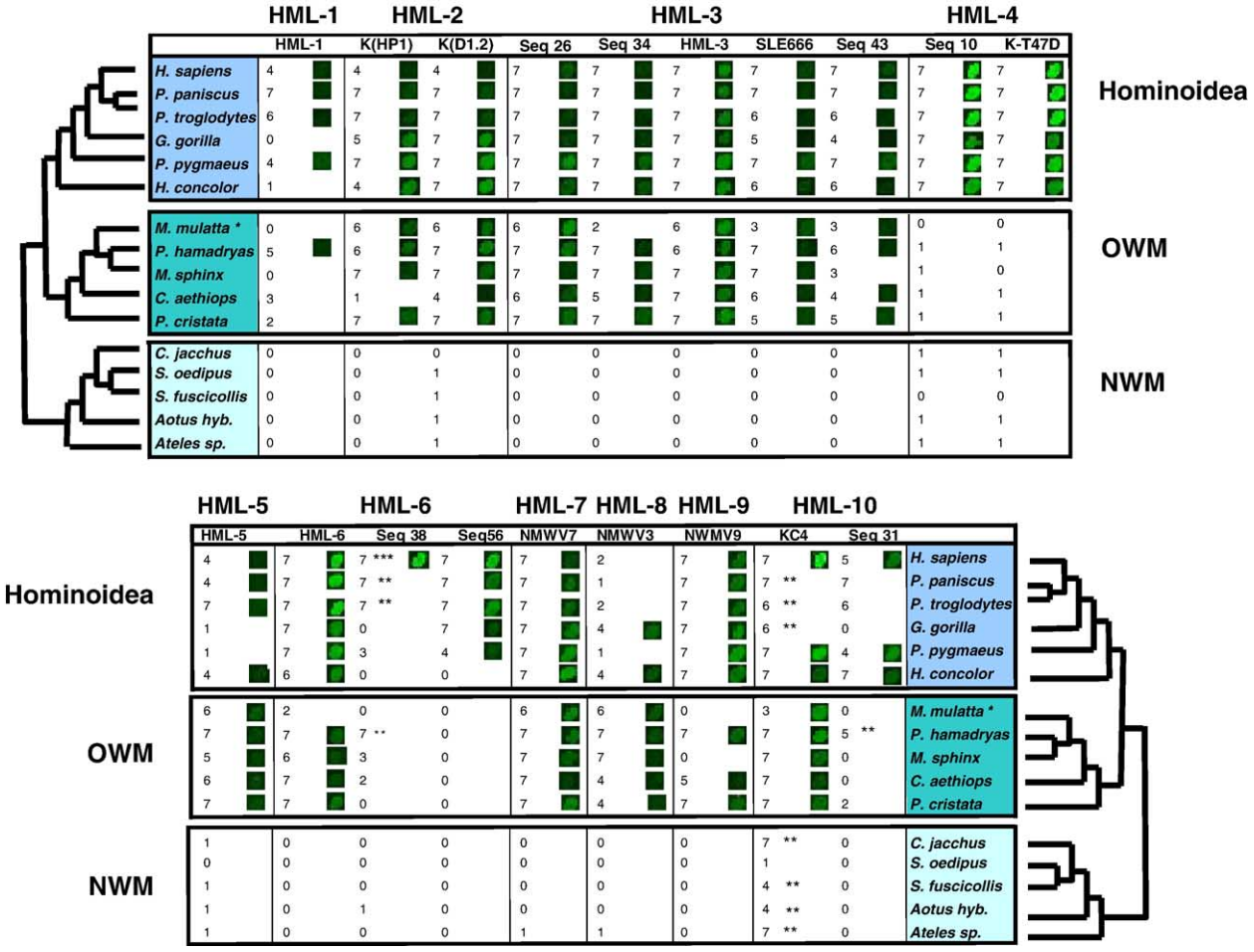

Fig. 2. Class II betaretrovirus-like HERVs (HERV-K superfamily). All nomenclature and designations are the same as in the legend of Fig. 1. "**" indicates that for these elements, there was a detectable contamination in the water control and the signal detected in the species indicated was not above this background.

HML-4 (Seq 10, K-T47D) is known to occur in OWM and apes but is absent from NWMs (Seifarth et al., 1998). Consistent with this conclusion, we did not detect the element in any NWM sample. However, our results demonstrate that the sequence is restricted to apes and is not detected in OWMs. The discrepancy suggests that the apes had an expansion of this element whereas OWMs have a lower copy, or *pol* deletions, and are thus below the detection threshold of our methods.

HML-5 distribution in non-human primates has recently been investigated by Lavie et al. (2004). Our results indicate the element is present in all apes and OWMs with the exception of *G. gorilla* and *Pongo pygmaeus*. The element was not detected in any NWM sample tested. PCR screening for HML-5 *gag* and *env* gene sequences in NWMs was positive although *pol* was not examined (Lavie et al., 2004). Also, based on LTR sequence divergence and the genetic distances from the consensus sequence for five different proviral regions suggests HML-5 integration predates the Catarrhini/NWM split. The discrepancy between our results and those of Lavie et al. (2004) may reflect the gene region chosen for analysis and the relative divergence of the different HERV genes.

HML-6 (HML-6, Seq 38, Seq 56) have not been examined for presence or absence in NWMs extensively. HML-6 was present in all apes and OWM's except for *M. mulatta*. This is consistent with results for HERV-K in general (Mayer et al., 1998). Seq 38 was only detected in human while Seq 56 was restricted to the great apes excluding *H. concolor*. Again, these elements appear to have invaded the primate germline after the platyrrhine and catarrhine lineages diverged.

HML-7,8 and 9 have not been investigated for their distributions among non-human primates. Our results demonstrate for each subclass a broad distribution among OWMs and apes. HML-7 was observed in all apes and OWMs tested in this study. HML-8 was not observed in human, *Pan paniscus* or *Pan troglodytes* most likely due to deletions in the majority of *pol* genes in these species. It was also absent from *P. pygmaeus*. However, it was detected in all OWMs tested. HML-9 was absent from *M. mulatta* and *M. sphinx*. It was also absent from all NWMs tested. HML-10 (KC4) elements are known to exist in apes and OWMs (Dangel et al., 1994; Johnson and Coffin, 1999). KC4 was detected in apes and OWMs whereas Seq 31 was restricted to apes.

Class III HERVs

HERV-L element LG895 was detected in all primate DNAs tested except for *Ateles* sp. (Fig. 3). Seq 39 was not detected in *P. pygmaeus*, *M. mulatta*, and all NWMs with the exception of *Aotus* hyb. Seq 45 was absent from all NWMs with the exception of *Aotus* hyb., and was not detected in *P. cristata*, *M. mulatta* and *G. gorilla*. Seq 51 was absent in *C. jacchus* and *Ateles* sp. Seq 58 was absent in all NWM except for *S. fuscicollis*, all OWM except for *P. hamadryas* and was not detected in *G. gorilla* or surprisingly in *H. sapiens*. While it is clear that HERV-L is present in NWMs, the weak hybridization to human DNA and the somewhat inconsistent results with Seq 45 and 39 could reflect poor amplification, hybridization and sequence copy number differences and sequence divergence.

HERV-L presence in all non-human primates is consistent with the known evolution of this class of element. HERV-L related sequences have been detected in all mammals tested to date (Bénit et al., 1999; Greenwood et al., 2001, 2004). NWM sequences have been characterized and copy number estimated at approximately 200 copies much like in the human genome (Bénit et al., 1999). This class of element is thus similar to HERV-I which, in addition to having a germline entry time predating the separation of NWMs from OWMs, is pervasive throughout Mammalia. HERV-L together with HERV-I, may represent two of the oldest and most conserved HERV classes yet discovered. It also suggests that the methodology employed in this study is robust enough to detect element groups that are ancient and contain members that are highly divergent which is the case for ERV-L (Bénit et al., 1999).

Mammalian ERVs and housekeeping genes

Fig. 4 shows the microarray results for several mammalian ERVs and the RPL19 control gene. Several mammalian exogenous and endogenous retroviruses were spotted on the microarray (MMTV, MPMV, MLV, PERV, FIV, BaEV GaLV, BoEV, OvEV-1, OvEV-2, JSRV). However, no human exogenous retrovirus was detected in any experiment performed from any sample tested (HIV-1, HIV-2, HTLV-1, HTLV-2, Foamy, data not shown). Although all PCR reaction water controls were negative as determined by gel electrophoresis, there was a low but clear background in all water controls on the microarray for several of the house keeping genes. In the case of RPL19, all samples tested greatly exceeded the background signal. This was not the case for actin or HPRT (not shown). Given the contamination prevention precautions taken and the negative result for the water controls as determined by gel electrophoretic analysis, the background may represent an artifact signal derived from the primer dimers. Additional support for this comes from the quantitative PCR analysis. Contamination of the water control for RPL19 was never observed from the same samples used for the microarray work.

Of interest, BaEV was detected in all experiments with *M. mulatta*, *P. hamadryas*, *M. sphinx*, and *C. aethiops*. *P. cristata* in contrast, did not demonstrate the presence of this element. None of the apes or NMW samples gave a positive signal either. This is consistent with the available information for this cross species jumping element (van der Kuyl et al., 1995). Similarly, GaLV was detected in the same samples consistent with the known cross species infectivity of this class of retroelement (Hanger et al., 2000; Martin et

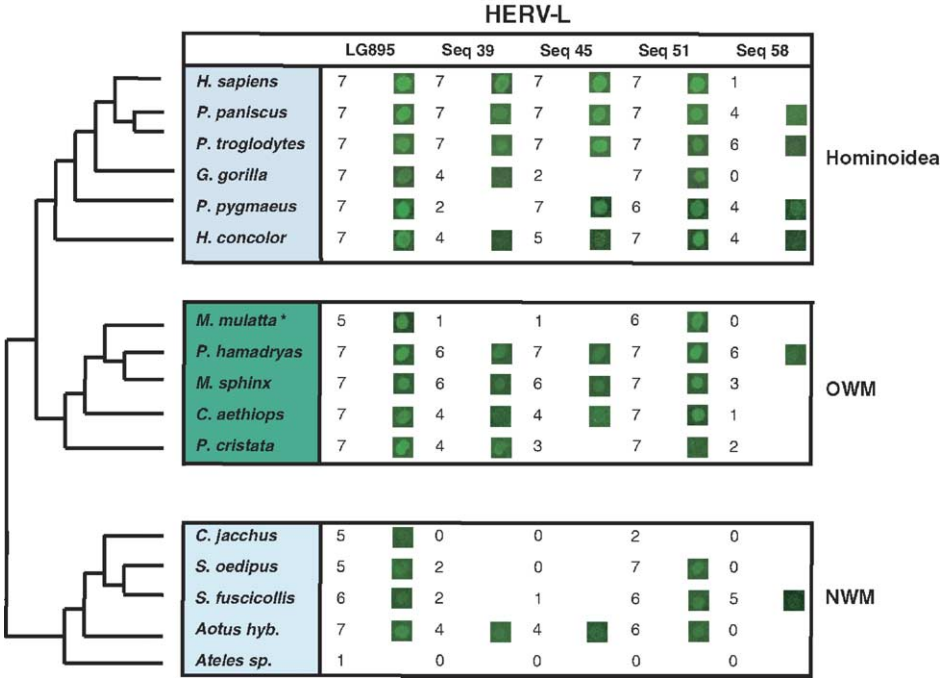

Fig. 3. Class III spuma-like HERVs. See Figs. 1 and 2 legends for description of the layout of the figure.

al., 1999). The Mason–Pfizer monkey virus (MPMV) was detected in *P. cristata* but not in the NWMs or hominoid apes. MPMV is a known OWM exogenous virus and thus its presence in *P. cristata* is consistent with the available information (Barker et al., 1985).

Quantitative PCR results and sequence analysis

Results for RPL19, HML-3 (Seq 26), HERV-E, and HML-5 were confirmed using SYBR I Green and quantitative PCR analysis (Fig. 5). The samples tested were a water-negative control, human, *C. aethiops*, *C. jacchus*, *Aotus* hyb., and *S. fuscicollis* DNA. Quantitation and subsequent gel electrophoresis of the amplified products confirmed the results for RPL19 which was detected in all samples except for the water control. Seq 26 (HML-3) was restricted to human and *C. aethiops*. There are approximately 150 *pol* containing copies per genome of HML-3 as a group (Mager and Medstrand, 2003). *C. aethiops* yielded an approximately 13% higher relative ratio for Seq 26. The exact copy number for the HML-3 subgroup represented by Seq 26 is not known. However, the results

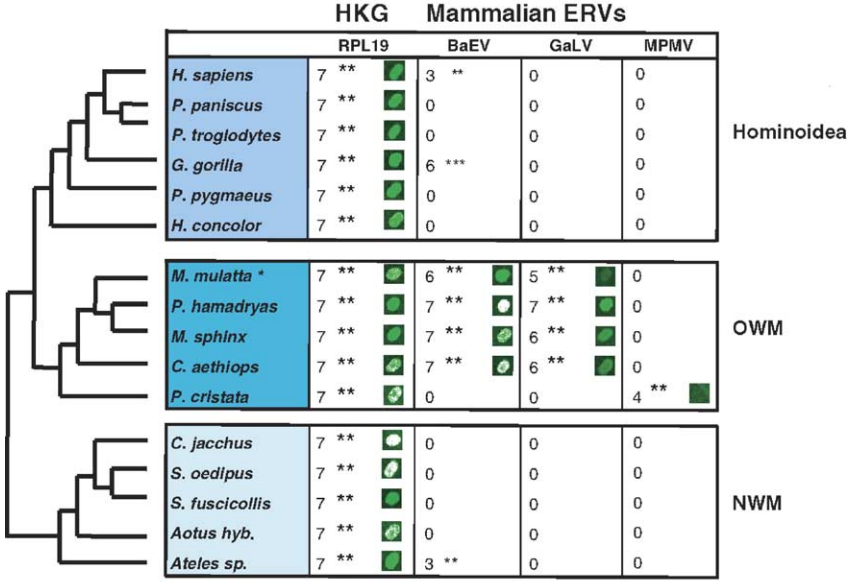

Fig. 4. Housekeeping genes, human exogenous retroviruses, and mammalian ERVs. See Figs. 1 and 2 legends for a description of the figure layout.

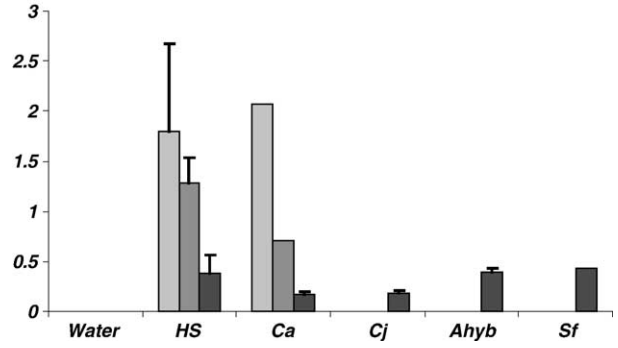

Fig. 5. Quantitative PCR results. The averaged results of triplicate quantitative PCR experiments for HML-3 (Seq 26), HERV-E, and HML-5 are shown. The data are presented as the ratios of the target gene/control gene (RPL19). Standard error bars are indicated. Samples tested were human (Hs), *C. athieops* (Ca), *C. jacchus* (Cj), *Aotus* hyb., and *S. fuscicollis* (Sf). HML-3 is represented by the light shaded bars, HERV-E by the grey bars, and HML-5 by the dark shaded bars.

from both the microarray and quantitative PCR suggest this element is similar in copy number in both apes and OWMs.

HERV-E was also restricted to human and the OWM sample. *C. aethiops* yielded an average relative ratio approximately 55% lower than human. There are about 250 copies of HERV-E in the human genome (Mager and Medstrand, 2003). Thus, *C. aethiops* has approximately 140 copies per genome.

HML-5 however, produced a signal in all samples contrary to the microarray results. Similarly, PCR with the HML-5 primers used in the degenerate primer mix for the microarray experiments followed by gel electrophoresis yielded a product in NWMs (data not shown). HML-5 exists in approximately 100 *pol* containing copies in the human genome (Mager and Medstrand, 2003). Thus, from the average determined relative ratios, there are about 50 copies per genome for *C. aethiops* and *C. jacchus* and 100 for *Aotus* hyb. and *S. fuscicollis*. An explanation for the failure of HML-5 to hybridize with NWMs on the microarray may be that the amplified sequences are too divergent. Alternatively, the homologous sequences to the probe may be frequently deleted and a more distantly related element amplified instead. To address this, the HML-5 products from *S. fuscicollis* and *S. oedipus*, respectively, were cloned and sequenced. HML-5 was also cloned and sequenced from human. In the case of HML-5, the expected sequence was amplified from human as determined by sequencing (data not shown). However, the sequence underlying the 5′ primer of the HML-5 primer used in the microarray experiment (which was contained in the fragment amplified) was divergent by up to 1–5 bp in NWMs. This could account for the failure of the microarray experiment primers to specifically amplify and/or the product to hybridize to the capture probe in NWMs.

HERV-H sequences amplified with the HERV-H-specific primers used in the degenerate primer mix were also sequenced. HERV-H from *S. oedipus* produced a sequence very similar to the human HERV-H sequence spotted on the microarray (5–6 differences among various clones sequenced, data not shown). However, unlike all catarrhines tested, *S. oedipus* and several other tested NWMs produced multiple larger PCR products or alternatively no product. Sequencing of these products revealed homology to LINE-1 like elements. Although HERV-H like sequence was detected, the primers appear to bind to other retroelements in NWMs at the expense of the HERV-H product. The diversion of the PCR primers to the other sequences explains the difficulty in detecting HERV-H in these samples using the microarray.

Similarly, HML-4 which is known to occur in apes and OWMs (Seifarth et al., 1998) was only detected in apes in this study. PCR amplification from human, *C. aethiops*, *C. jacchus*, *Aotus* hyb., and *S. fuscicollis* with the primers for HML-4 (non-multiplex) yielded a strong band in human, no product in *C. aethiops* and multiple products in most of the NWMs tested. A clear though weaker band was detected in *Aotus* hyb. Additional non-specific bands were also detected in the PCR from this sample. The fragment from *Aotus* hyb. was cloned and 14 clones sequenced revealing two classes of products, one weakly related to HERV-H and another with weak homology to an MalR repeat element (Smit, 1993). The combination of weaker product, sequence divergence, non-specific amplification, and further difficulties related to conditions of multiplex PCR suggest this is why detection with the microarray was restricted to apes much like HERV-H.

Discussion

PCR using mixed degenerate primers based on the most conserved region of the HERV *pol* gene followed by selection with HERV family specific probes is a rapid and sensitive way to score many sequences in a large number of samples for the presence or absence of specific *pol* containing HERVs. This method has been used successfully for detecting differential expression of HERVs in human tissues (Seifarth et al., 2003, 2005). The results of this study were reproducible and confirmable with additional methods. In the few cases where there was a discrepancy between our data and the known distribution of a given HERV, the results could be reconciled, that is HERV-H, HML-4, and HML-5. In cases of lack of amplification in non-human primates that were scored negative although 1 to 3 experiments produced faint signals, the complex PCR could potentially generate artifacts or random cross reactions. Multiple replications were thus performed to insure data quality and a conservative scoring protocol was used to avoid false-positives. This is especially important for the non-human primate samples which may vary in the primer binding sequences and thus the multiplex PCR could potentially be skewed in favor of some elements over others. The control gene RPL19 was detectable throughout the samples tested and thus, conserved loci in principle should be detectable by this

method for the evolutionary time scale examined in this study. Finally, the results are consistent with those for which elements have been studied in non-human primates such as HERV-W, HERV-L and several of the HERV-K subgroups.

While the possibility exists that NWM ERV detection would be hindered by greater sequence divergence of specific elements, it is telling that the general trend was absence of elements regardless of class with the exception of HERV-L elements which are known to be widely distributed among mammals. The control gene was clearly detectable in all NWMs so general sequence divergence would not explain the absence of many class I and II HERVs. The length of time since the lineage leading to OWMs and the lineage leading to apes diverged is also quite large and yet most sequences were clearly detectable. Some elements were absent among some apes but detected in OWMs, an inverse of the expectation and suggests taxon-specific levels of sequence amplification for given elements. In the cases where an element is known to be present in NWMs but was not detected in this study, that is, HERV-H, it is known that they are in far lower copy number than in catarrhines which in itself represents an evolutionary difference between platyrrhines and catarrhines. Finally, the study was based on the most conserved portion of the *pol* gene. This cannot exclude the possibility that other regions of some element classes exist in NWM genomes.

Fig. 6 summarizes the evolutionary history of HERVs in non-human primates as determined from this study combined with work that precedes it. It is clear that the major evolutionary events whereby *pol* containing HERVs entered the primate genome and/or expanded were subsequent to the split of NWMs from the Tarsii or Strepsirrhine primates and then more profoundly, after the NWMs and OWMs diverged. Class III (HERV-L) precedes the evolution of primates as they are widely distributed among mammals (Bénit et al., 1999). Of interest, virtually all class II elements, except HML-5, appear to be unique to the Catarrhini. Class II elements have been the focus of considerable research in Hominoid apes and this study provides empirical data for many evolutionary predictions for this class of elements. It re-affirms that the majority of the HERV-K superfamily has an origin subsequent to the separation NWMs from other anthropoid primates. Given that this group includes recently or currently active HERVs, it suggests this type of element is still in the process of endogenization and maintains many viral functions. Class I elements are more varied in their distribution, some like HERV-I predating the platyrrhine/catarrhine split and others with a younger distribution, such as HERV-T.

Almost all HERVs were detected in OWMs. Several exceptions included ERV3, HERV-H, Seq 31 (HML-10 subgroup), HML-4, Seq 38, and Seq 56 (both HML-6 subgroups). We cannot, however, rule out the presence of these elements in some form from OWMs. Some may have deletions in the *pol* gene and would be undetectable with this assay due to the relative short *pol* sequence spanned by the capture probes of the microarray. For example, the HML-4 subgroup is known to be present in OWMs and absent from NWMs (Seifarth et al., 1998). In our study, it was only detected in the Hominoidea. In humans, HERV-K T47D exists in 6 copies containing *pol*. This may be lower in OWMs or only solo LTRs remain and thus would not be detectable with our assay.

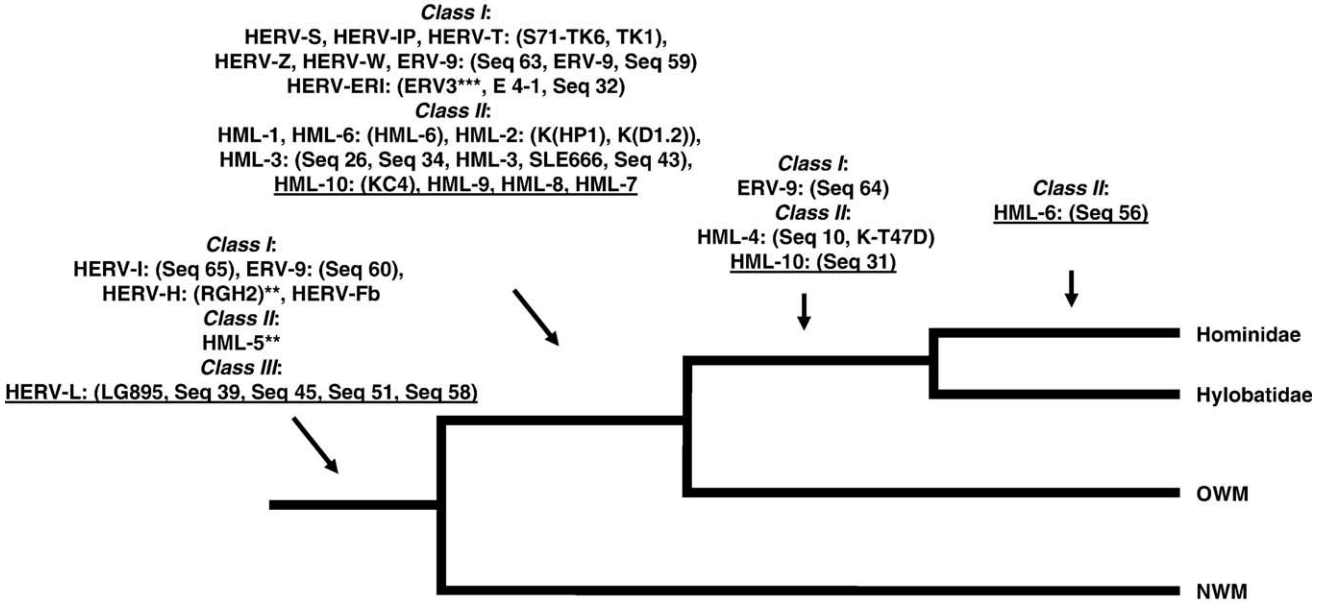

Fig. 6. Evolutionary history of HERVs in non-human primates. The phylogenetic tree shows the broad separation of Hominoid apes, Hylobatidae, OWMs, and NWMs. Presumed integration/expansion events for each HERV subclass based on this study are indicated with arrows. "**" indicates conflicts with the microarray data and literature that were resolved with quantitative PCR and sequence analysis. "***" indicates a conflict with the microarray results and the literature that is likely due to sequence divergence of the element.

Considering HERVs make up 8–9% of the human genome (Lander et al., 2001), the major expansion of this class of retroelements after the separation of platyrrhines and catarrhines suggests a significant evolutionary event occurred subsequent to the divergence. Unlike most genes, pseudogenes, and other genomic elements, the HERV component of the mammalian genome appears to be capable of significant dynamic plasticity. Of all autonomous retrotransposons, only LTR elements possess the ability to infect other organisms, whereas the spread of non-LTR retrotransposons such as LINE elements is restricted to retrotransposition within a cell. LTR elements with an *env* gene may escape the cell and exist for long periods as exogenous retroviruses. For example, there is evidence that the proliferation of the HERV-K(HML-2) family has been almost entirely due to germ-line re-infection rather than retrotransposition. Infectious members of this family have persisted within the primate lineage over the past 30 million years (Belshaw et al., 2004). Interspecies transmission of retroviruses have often occurred during the evolution of placental mammals (Martin et al., 1999). This may explain the discontinuous evolutionary history of HERVs in primates compared with other retrotransposons. Horizontal transmission into and among primate lineages could be an ongoing evolutionary process. An example is the transfer of an endogenous koala retrovirus (KoRV) into gibbons as suggested by the remarkably high sequence similarity between KoRV and the highly oncogenic gibbon ape leukemia virus (GaLV). The level of sequence divergence between GaLV and KoRV is comparable to that observed among different strains of GaLV indicating a very recent interspecies transfer (Hanger et al., 2000; Martin et al., 1999). HIV-1 and -2 are prominent examples of primate retroviruses that have crossed the species barrier into humans during the last century (Hahn et al., 2000).

The data in this study suggest among other things, a biological uniqueness of the catarrhine primates relative to the platyrrhines. Tracing the origin and the mechanisms responsible for these differences should guide future work on non-human primate endogenous retroviral research.

Materials and methods

DNA sample

Non-human-primate DNAs were kindly provided by J. Blusch (Novartis, Basel, Switzerland). Hominoid genomic DNA samples included common chimpanzee (*P. troglodytes*), bonobo (*P. paniscus*), gorilla (*G. gorilla*), and orangutan (*P. pygmaeus*). Hylobatidae was represented by a gibbon sample (*H. concolor*). OWM samples included rhesus macaque (*M. mulatta*), hamadryas baboon (*Papio hamadryas*), mandril (*M. sphinx*), vervet monkey (*C. aethiops*), and silvered langur (*P. cristata*). NWM samples included the common marmoset (*C. jacchus*), cotton top tamarin (*S. oedipus*), saddle back tamarin (*S. fuscicollis*), a hybrid owl monkey (*Aotus azarae x leumurinus*, here called *Aotus* hyb.), and spider monkey (*Ateles* sp.). Human genomic DNA was prepared from blood by standard methods (Sambrook et al., 1989).

Multiplex PCR and DNA microarray analysis

To inhibit amplification of contaminating DNA, all PCRs were carried out in a flow hood after 10–15 min of UV irradiation. The room and reagents used are dedicated to the setup of PCR reactions and no DNA or amplified products are ever present in the room. One repetition was carried out in a separate lab under similar conditions (Medical Clinic III, Mannheim). The PCR primers used in this study are based on conserved regions of the *pol* gene for the different HERV classes and are described in Seifarth et al. (2003). PCRs were performed as described in Seifarth et al. (2003) for 35 cycles for each primate sample. A human DNA-positive control and a water-negative control were included in most repetitions. 5 μl of each reaction was visualized on an ethidium bromide stained 2% agarose gel to insure that amplification had occurred and that the water control was negative. In all cases, primate DNAs generated strong PCR products and the water controls were negative.

The DNA microarray consists of 90 bp RT sequences specific for subfamilies of type class I, II, and III HERVs. The layout of the chip is detailed in Seifarth et al. (2003) with additional newly characterized HERV *pol* sequences spotted (Seifarth et al., 2005). Hybridization was performed for each primate sample in triplicate. Water controls were hybridized in all but one experiment. As each microarray is spotted in triplicate with the exception of one experiment for which data were only obtained for one replicate, a total of 7 spots were scored for each HERV for each sample. In a few instances, there was no signal detected at all or none in human. These were eliminated from further analysis. All procedures were as described in Seifarth et al. (2003) except that the hybridization solution consisted of 3× SSC and 0.2% BSA and hybridizations were carried out at 58 °C.

Scanning and microarray evaluation

The program GenePix Pro 3.0 (Biodiscovery Inc., Los Angeles, CA, USA) was used to generate JPEG files for evaluating background, contamination, and presence or absence of each spot in a given sample. If a signal was detected for a given HERV in 50% or more experiments it was scored as positive. In addition, for a few specific elements, faint water-positive signals were detected in one or more experiments, although PCR reactions were always devoid of bands. This background was also taken into consideration when scoring the specific elements (see Fig. 1 legend).

Quantitative PCR

To validate a portion of the microarray results quantitative PCR using LightCycler technology (Roche) and SYBR I Green was used for HML-3 (Seq 26), HERV-E, and HML-5 and RPL19 as a control. Primers used were RPL19 (5′ CCCGAATGCCAGAGAAGG 3′, 5′ CTTCCTTGGTCTTAGACCTG 3′), Seq 26 (5′ CTGCAGCCTGCTAAGCG 3′, 5′ CACTGTGAAAATTTTTTACGAG 3′), HERV-E (5′ GGTGTCACTACTCAATACAC 3′, 5′ GCAGCCTAGGTCTCTGG 3′), HML-5 (5′ TGAAAGGCCAGCTTGCTG 3′, 5′ CAATTAGGAAATTCTTTTCTAC 3′). The expected product size for each PCR reaction is approximately 100 bp. Experiments were done in triplicate and according to manufacturer's instruction. The data were normalized to an RPL19 control amplification. Melting curve analysis was performed after each experiment to exclude primer dimer measurement. 3–5 μl of each reaction was run on a gel to confirm that the results reflect the measurement of a specific product and not an artifact.

Acknowledgments

The authors wish to thank J. Blusch, T. Werner, and R. Brack-Werner for providing the primate samples and making this study possible. We thank C. Englbrecht for helpful discussion. We also thank O. Frank who provided advice for the quantitative PCR optimization.

References

Barbulescu, M., Turner, G., Seaman, M.I., Deinard, A.S., Kidd, K.K., Lenz, J., 1999. Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. Curr. Biol. 9, 861–868.

Barbulescu, M., Turner, G., Su, M., Kim, R., Jensen-Seaman, M.I., Deinard, A.S., Kidd, K.K., Lenz, J., 2001. A HERV-K provirus in chimpanzees, bonobos and gorillas, but not humans. Curr. Biol. 11, 779–783.

Barker, C.S., Wills, J.W., Bradac, J.A., Hunter, E., 1985. Molecular cloning of the Mason–Pfizer monkey virus genome: characterization and cloning of subgenomic fragments. Virology 142, 223–240.

Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Paces, J., Burt, A., Tristem, M., 2004. Long-term reinfection of the human genome by endogenous retroviruses. Proc. Natl. Acad. Sci. U.S.A. 101, 4894–4899.

Bénit, L., Lallemand, J.B., Casella, J.F., Philippe, H., Heidmann, T., 1999. ERV-L elements: a family of endogenous retrovirus-like elements active throughout the evolution of mammals. J. Virol. 73, 3301–3308.

Blaise, S., de Parseval, N., Bénit, L., Heidmann, T., 2003. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc. Natl. Acad. Sci. U.S.A. 100, 13013–13018.

Blaise, S., Ruggieri, A., Dewannieux, M., Cosset, F.L., Heidmann, T., 2004. Identification of an envelope protein from the FRD family of human endogenous retroviruses (HERV-FRD) conferring infectivity and functional conservation among simians. J. Virol. 78, 1050–1054.

Boissinot, S., Roos, C., Furano, A.V., 2004. Different rates of LINE-1 (L1) retrotransposon amplification and evolution in New World monkeys. J. Mol. Evol. 58, 122–130.

Buzdin, A.A., Lebedev, Y.B., Sverdlov, E.D., 2003. Human genome-specific HERV-K intron LTR genes have a random orientation relative to the direction of transcription, and, possibly, participated in antisense gene expression regulation. Bioorg. Khim. 29, 103–106.

Dangel, A.W., Mendoza, A.R., Baker, B.J., Daniel, C.M., Carroll, M.C., Wu, L.C., Yu, C.Y., 1994. The dichotomous size variation of human complement C4 genes is mediated by a novel family of endogenous retroviruses, which also establishes species-specific genomic patterns among Old World primates. Immunogenetics 40, 425–436.

Glazko, G.V., Nei, M., 2003. Estimation of divergence times for major lineages of primate species. Mol. Biol. Evol. 20, 424–434.

Goodchild, N.L., Wilkinson, D.A., Mager, D.L., 1993. Recent evolutionary expansion of a subfamily of RTVL-H human endogenous retrovirus-like elements. Virology 196, 778–788.

Greenwood, A.D., Lee, F., Capelli, C., DeSalle, R., Tikhonov, A., Marx, P.A., MacPhee, R.D.E., 2001. Evolution of endogenous retrovirus-like elements of the woolly mammoth (*Mammuthus primigenius*) and its relatives. Mol. Biol. Evol. 18, 840–847.

Greenwood, A.D., Englbrecht, C.C., MacPhee, R.D.E., 2004. Characterization of an endogenous retrovirus class in elephants and their relatives. BMC Evol. Biol. 4, 38.

Hahn, B.H., Shaw, G.M., De Cock, K.M., Sharp, P.M., 2000. AIDS as a zoonosis: scientific and public health implications. Science 287, 607–614.

Haltmeier, M., Seifarth, W., Blusch, J., Erfle, V., Hehlmann, R., Leib-Mosch, C., 1995. Identification of S71-related human endogenous retroviral sequences with full-length pol genes. Virology 209, 550–560.

Hanger, J.J., Bromham, L.D., McKee, J.J., O'Brien, T.M., Robinson, W.F., 2000. The nucleotide sequence of koala (*Phascolarctos cinereus*) retrovirus: a novel type C endogenous virus related to Gibbon ape leukemia virus. J. Virol. 74, 4264–4272.

Herve, C.A., Forrest, G., Löwer, R., Griffiths, D.J., Venables, P.J., 2004. Conservation and loss of the ERV3 open reading frame in primates. Genomics 83, 940–943.

Johnson, W.E., Coffin, J.M., 1999. Constructing primate phylogenies from ancient retrovirus sequences. Proc. Natl. Acad. Sci. U.S.A. 96, 10254–10260.

Kim, H.S., Takenaka, O., Crow, T.J., 1999. Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates. J. Gen. Virol. 80, 2613–2619.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860–921.

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

Lebedev, Y.B., Belonovitch, O.S., Zybrova, N.V., Khil, P.P., Kurdyukov, S.G., Vinogradova, T.V., Hunsmann, G., Sverdlov, E.D., 2000.

Differences in HERV-K LTR insertions in orthologous loci of humans and great apes. Gene 247, 265–277.
Mager, D.L., Freeman, J.D., 1995. HERV-H endogenous retroviruses: presence in the New World branch but amplification in the Old World primate lineage. Virology 213, 395–404.
Mager, D.L., Medstrand, P., 2003. Retroviral repeat sequences. In: Cooper, D. (Ed.), Nature Encyclopedia of the Human Genome. Nature Publishing Group, London.
Martin, J., Herniou, E., Cook, J., Waugh O'Neill, R., Tristem, M., 1997. Human endogenous retrovirus type I-related viruses have an apparently widespread distribution within vertebrates. J. Virol. 71, 437–443.
Martin, J., Herniou, E., Cook, J., O'Neill, R.W., Tristem, M., 1999. Interclass transmission and phyletic host tracking in murine leukemia virus-related retroviruses. J. Virol. 73, 2442–2449.
Mayer, J., Meese, E.U., 2002. The human endogenous retrovirus family HERV-K(HML-3). Genomics 80, 331–343.
Mayer, J., Meese, E., Mueller-Lantzsch, N., 1998. Human endogenous retrovirus K homologous sequences and their coding capacity in Old World primates. J. Virol. 72, 1870–1875.
Medstrand, P., Mager, D.L., 1998. Human-specific integrations of the HERV-K endogenous retrovirus family. J. Virol. 72, 9782–9787.
Reus, K., Mayer, J., Sauter, M., Zischler, H., Muller-Lantzsch, N., Meese, E., 2001. HERV-K(OLD): ancestor sequences of the human endogenous retrovirus family HERV-K(HML-2). J. Virol. 75, 8917–8926.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory.
Seifarth, W., Baust, C., Murr, A., Skladny, H., Krieg-Schneider, F., Blusch, J., Werner, T., Hehlmann, R., Leib-Mosch, C., 1998. Proviral structure, chromosomal location, and expression of HERV-K-T47D, a novel human endogenous retrovirus derived from T47D particles. J. Virol. 72, 8384–8391.
Seifarth, W., Baust, C., Schon, U., Reichert, A., Hehlmann, R., Leib-Mosch, C., 2000. HERV-IP-T47D, a novel type C-related human endogenous retroviral sequence derived from T47D particles. AIDS Res. Hum. Retroviruses 16, 471–480.
Seifarth, W., Spiess, B., Zeilfelder, U., Speth, C., Hehlmann, R., Leib-Mosch, C., 2003. Assessment of retroviral activity using a universal retrovirus chip. J. Virol. Methods 112, 79–91.
Seifarth, W., Frank, O., Zeilfelder, U., Spiess, B., Greenwood, A.D., Hehlmann, R., Leib-Mösch, C., 2005. Comprehensive analysis of human endogenous retrovirus (HERV) transcriptional activity in human tissues with a retrovirus-specific microarray. J. Virol. 79, 341–352.
Shih, A., Coutavas, E.E., Rush, M.G., 1991. Evolutionary implications of primate endogenous retroviruses. Virology 182, 495–502.
Smit, A.F.A., 1993. Identification of a new, abundant superfamily of mammalian LTR-transposons. Nucleic Acids Res. 21, 1863–1872.
Steinhuber, S., Brack, M., Hunsmann, G., Schwelberger, H., Dierich, M.P., Vogetseder, W., 1995. Distribution of human endogenous retrovirus HERV-K genomes in humans and different primates. Hum. Genet. 96, 188–192.
Takahata, N., Satta, Y., 1997. Evolution of the primate lineage leading to modern humans: phylogenetic and demographic inferences from DNA sequences. Proc. Natl. Acad. Sci. U.S.A. 94, 4811–4815.
Tristem, M., 2000. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. J. Virol. 74, 3715–3730.
Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M.I., Kidd, K.K., Lenz, J., 2001. Insertional polymorphisms of full-length endogenous retroviruses in humans. Curr. Biol. 19, 1531–1535.
van der Kuyl, A.C., Dekker, J.T., Goudsmit, J., 1995. Distribution of baboon endogenous virus among species of African monkeys suggests multiple ancient cross-species transmissions in shared habitats. J. Virol. 69, 7877–7887.
Werner, T., Brack-Werner, R., Leib-Mösch, C., Backhaus, H., Erfle, V., Hehlmann, R., 1990. S71 is a phylogenetically distinct human endogenous retroviral element with structural and sequence homology to simian sarcoma virus (SSV). Virology 174, 225–238.
Widegren, B., Kjellman, C., Aminoff, S., Sahlford, L.G., Sjogren, H.O., 1996. The structure and phylogeny of a new family of human endogenous retroviruses. J. Gen. Virol. 77, 1631–1641.