Cell. Mol. Life Sci. 65 (2008) 3348 – 3365
1420-682X/08/213348-18
DOI 10.1007/s00018-008-8495-2

Cellular and Molecular Life Sciences

Evolution of human endogenous retroviral sequences: a conceptual account

V. Blikstad[a], F. Benachenhou[a], G. O. Sperber[b] and J. Blomberg[a,*]

[a] Section of Virology, Department of Medical Sciences, Uppsala University, Academic Hospital, 751 85 Uppsala (Sweden), Fax: +46 18 55 10 12, e-mail: Jonas.Blomberg@medsci.uu.se
[b] Department of Neuroscience, Physiology, Uppsala University, Uppsala (Sweden)

Online First 27 September 2008

Abstract. Endogenous retroviruses (ERVs) most likely are remnants of ancient retroviral infections. ERVs preserve functions of exogenous retroviruses to a varying extent, and can be parasites, symbionts or more or less neutral genetic 'junk'. Their evolution has two facets, pre- and post-endogenization. Although the two are not clearly separated, the first pertains to retroviral evolution in general and the second to the fate of repetitive DNA and the evolution of the host organism and its genome. The study of ERVs provides much material for the understanding of retroviral evolution. This sequence archive reflects the history of successes and shortcomings of antiviral resistance, but also of strategic evolutionary decisions regarding genome organization and new gene acquisition. This review discusses retroviral evolution illustrated through HERVs, bioinformatic prerequisites for ERV studies, the endogenization process and HERV evolution post-endogenization, including relation to disease. (Part of a Multi-author Review)

Keywords. Endogenous retrovirus, evolution, lateral transfer, virulence, vertebrate genome, bioinformatics, sequence recognition.

Introduction

The human genome harbours groups of retrovirus-like sequences (HERVs) that are remnants of infections early during primate evolution. Thus, retroviral genomes have been inherited in a Mendelian manner as endogenous retroviruses. Initially, the provirus is a 'selfish' gene [1] i.e., it is optimizing its own replication and replicates like an exogenous retrovirus in somatic cells, occasionally infecting other individuals, being more or less damaging to the host. Unlike HIV, most ERVs accumulate outside active transcription units; it is suggested that purifying selection somehow dictates their distribution [2]. In general, mobile genetic elements (like ERVs) cause disease by transposition [3]. Retrotransposons have greatly influenced the evolution of the primate genome through transposition, translocation and recombination [4–7]. Due to their age, the majority of the ERVs contain numerous deletions and mutations that compromise or abrogate their function. However, some HERVs in the human genome seem to have retained function. Potentially useful functional modules offered by retroviruses are tissue-specific enhancers [8, 9], polymerase II promoters [10], splice donors [11], splice acceptors [12], a myristylated cytoplasmic membrane protein [13], structural and nucleic acid binding proteins capable of packaging RNA, a dUTPase [14], a protease [15, 16], an RNA-dependent DNA polymerase with RNAse H [17] (the reverse transcriptase may also be involved in the pathogenesis of nonretroviral RNA viruses due to their persistence as DNA copies [18]), an integrase [19], an envelope protein which binds to a host surface protein, a spring-loaded transmembrane protein ready to fuse membranes and possibly cause immunosuppression, and a polyadenylation signal and site [20–22]. The thousands of HERVs can be viewed as a 'proviral quasispecies'.

* Corresponding author.

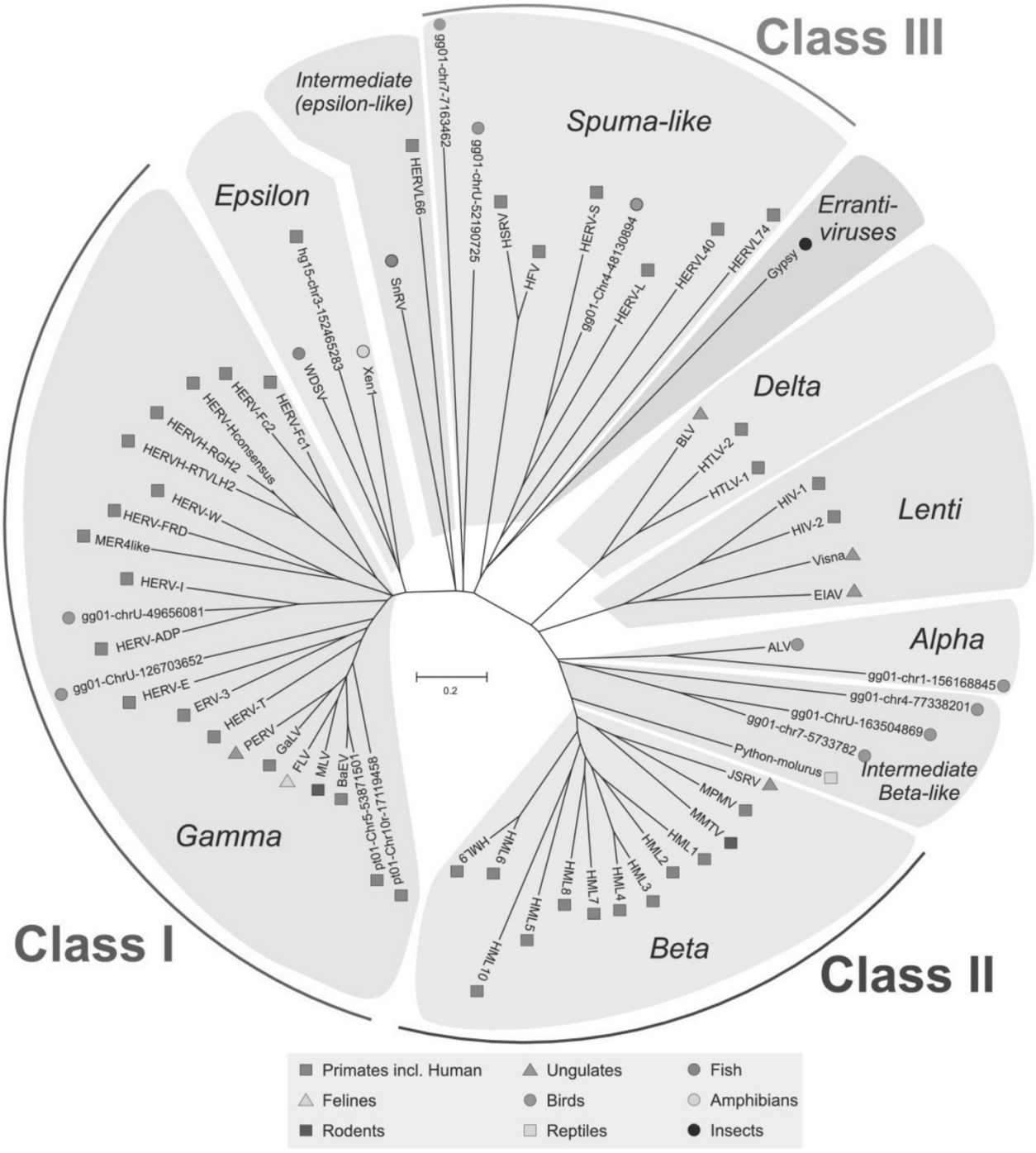

Figure 1. General tree of retroviruses based on the Pol proteins of exogenous and endogenous retroviral sequences. From [32]. Chicken and primate elements occur together in the tree, despite their hosts being separated by 300 million years. Data from a RetroTector (see the section on bioinformatics) analysis of the chicken, human, chimpanzee and dog genomes, and from the literature, as specified in [32].

This genomic quasispecies is not static. From time to time it is modified by gene conversion, which pastes one sequence over a highly similar one [23–25]. Such nonreciprocal transfer of genetic information between homologous sequences is a fundamental mechanism of genome variation and rearrangement. Although the frequency is in general uncertain, a hotspot of gene conversion in the distal inter-AD segment between two endogenous retrovirus HERVs and the paralogous AZFa-repeats on the Y-chromosome has recently been described [26]. The same locus can also suffer major deletions through a related mechanism, unequal crossing over, leading to azoospermia [27]. A number of funda-

mental observations regarding epigenetic control of gene expression were published in the last 10 years. These include methylation and heterochromatic silencing through histone modification, posttranscriptional expression regulation with various forms of inhibitory RNA and nonsense-mediated degradation of untranslatable RNA. Most of these developments are beyond the scope of this brief review. However, it is now clear that there is a regulatory web of expression and replication controls which a retrovirus must overcome. The ubiquity of transposons probably made it necessary for the host to keep their expression and reintegration under control. Although many retrotransposons become epigenetically silenced by methylation [28], their degree of methylation varies [29]. Despite these multilayered controls, HERV RNA is abundant in most cells [30]. Finally, an aspect of virus-host interaction, pathogenicity, is also addressed briefly. Suggestions of links with human disease have been obtained with HERVK(HML2), HERVW, HERVE and HERVH during the past 5–6 years [31].

Retroviral evolution illustrated through ERVs

ERVs provide a glimpse of a large evolutionary retroviral tree of which only some branches have extant exogenous members. Similar to exogenous retroviral (XRV) evolution, the evolution of ERVs depends on the high error rate of the reverse transcriptase, and the recombination caused by its jumps between the two RNA strands of the virus particle. ERVs provide a far more comprehensive view of retroviral evolution than can be obtained from studies on the much fewer extant XRVs [32].

Figure 1 shows a continuum of retroviral evolution. The three major branches of the retroviral tree, the gamma-epsilon, the spuma and the delta-lenti-alpha-beta ones, are all represented in the human genome as HERVs. A classical division of ERVs is Class I (gammaretrovirus-like), Class II (betaretrovirus-like) and Class III (spumaretrovirus-like). Bacteria are of polyphyletic origin, largely due to activities of transposable elements, giving rise to a multitude of overlapping trees depending on the gene under study [33]. Likewise, the retroviral tree is partially incongruent with the vertebrate host tree (a brief version is given in Fig. 2), compatible with occasional lateral transfers.

The ability of retroviral sequences to 'freeze' at multiple evolutionary stages in various host lineages leads to accumulation of a vast range of retroviral structures. Recombination arising from copackaging of such past and present retroviral RNAs may lead to reuse of retroviral structures optimized for viral survival under conditions which occurred many millions of years ago. This is a gene transfer which is 'horizontal' in certain respects, and 'vertical' in others. This evolutionary aspect is rather unique to retroviruses. Many recognizable, i.e. complete or relatively complete, proviral sequences entered the primate lineage during the last 100 million years. For unknown reasons, a discontinuity was introduced at the time of separation of the Old and New Worlds, 30–40 million years ago.

The eight vertebrate genomes shown in Figure 2 have both common and distinct features. Common to all is the content of both gamma- and betaretrovirus-like sequences. Betaretrovirus-like proviruses dominate the mouse ERVs, while humans, chimpanzee and opossum have proportionally more gammaretrovirus-like sequences. Distinct are the occurrences of errantiviral and epsilonretrovirus-like sequences in the zebrafish, and alpharetroviruses in the chicken genome. However, a few erranti-, epsilon- and pseudoretroviral sequences occur also in the human genome [32, 34–38]. Judging from their LTR divergence and mutated interior, they seem to be old acquisitions [Blomberg et al., unpublished].

Recently integrated proviral sequences in the genome of humans and other vertebrates: indirect evidence for horizontal transfer

A pattern of occasional horizontal transfers of XRVs between primates during relatively recent evolutionary time is emerging. Prime examples are HIV-1 and HIV-2 [39–47], HTLV-1, HTLV-2, HTLV-3 and HTLV-4 [48–52], and HSRV [53, 54] . The recent discovery of a mouse retrovirus in low copy number, in some somatic cells, in a subset of human prostatic carcinoma (XMRV) cases could be an indication that this is still ongoing in humans [55].

Judging from similarities of the 26 new chimpanzee integrations to other primate retroviruses (Fig. 3), their most likely origin is from viruses related to the baboon endogenous retrovirus, a replication-competent gammaretrovirus. One new betaretroviruslike integration, due to an HML2 element, occurred during the same period. The new integrations in the human genome consisted of 12 HML2 proviruses and two HERVHs. The former constitute the 'human-specific' HML2 proviruses, described earlier [56–58]. Besides the 12, there are several polymorphic integrations in the human population [59]. These betaretrovirus-like proviruses are intact or almost intact, with open reading frames in the major genes, and have a very low LTR divergence. Replication-competent sequences were generated from them by minor genetic surgery [60, 61]. The new gammaretrovirus-

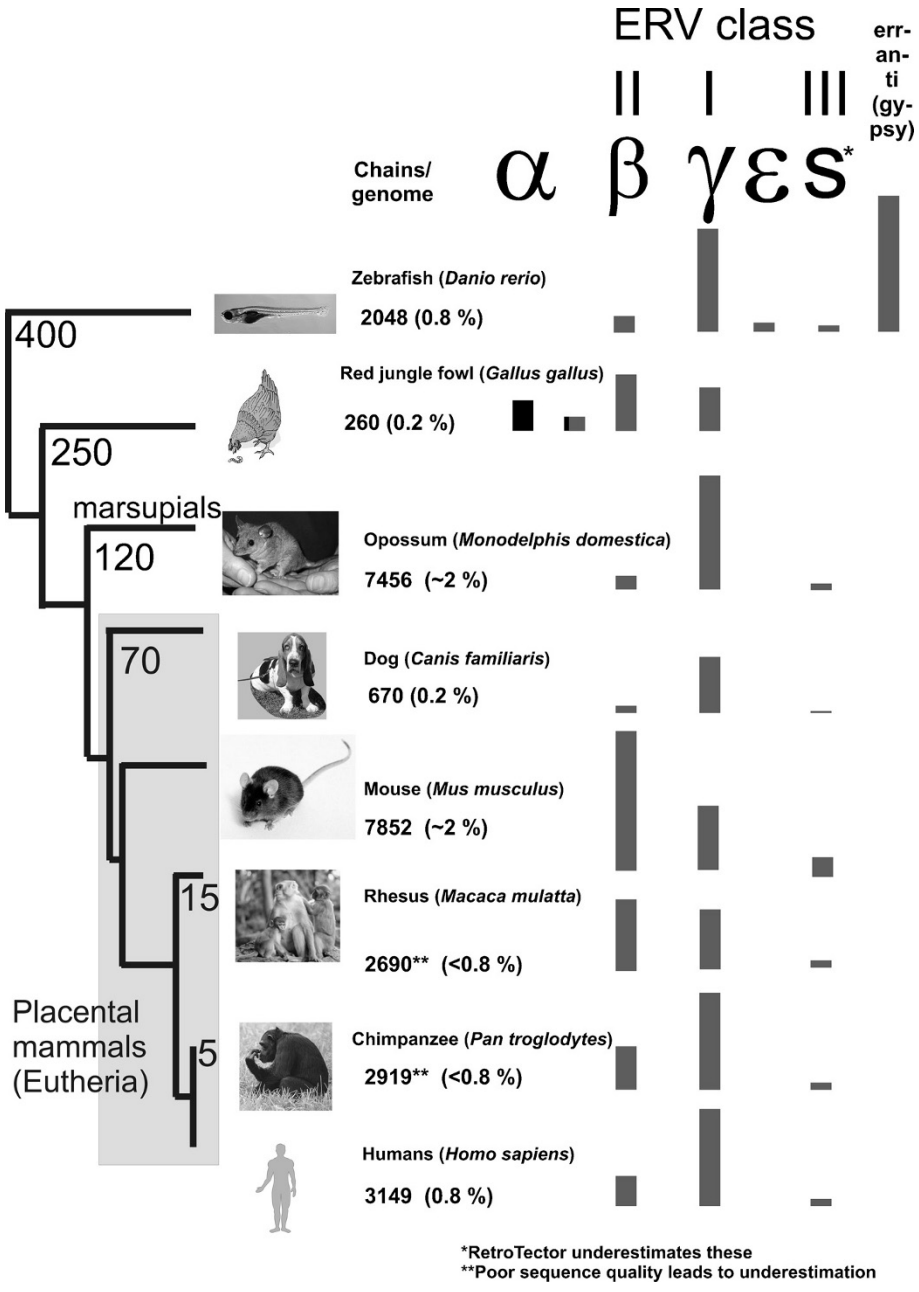

Figure 2. Grand overview of Class I, II and III ERVs in eight vertebrates, as detected by RetroTector (see the section on bioinformatics). RetroTector detects relatively complete proviruses, which gives a lower genomic ERV percentage than other techniques. Note also that class III is underrepresented. Approximate times, in millions of years, of separation of major taxa are shown in the tree. Absolute numbers of detected proviruses are given for each species. The relative distribution of the classes is indicated by bar heights. More exact figures are given in [34].

like integrations are defective HERVHs with low LTR divergence, presumably arising from copackaging into particles provided by less-damaged related, 'midwife', elements [62, 63].

The likely source of the new chimpanzee integrations are baboons and/or other primates. The source of the new human integrations is uncertain. The simplest explanation is that they came from pre-existing HML2 proviruses [55, 64–67]. Recombination and/or gene conversion could have mobilized a replication-competent HML2. However, they may also have come from animals close to humans during this evolutionary period. A bovine endogenous provirus similar to the HMLs, BERV, was recently described [68] (see below). If chimpanzees acquired baboon retroviruses from their prey, it is possible that humans acquired HML2 from prey artiodactyls, like antelopes. In contrast to chimpanzees, in this same time period, humans evolved into bipedalism and high mobility on the savannah, being able to catch fast prey [69]. The ongoing accumulation of sequence information from many species will undoubtedly provide additional surprises.

It was recently reported that the cow genome, on chromosome 5, contains a betaretrovirus-like sequence, here refererred to as BERVK(beta3), with

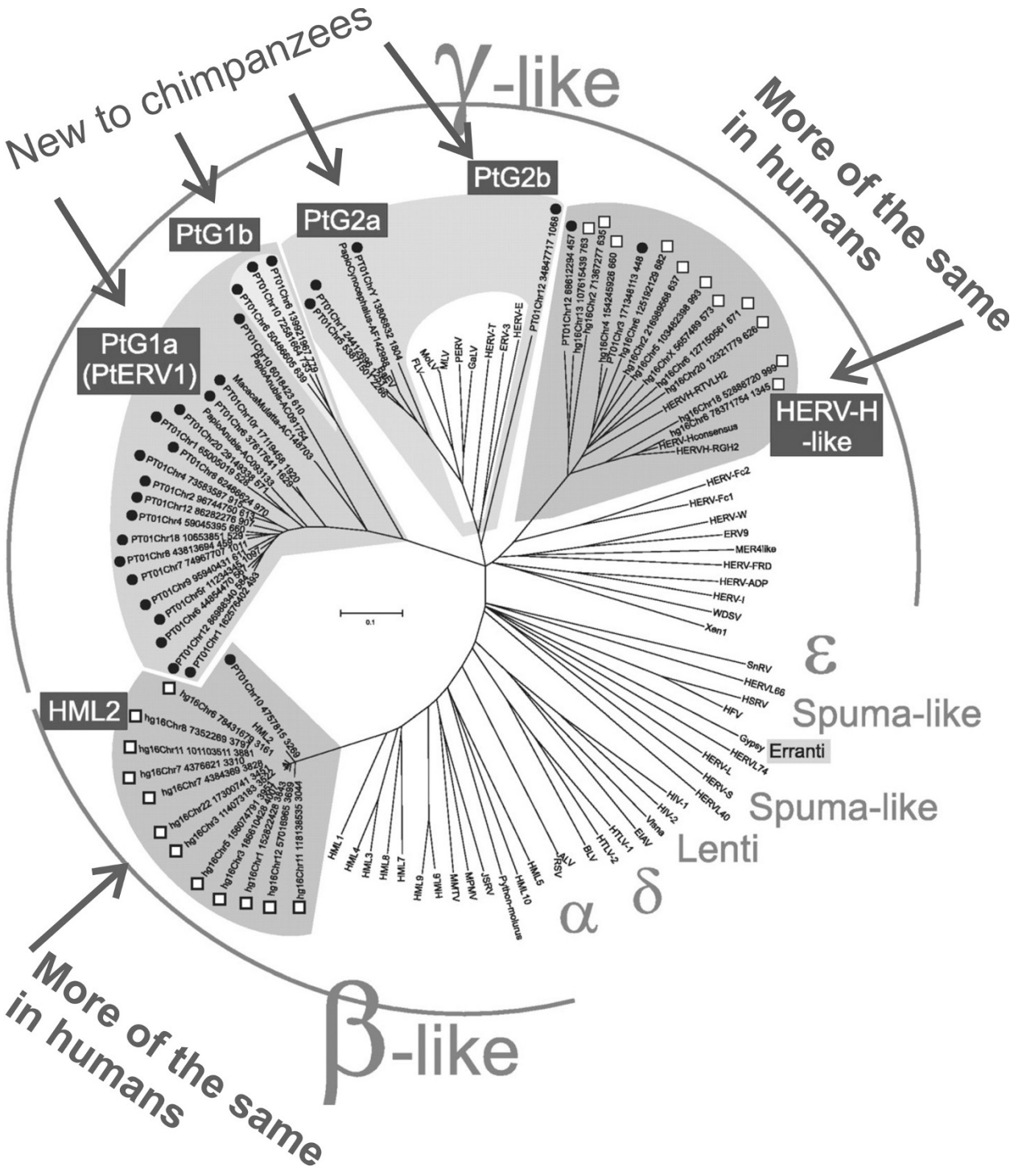

Figure 3. New retroviral integrations in the human (black squares) and chimpanzee (black dots) genomes since the split between the two ca. 5 million years ago. Reference sequences are also included. Reference HERVs are shown as open squares. Modified, from [40].

open reading frames in all major genes. Interestingly, it is intermediate to the highly degenerated HML5 and HML6 elements, the oldest betaretrovirus-like sequences in the human genome [66] (Fig. 4).

Genomes of higher primates, with the exception of orangutan, contain a small group of recently integrated gammaretrovirus-like two-zinc-finger elements, HERVFc [70]. Two of them have *env* ORFs. They must have entered the human genome in recent evolutionary times. Among the most intact of dog ERVs are a related group of HERVFc-like sequences (although they have a PBS complementary to proline

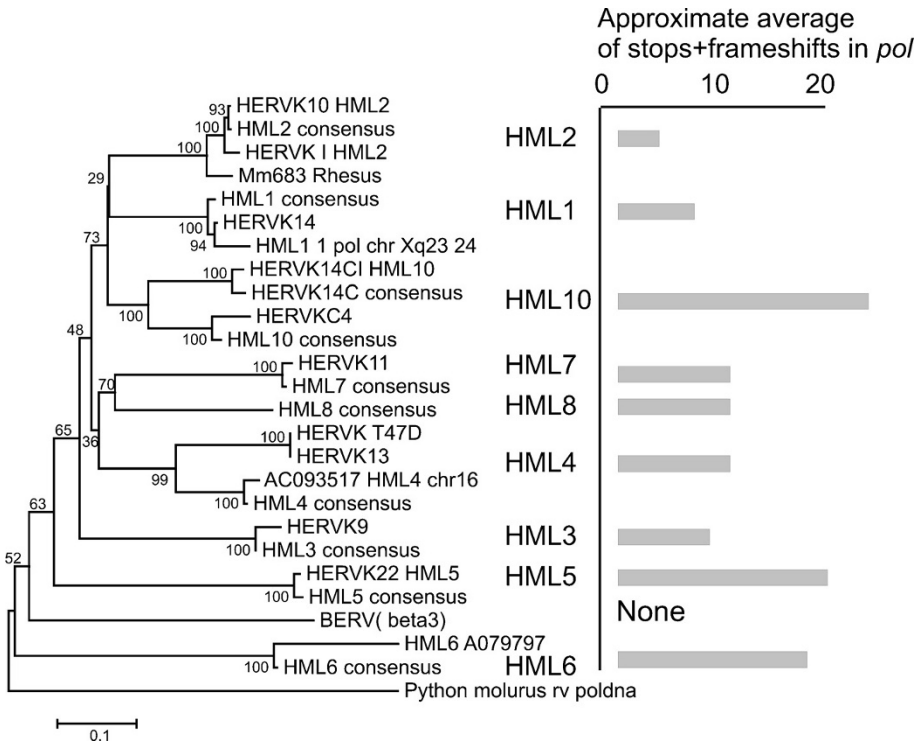

Figure 4. Neighbour-joining tree based on an alignment of *pol* sequences of human betaretrovirus-like (HML) sequences, together with the structurally intact BERVK(beta3) element from cow chromosome 5 [70]. The average of number of stop codons plus number of frameshifts for the *pol* reading frame, based on a RetroTector analysis of selected elements, is shown for each HML group. Constituent sequences were either consensus sequences from RepBase [72] or from a compilation of HML groups [Blikstad, unpublished]. The HML9 group is not represented. The rhesus Mm683 sequence was kindly provided by Jack Lenz. Nodes are shown with % bootstrap value, from 500 iterations. The tree was rooted with the python molurus ERV (GenBank id AF500296).

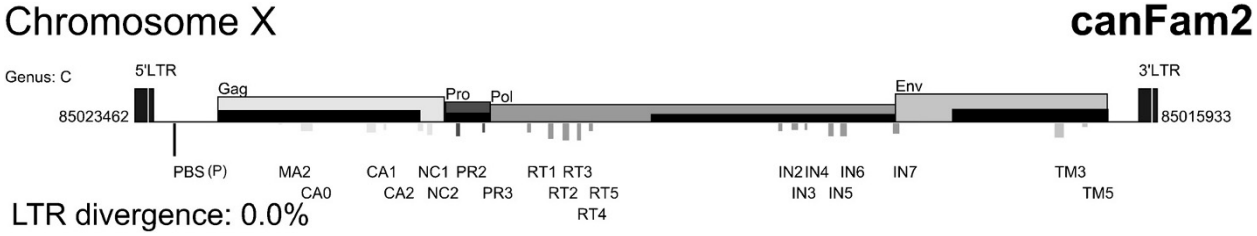

Figure 5. An almost intact provirus, related to HERVFc1, in the dog genome, on chromosome X [Blomberg et al., unpublished; Martinez-Barrio et al., in preparation], as depicted by RetroTector (see the section on bioinformatics). Because of the low LTR divergence and interior intactness it should be a recent acquisition. It has two zinc-finger motifs (NC1 and NC2) in *gag* [32, 65, 73, 74]. Frames with up to one stop and one frameshift are shown as open rectangles. The longest totally open reading frame is indicated by a black bar within the rectangle. Characteristic motifs are shown as vertical bars with a motif label.

tRNA) (Figs. 5 and 7). HERVFc-related viruses thus seem to have been transmitted to several vertebrates recently. One-zinc-finger MLV-related gammaretroviruses seem to have spread even more widely in a comparable time frame, see e.g. this paper and [71, 72].

The opossum genome contains a large number of gammaretrovirus-like one-zinc-finger ERVs related to mouse leukemia virus (MLV) and to HERVT (Fig. 6). Due to a low LTR divergence, absence of stop codons, frameshifts, they probably represent a recent expansion. The most-related elements in the human genome are HERVTs which in most cases are highly degenerated (Fig. 7).

The mutational rate of a neutral sequence is around 0.2 % per million years for many vertebrates. Thus, the given examples of structurally intact or almost intact proviruses, which are similar to intact or decayed proviruses in distant vertebrate species, are compatible with recent occasional horizontal transfers of infectious retroviruses. The alternative explanation, that some proviruses are immune to mutation in some host genomes, is not likely. Some of the infecting XRVs become fixed as endogenous elements. Others continue to undergo a fast evolution as exogenous viruses. A further example of this process has been demonstrated in the rhesus macaque genome where eight proviruses were predicted to be recent acquisitions [73].

In what ways are ERVs different from XRVs? For nearly all ERVs, imperfect structural information inferred from sequence is what is available. The proviral model used by RetroTector (see below) is able to detect ERV sequences in many different

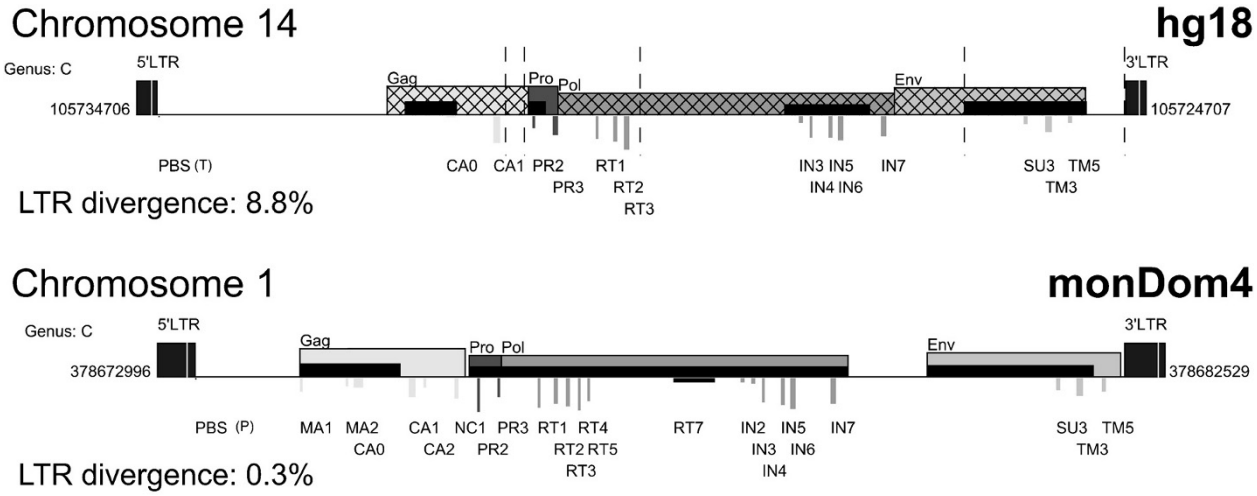

Figure 6. An almost intact HERVT-like sequence in the opossum genome (monDom4) on chromosome 1 [Blomberg et al., unpublished]. From RetroTector (see the section on bioinformatics). A relatively intact HERVT from human chromosome 14 (assembly hg18) is also shown. Reading frames with more than one stop or one framshift are cross-hatched. Otherwise, see the legend of Figure 5.

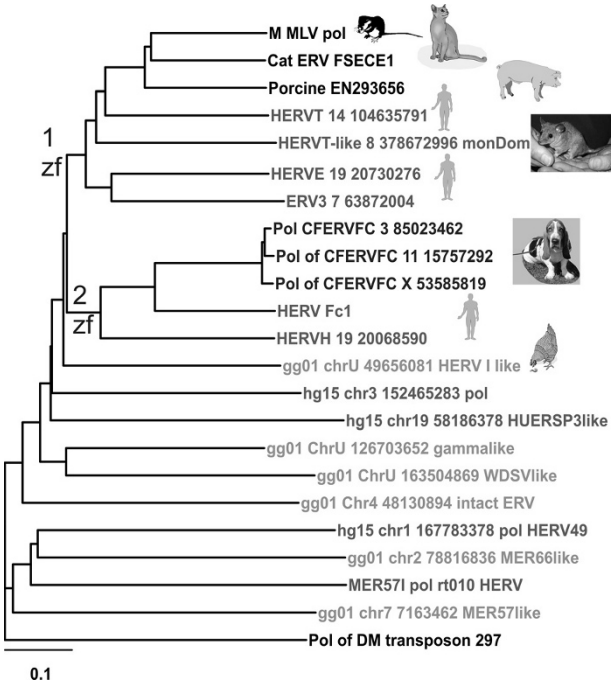

Figure 7. Pol protein similarity based neighbour-joining tree depicting the degree of relationship of selected human, mouse, chicken, dog and opossum gammaretrovirus-like proviruses detected by RetroTector (see the section on bioinformatics). The tree was rooted with the Pol sequence of a *Drosophila* transposon. ERV sequences from chicken genome assembly gg01 are shown in light gray. The recently active *gag* one- and two-zinc finger elements, shown as '1 zf' and '2 zf', respectively, of opossum and dog, respectively, are shown in the context of related elements.

vertebrate genomes. Thus, ERV proviral structure conforms to the proviral structure of known extant XRVs, and there are no indications that ERVs are fundamentally different from XRVs. However, they must fulfill at least two criteria: low virulence [74] and the ability to infect germ line cells and their progenitors. The envelope genes of ERVs therefore define germ line tropism, at least as it was in the past. The great number of HERVs necessitates clustering them into groups. Current taxonomy is mainly based on the tRNA primer preference [75]. Owing to the variability of tRNA usage within an HERV group [76], new taxonomic principles are needed. We will briefly discuss this in the next section, on ERVs and bioinformatics.

The bioinformatic study of ERVs

Techniques for detection of ERV have been built based on several principles.

i. Detection of repetitive sequences. These are collected in Repbase [77], which provides a valuable basis for ERV studies. The program RepeatMasker [78] uses RepBase in genomewide searches of a great variety of species. According to our experience, RepeatMasker efficiently detects most ERVs. It is the *de facto* standard for repetitive element detection. Drawbacks are an inability to detect low copy number elements, a nonsystematic nomenclature and fragmentation of proviral sequences into short repeats.

ii. Detection of retroviruslike structures. LTR_Struc is based on the presence of LTR-like repeats at a predetermined distance [79]. Its focus is the detection of reverse-transcriptase motifs. LTR_Struc can detect many kinds of autonomous LTR and non-LTR retroid elements. It has evolved to also detect some of the internal proviral structure and has been used on several vertebrate genomes. The retroid element detection Genome Parsing Suite (GPS) [80] is a

generic multistep automated process. It is used to detect reverse-transcriptase motifs as well as many kinds of autonomous LTR and non-LTR retroid elements.

The program developed in our group, RetroTector [81], is based on a collection of conserved motifs from *gag*, *pro*, *pol* and *env*, and constraints on intermotif distances of known retroviruses. It also suggests a reconstruction of the proteins of even highly mutated ERV sequences. The detailed structural characterization gives confidence regarding the nature of the detected provirus, and the ability to look for long-range similarity and classify a wide range of retroviruses via protein alignments. The detailed and exact characterization of the proviral contents is unique to RetroTector. A complete analysis of a 3-Gbase genome takes 1–2 days on the Uppmax computer cluster. RetroTector currently misses many of the highly mutated Class III ERV elements. Its modular structure may allow modifications to remedy this shortcoming. Neither is it currently suited for detection of single LTRs. Relatively long stretches (around 5000 bp) of proviral structure are needed for a clearly positive identification. Therefore, sequence quality should be high. The relatively poor quality of the first chicken and chimpanze genome assemblies resulted in significantly fewer detected ERV chains than with subsequent versions [Sperber, Jern, Benachenhou and Blomberg, unpublished]. The availability of a panel of approaches to retrotransposon detection now allows a more complete detection and understanding of this dynamic and dominating genomic component.

iv. Detection of single LTRs. Retroviral LTRs, paired or single, are plastic structures which contribute profoundly to vertebrate genomic and transcriptional diversity [82]. Single LTRs are remnants of past complete proviruses, which underwent homologous recombination betwen their LTRs (Fig. 8). As mentioned earlier, single as well as proviral LTRs, can act as enhancers, promoters (sometimes bidirectional), polyadenylation sites and building blocks for exons (neo-exons).

Like a complete provirus, single LTRs are surrounded by direct repeats. Although they possess a variety of retroviral expression regulatory sequences, these are hard to pinpoint bioinformatically as discrete sequence features. Their most clear internal feature is the AATAAA polyadenylation signal or variations thereof. Single LTR detection is necessary to understand the genomic impact of ERVs. Single LTRs are far more abundant in the host genome than proviruses [83]. The structural features of LTRs are, however, too general for unequivocal LTR identification. RepeatMasker can detect single LTRs by virtue of their multitude, and their occasional presence in complete

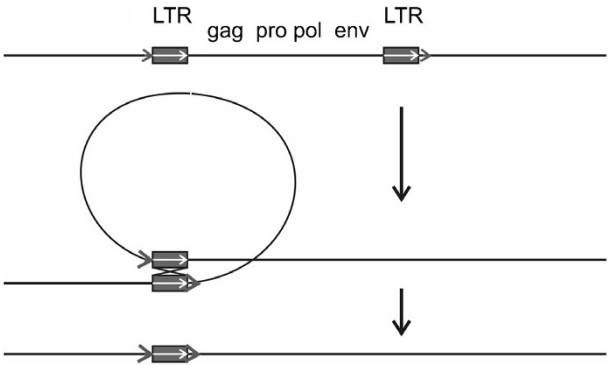

Figure 8. Generation of single LTRs by homologous recombination.

proviruses. RetroTector uses a combination of features to select LTR candidate sequences before attempting to construct proviral chains [81]. We recently found that the specificity can be increased with the help of hidden Markov models trained on known LTRs [Benachenhou, unpublished]. Although an all-pervasive bioinformatic solution seems far away, it may become possible to identify many single LTRs *ab initio*, without reference to related proviral LTRs.

Detection and alignment of single LTRs is a bioinformatic challenge. In an ongoing project [Benachenhou et al., submitted], all or subsets of over 6000 LTRs collected from five genomes (human, chimp, mouse, opossum and chicken) using RetroTector, were used for training of Hidden Markov Models (HMMs). The more narrowly targeted HMMs had a high sensitivity and specificity, but also had a limited detection range. One of the narrowly HML-specific HMMs could detect 90% of all known HML LTRs, with a false positive rate (compared to RepeatMasker) of 10%. More general HMMs could detect around 20% of LTRs from a broad collection of genomes, from zebrafish, chicken, opossum, mouse, dog, rhesus, chimpanzee and humans. Apart from this novel source of information, LTRs are annotated through the RepBase/RepeatMasker system. Around 1% of the human genome consists of single LTRs, and around 7% of largely damaged proviruses [83]. Whether this figure is inclusive of all LTRs is unknown. The RepBase/RepeatMasker system is based on multiple ocurrences of a sequence, and will therefore miss low copy number single LTRs. It may also misclassify LTRs. However, the experience from our HMM-based detection systems, which are repeat-independent, does not indicate this to be a common problem.

v. The bioinformatic use of LTR divergence. A special property of retroviral sequences is that their LTRs are identical at the time of integration. Following the

gradual divergence of LTRs post-endogenization is thus a convenient way of following the mutational decay of proviruses. However, this parameter is compromised by several factors. Gene conversion may obscure the pattern of accumulated mutations. Mobilization of old and defective elements by packaging into particles formed by more intact elements leads to a newly integrated, but internally mutated, provirus with low LTR divergence [55, 63, 84]. Moreover, genomic positions may differ in mutational rate. Inference of element age on the basis of LTR divergence alone therefore should be complemented by estimates of the degree of internal mutation. RetroTector provides this information.

vi. Classification of HERVs, taxonomy. It is reasonable to base a viral taxonomy on viral pheno- and genotypic properties. Because these are undisturbed in the virus prior to endogenization, ERV taxonomy should be based on the pre-endogenized virus. For old ERVs this ideal is hard to live up to. Obscure structure leads to weakly founded functional inference and classification. Viral taxonomies can be based on sequence divergence, structural features like presence of dUTPase, inferred functional features like translational strategy (gag-pro-pol frameshifting) and phenotypic properties such as host preference [32, 85–92]. In order to get a comprehensive view, ERVs of many different hosts, and XRVs, should be considered. HERV evolution and HERV taxonomy are only special cases of general retroviral evolution and taxonomy. An optimal retroviral taxonomy has yet to be constructed, in large part because of its relationship to the larger question of transposon classification [93].

The endogenization process

The endogenization process (Fig. 9) is dependent on several factors. These include the ability of the virus to withstand restrictions common to all cells, to infect reproductive tissues in order to reach the germ line, to infect germ cell progenitors, to allow the infected germ line cell survive, to let the progeny organism to survive without loss of fitness and finally to proceed to fixation, and survive population bottlenecks, in the host species. A highly pathogenic ERV cannot go through this fine-meshed sieve.

It is logical that HERVs are highly expressed in reproductive tissue, as this is how they were introduced into the germ line originally. As further described below, Hu looked at expression in many different human tissues with the help of real-time PCRs for gammaretrovirus-like and betaretrovirus-like HERV groups [30, 94–96]. An especially high expression was seen in reproductive tissue.

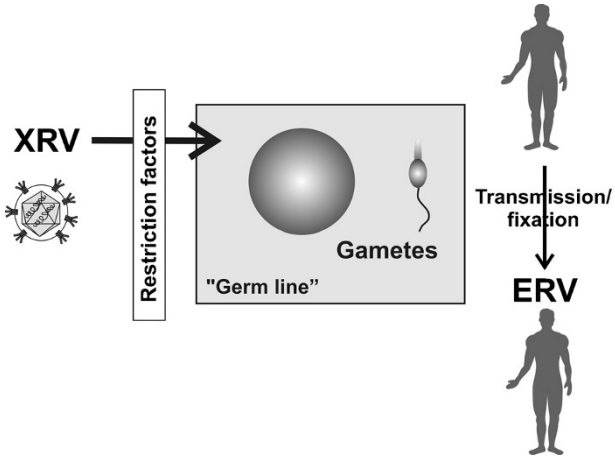

Figure 9. The endogenization process. Before an exogenous retrovirus (XRV) can reach the germ line, it is prevented by an array of antiretroviral restrictions. Once the germ line cells or their progenitors have been successfully integrated, the progeny must be fit, and the new trait must be fixed, before the endogenous retrovirus (ERV) is established.

Obstacles to endogenization

An incoming infectious retrovirus can vary in its pathogenic potential, and it can be of vital importance for the organism to defend itself. The ubiquity of retrotransposons in most genomes indicates that the interaction between organism and 'selfish' nucleic acid is a profound aspect of both host and viral evolution.

The interactions with defenses against transposons and foreign nucleic acid in general are powerful selective forces which profoundly influence the evolution of retroviruses (Fig. 10). Retroviral restriction factors have been known since the classical work of Lilly, see e.g. [97]. Restrictions may be due to other ERVs, trans-dominant Gag interference, FV1 [98] or enJSRV [99, 100], and cell surface receptor interference [101, 102]. Thus, the previous ERV protects the genome against ensuing similar XRVs. APOBEC3G is a cytosine deaminase which can be encapsidated and hypermutate the new negative cDNA strand inside the virion during reverse transcription. It restricts retroviruses in many hosts and cell types [103, 104]. TRIM5alpha is another host protein which probably ubiquitinylates the newly incoming preintegration complex, which may send it to degradation at the proteasome [105–110]. Other defenses against retroviruses are shown in Figure 10. The field of RNA-mediated transposon control is too large and too rapidly evolving to be covered here, see [111, 112]. All these restriction factors are obstacles that must be overcome by an ERV. The same restriction factors may also restrict XRVs, like HIV. Thus, ERV research can indicate defects in the antiretroviral armour existing in a host species at the time of endogenization.

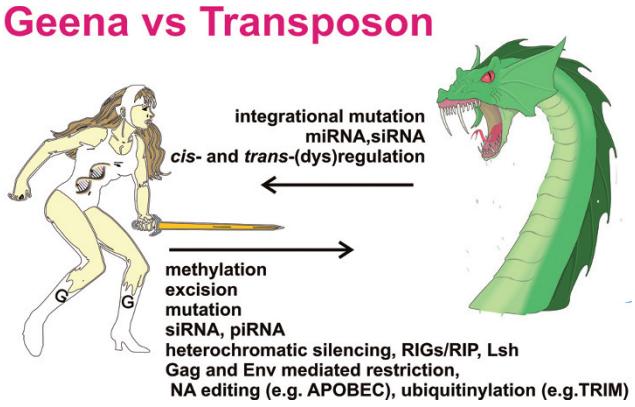

Figure 10. Retroviral restrictions. An allegorical rendition of an imaginary fight between target cell defenses of the organism ('Geena') and parasite (dragon) at the genetic battlefield. A collection of measures and countermeasures of the two combatants is shown. They are not all present in all hosts and in all parasites. Abbreviations: RIGs, repeat-induced gene silencing [98, 99]; RIP, repeat-induced point mutation [100, 101]; Lsh, lymphoid-specific helicase [102]; miRNA, micro inhibitory RNA; siRNA, short interfering RNA; piRNA, piwiRNA. Further explanations are given in the text.

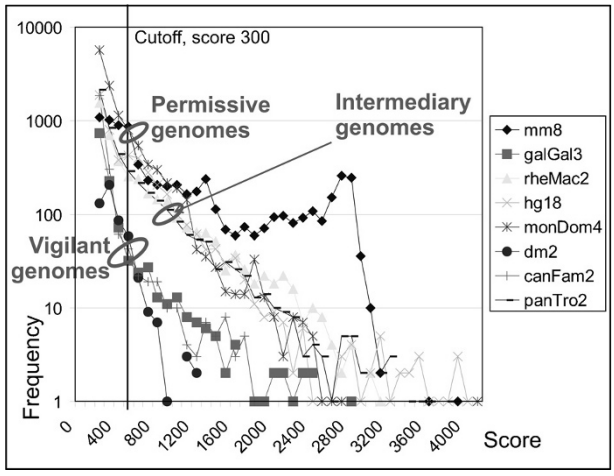

Figure 11. Distribution of RetroTector chain scores in eight genome assemblies [34, and Blomberg, Jern and Sperber, unpublished]. The chain score is a measure of degree of fit to the retroviral structural model used by RetroTector. The chains of the genomes segregate into three groups with different degrees of permissiveness of the species to retroviral integration. Abbreviations: mm, *Mus musculus*, mouse; galGal, *Gallus gallus*, chicken; rheMac, Rhesus macaque; hg, human genome; monDom, *Monodelphis domestica*, opossum; dm, *Drosophila melanogaster*, fruit fly; canFam, *Canis familiaris*, dog; panTro, *Pan troglodytes*, chimpanzee.

Certain genomes are more vigilant in their antiretroviral defenses. In a recent study using the RetroTector proviral detection system, Blomberg et al. [in preparation] and Martinez-Barrio et al. [in preparation] found that dog and chicken genomes contained approximately five times fewer relatively complete proviruses than the human, mouse and opossum genomes. ERV content of various genomes as detected by RetroTector is shown in Fig. 11.

HERV evolution post-endogenization

Functional role of LTRs

Retroviral LTRs, paired or single, are plastic structures which contribute profoundly to vertebrate genomic and transcriptional diversity. A number of examples demonstrate that single or proviral LTRs can act as enhancers [9, 113–115], promoters (sometimes bidirectional) [10, 22, 116–123], polyadenylation sites [124] and building blocks for protein sequence [125, 126].

HERV polymorphisms

These may be a sign of proliferation, i.e. retroviral function, within a cell, an organism and within a host species. There are at least three kinds of polymorphisms: insertional polymorphism [59, 127, 128], variable tandem repeats [7, 129] and sequence polymorphism due to single-nucleotide polymorphisms (SNPs), see e.g. [130].

Copy number variation can occur after unequal crossing over (Fig. 12). Copy number variation in repetitive sequences seems to be a major source of genetic difference between individual humans [131–133]. Gain or loss of tandem repeats also affects HERVs [7, 65, 134, 135].

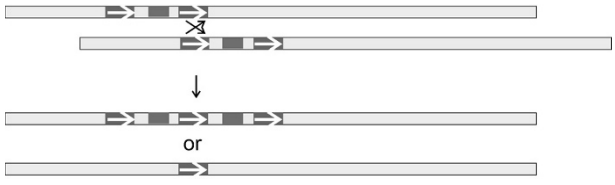

Figure 12. Generation or loss of chromosomal sequence by unequal homologous crossing over between repetitive sequences during meiosis.

HERVs: some functional considerations

Antisense signalling

Transposable elements constitute half the human genome [83]. Both LINE elements and several HERVs have bidirectional promoters [4, 111, 123, 136–145]. This is a natural source of antisense RNA which may regulate both transposon expression and any gene in the vicinity which becomes included into a complementary transposon-promoted transcript. Thus, HERVs may be part of an intricate transcriptional network [146–148].

Expression of HERVs

A high transposon expression in germ line cells, and in adjacent tissue, i.e. reproductive tissue, is common in

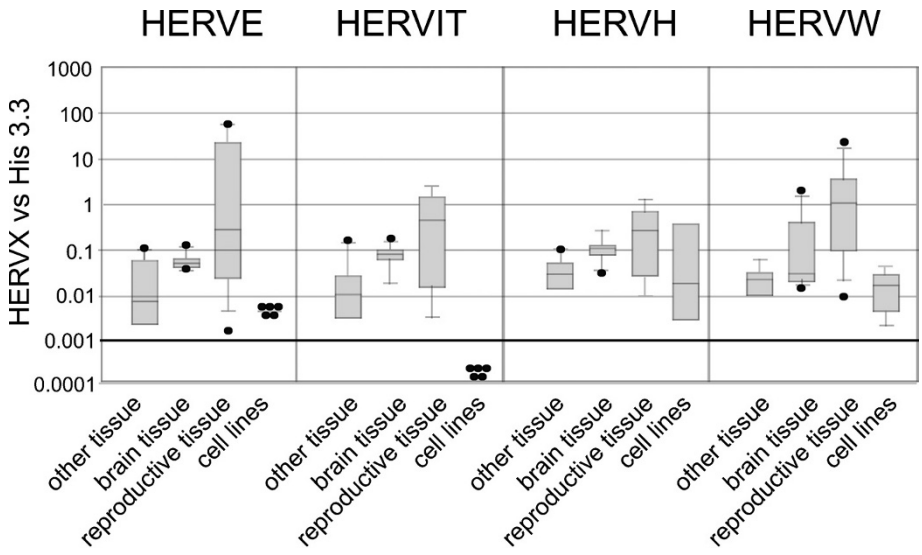

Figure 13. Overview of expression of Class I HERVs in different tissues [30], available at the URL http://publications.uu.se/theses/abstract.xsql?dbid=8213. Expression was measured with four broadly targeted real-time PCRs for gammaretrovirus-like HERVs [73]. The dotted line indicates the approximate position of one HERV RNA equivalent per cell (low-abundance transcripts). Some reproductive tissues can have a 10000-fold higher expression than this. Expression is normalized versus histone 3.3 RNA [30, 73, 96, 97]. Histone 3.3 expression is approximately evenly expressed at around 1000 copies per cell. One HERV RNA equivalent is calculated from a standard curve of an appropriate reference plasmid DNA. PCR efficiencies vary for different target ERV sequences. Hence, an equivalent may represent one or several copies of an aberrant target ERV DNA. 'Reproductive tissues' are placenta, ovary, endometrium and testis. 'Brain tissue' are brain (different areas) and cerebellum. 'Other tissue' includes adrenals, bone marrow, colon, small intestine, kidney, prostate, salivary gland, heart, spleen, skeletal muscle, thymus and thyroid glands. 'Cell lines' are A549, human fibroblasts, RD, VERO A6 and GMK AH1. The latter two are of primate, but not of human, origin. From [30].

many animals, see e.g. [149]. HERVs seem to be no exception (Fig. 13). A simplistic view is that ERVs had to go through several somatic replication rounds in reproductive tissue before reaching the germ line, and therefore have a natural tropism for these tissues. The germ line is guarded against transposons by several mechanisms (Fig. 11). A primary line of defense is the piwi (rasi) RNA pathway, which is specifically expressed in germ line cells [112, 149]. This provides a strong selection barrier for incoming XRVs before they can become ERVs. From our experience with reverse transcription real-time PCRs for expression of gamma- and betaretrovirus-like HERVs, it is evident that many HERV loci, even highly mutated ones, are expressed. This parallels the recent findings of the ENCODE project, which indicate that many more genes than previously understood are expressed, albeit at a low frequency [150]. The functions of the low copy number defective ERV RNAs are uncertain. One suggestion is that they might be a training set for antiretroviral RNA functions, preventing similar XRVs from entering a new host, especially its germ line.

Frequency of ORFs or near-ORFs

A central question for the understanding of late events during ERV evolution is the balance between cooption and decay. How and why are open reading frames kept open? Any element with a high LTR sequence divergence, a high defectiveness of other retroviral genes, and a gene which is open from translational start, with a high portion of synonymous mutation, to a significant portion of its length, is a candidate for having been coopted for a protein function. With this definition, random mutation will favour *gag* and *env* ORFs, and disfavour *pro* and *pol* ones, because of their downstream position in the transcript. Another factor is gene length. Random mutation, even without selection, will tend to spare the small *pro* and damage the large *pol* gene. RetroTector reports stops and frameshifts. A definition of open or near-open frames (near)ORFs with 0–1 stops and 0–2 frameshifts shows that the human genome (hg18 assembly) contains 131 Gag, 921 Pro, 27 Pol and 146 Env (near)ORFs [Blomberg et al., unpublished]. HERV groups with (near)ORFs were HERVW, HERVH, HERVF, HERVT, HERVE, HML1, HML2 (of which many human-specific), HML3 and HML6. Genome-wide searches for retroviral ORFs, sometimes combined with cloning and the protein truncation test, have been made by others [151–154]. However, the exact identification of all of them is not yet complete. Deficiences of current genomic assemblies make it hard to predict ORFs with absolute certainty. The relatively high frequency of (near)ORFs raises the possibility of additional complete ORFs in some

individuals. It is difficult to dissociate assembly and sequencing errors from true interindividual variation. Therefore, PCR analysis of ORF loci of several individuals, with a protein translation test and studies on protein expression in cell lines and *in vivo*, are desirable.

Virulence of ERVs

Adaptation of an XRV to ERV existence should involve a decreased pathogenicity. Eiden et al. found that endogenization of Koala retrovirus (KoRV) was accompanied by an envelope mutation which reduced the ability to cause syncytia [74]. This is reminiscent of the minor mutations in H5 influenza hemagglutinin which govern the highly pathogenic phenotype. Virulence switches are a fundamental property of pathogens (Fig. 14). ERVs probably are the most persistent of all transmissible agents. They are therefore expected to have a low pathogenicity. The experience so far is that generally ERVs are of low pathogenicity. There are no straightforward examples of human diseases caused by ERVs, contrasting to the evident association of HIV with AIDS [155] and HTLV-I with adult T-cell leukemia (ATL) [156], which are fatal infections, but often after a delay of many years which allows the virus to transmit before death of the host. HIV transgenic mice have a low fitness, if they survive the fetal stage [157]. Indeed, lentiviruses, which have a set of host-cell regulating *trans*-active proteins, endogenize only rarely [158]. However, this reasoning does not mean that ERVs are completely apathogenic. The experience from mice, chicken, felines and sheep shows that newly endogenized ERVs may retain a considerable pathogenicity. Moreover, there are genetic diseases which are caused by mutations occurring generations ago. Thus, the removal of disease-causing genes by the evolutionary 'selection filter' is imperfect. The same imperfection should pertain to ERVs. Indeed, a number of observations indicate an association between the most recently endogenized ERVs, the human-specific subset of the HML2 betaretrovirus-like ERVs, and seminoma [159–166]. We therefore still think that it is worthwhile to look for disease associations of ERVs.

HERVs and disease

This subject was covered in a previous review [167]. However, a few new findings may contribute to the discussion on HERV evolution. *Multiple Sclerosis (MS)* is a disease with bouts of neurological impairment. HERVW and HERVH have been implicated [168–174]. The syncytin-1 encoding ERVWE1 locus on chromosome 7 may be selectively expressed in MS [175], and may contribute to cytotoxicity in lesions [176]. However, a causal link between HERVW and

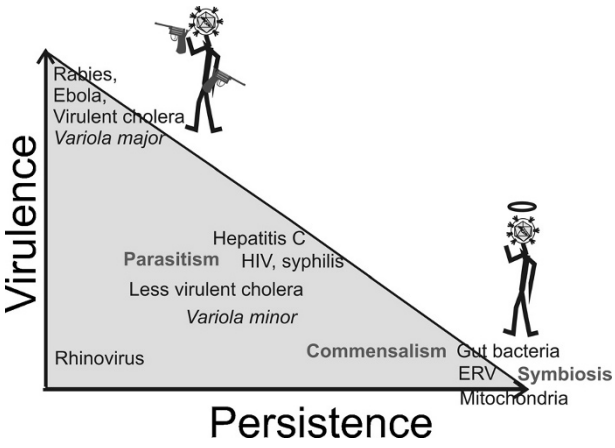

Figure 14. Approximate relation between virulence (the tendency of the microbe to cause disease) and persistence. The longer the coexistence of a microbe (depicted as 'Mr Microbe') in a host, the stronger the selection against the tendency to cause disease. Among the most pathogenic (parasitic) microbes, *Variola major* was a lethal form of smallpox, while *Variola minor* was a milder, often endemic form of smallpox. Rhinovirus is a cause of common cold. HIV and syphilis are sexually transmitted diseases which initially are harmless, but then more or less rapidly become deadly. Likewise, hepatitis C often causes little short-term disease, but can have deadly consequences later in life. Commensals are microbes which may colonize a locale without normally causing disease, but may turn virulent under certain cirumstances. Gut bacteria and mitochondria are clear examples of bacteria with a very long persistence and are obviously symbiotic with the host. This view of host-microbe interaction is approximately the one promoted by Ewald [172].

MS is not established. During post-integrational HERVW evolution, one of the proviruses found a function useful for the organism before its sequence and promoter were damaged by mutation. However, the envelope function was previously optimized for use in a virus, not necessarily for a eukaryotic organism. Beneficial effects (syncytiotrophoblast formation in the human placenta) may have had to be traded against infrequent deleterious effects. This situation is reminiscent of the sickle cell anemia trait, which confers relative resistance against malaria. There may exist a paradoxical dualism during exaptation of ERV mechanisms. Positive effects on the host may overshadow negative ones. Thus, exapted proviruses could be especially prone to also cause disease. Regarding the HERV-cancer connection, malignant melanoma cells in culture were found to produce viral particles which contain RNA and regulatory proteins from HERVK (a sequence which belongs to the HML2 group of betaretrovirus-like sequences) [177, 178]. HML2- and HML6-encoded [179, 180] envelope antigens occur as tumour antigens on the surface of malignant melanoma cells. The *rec* gene, encoded by HML2, was found to give a high frequency of germ cell tumours in mice transgenic for it [165]. The Rec protein may thus promote growth of certain germ line

cells, and is a potential virally encoded oncogene. Immune recognition of a HERVE-encoded kidney tumour antigen was related to tumour regression [181]. A HERVH transcript was highly and selectively expressed in bladder carcinoma [182].

Other diseases

Two single-nucleotide polymorphisms in a retroviral dUTPase were strongly associated with psoriasis, a chronic inflammatory skin disease involving keratinocyte differentiation [183]. dUTPase is a gene typical of betaretroviruses [32, 184] which is colinear and coexpressed with the retroviral protease. dUTPase is involved in nucleic acid metabolism [185]. The findings indicate that this HML-encoded enzyme can somehow influence skin differentiation. The subject merits further investigation.

Concluding remarks

HERVs mirror exposure of our progenitors to various exogenous retroviruses, and the viral ability to penetrate anti-transposon and anti-retroviral defences. Retroviral evolution is a product of properties of both virus (transmissibility, virulence and ability to endogenize) and of host (fitness, habitat, habits and encounters with other animals). Although they often follow Mendelian rules, retroviruses have a unique potential to criss-cross between hosts and evolutionary stages, due to their endogenous and exogenous phases. Decayed HERVs, which entered the germ line many millions of years ago, can have close ERV relatives in other vertebrates which are structurally intact.

1 Dawkins, R. (1976) The Selfish Gene, Oxford Univerity Press. Oxford.
2 Barr, S. D., Leipzig, J., Shinn, P., Ecker, J. R. and Bushman, F. D. (2005) Integration targeting by avian sarcoma-leukosis virus and human immunodeficiency virus in the chicken genome. J. Virol. 79, 12035–12044.
3 Kazazian, H. H., Jr. and Moran, J. V. (1998) The impact of L1 retrotransposons on the human genome. Nat. Genet. 19, 19–24.
4 Baban, S., Freeman, J. D. and Mager, D. L. (1996) Transcripts from a novel human KRAB zinc finger gene contain spliced Alu and endogenous retroviral segments. Genomics 33, 463–472.
5 Goodchild, N. L., Freeman, J. D. and Mager, D. L. (1995) Spliced HERV-H endogenous retroviral sequences in human genomic DNA: evidence for amplification via retrotransposition. Virology 206, 164–173.
6 Jamain, S., Girondot, M., Leroy, P., Clergue, M., Quach, H., Fellous, M. and Bourgeron, T. (2001) Transduction of the human gene FAM8A1 by endogenous retrovirus during primate evolution. Genomics 78, 38–45.
7 Belshaw, R., Watson, J., Katzourakis, A., Howe, A., Woolven-Allen, J., Burt, A. and Tristem, M. (2007) Rate of recombinational deletion among human endogenous retroviruses. J. Virol. 81, 9437–942.
8 Ling, J., Pi, W., Bollag, R., Zeng, S., Keskintepe, M., Saliman, H., Krantz, S., Whitney, B. and Tuan, D. (2002) The solitary long terminal repeats of ERV-9 endogenous retrovirus are conserved during primate evolution and possess enhancer activities in embryonic and hematopoietic cells. J. Virol. 76, 2410–2423.
9 Ting, C. N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C. and Meisler, M. H. (1992) Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes Dev. 6, 1457–1465.
10 Medstrand, P., Landry, J. R. and Mager, D. L. (2001) Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. J. Biol. Chem. 276, 1896–1903.
11 Kato, N., Pfeifer-Ohlsson, S., Kato, M., Larsson, E., Rydnert, J., Ohlsson, R. and Cohen, M. (1987) Tissue-specific expression of human provirus ERV3 mRNA in human placenta: two of the three ERV3 mRNAs contain human cellular sequences. J. Virol. 61, 2182–2191.
12 Feuchter-Murthy, A. E., Freeman, J. D. and Mager, D. L. (1993) Splicing of a human endogenous retrovirus to a novel phospholipase A2 related gene. Nucleic Acids Res. 21, 135–143.
13 Juang, S. H., Huang, J., Li, Y., Salas, P. J., Fregien, N., Carraway, C. A. and Carraway, K. L. (1994) Molecular cloning and sequencing of a 58-kDa membrane- and microfilament-associated protein from ascites tumor cell microvilli with sequence similarities to retroviral Gag proteins. J. Biol. Chem. 269, 15067–15075.
14 Harris, J. M., McIntosh, E. M. and Muscat, G. E. (1999) Structure/function analysis of a dUTPase: catalytic mechanism of a potential chemotherapeutic target. J. Mol. Biol. 288, 275–287.
15 Mueller-Lantzsch, N., Sauter, M., Weiskircher, A., Kramer, K., Best, B., Buck, M. and Grasser, F. (1993) Human endogenous retroviral element K10 (HERV-K10) encodes a full-length gag homologous 73-kDa protein and a functional protease. AIDS Res. Hum. Retroviruses 9, 343–350.
16 Towler, E. M., Gulnik, S. V., Bhat, T. N., Xie, D., Gustschina, E., Sumpter, T. R., Robertson, N., Jones, C., Sauter, M., Mueller-Lantzsch, N. et al. (1998) Functional characterization of the protease of human endogenous retrovirus, K10: can it complement HIV-1 protease? Biochemistry 37, 17137–17144.
17 Berkhout, B., Jebbink, M. and Zsiros, J. (1999) Identification of an active reverse transcriptase enzyme encoded by a human endogenous HERV-K retrovirus. J. Virol. 73, 2365–2375.
18 Klenerman, P., Hengartner, H. and Zinkernagel, R. M. (1997) A non-retroviral RNA virus persists in DNA form. Nature 390, 298–301.
19 Kitamura, Y., Ayukawa, T., Ishikawa, T., Kanda, T. and Yoshiike, K. (1996) Human endogenous retrovirus K10 encodes a functional integrase. J. Virol. 70, 3302–3306.
20 Baust, C., Seifarth, W., Germaier, H., Hehlmann, R. and Leib-Mosch, C. (2000) HERV-K-T47D-Related long terminal repeats mediate polyadenylation of cellular transcripts. Genomics 66, 98–103.
21 Mager, D. L. (1989) Polyadenylation function and sequence variability of the long terminal repeats of the human endogenous retrovirus-like family RTVL-H. Virology 173, 591–599.
22 Mager, D. L., Hunter, D. G., Schertzer, M. and Freeman, J. D. (1999) Endogenous retroviruses provide the primary polyadenylation signal for two new human genes (HHLA2 and HHLA3). Genomics 59, 255–263.
23 de Parseval, N., Casella, J., Gressin, L. and Heidmann, T. (2001) Characterization of the three HERV-H proviruses with an open envelope reading frame encompassing the immunosuppressive domain and evolutionary history in primates. Virology 279, 558–569.

24 Johnson, W. E. and Coffin, J. M. (1999) Constructing primate phylogenies from ancient retrovirus sequences. Proc. Natl. Acad. Sci. USA 96, 10254–10260.

25 Liao, D., Pavelitz, T., Kidd, J. R., Kidd, K. K. and Weiner, A. M. (1997) Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the RNU2 locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion. EMBO J. 16, 588–598.

26 Bosch, E., Hurles, M. E., Navarro, A. and Jobling, M. A. (2004) Dynamics of a human interparalog gene conversion hotspot. Genome Res. 14, 835–844.

27 Kamp, C., Hirschmann, P., Voss, H., Huellen, K. and Vogt, P. H. (2000) Two long homologous retroviral sequence blocks in proximal Yq11 cause AZFa microdeletions as a result of intrachromosomal recombination events. Hum. Mol. Genet. 9, 2563–2572.

28 Yoder, J. A., Walsh, C. P. and Bestor, T. H. (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. 13, 335–340.

29 Reiss, D., Zhang, Y. and Mager, D. L. (2007) Widely variable endogenous retroviral methylation levels in human placenta. Nucleic Acids Res. 35, 4743–4754.

30 Hu, L. (2007) Expression of Endogenous Retroviruses in Humans, thesis, Uppsala University, Uppsala.

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

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

33 Woese, C. R. (2004) A new biology for a new century. Microbiol. Mol. Biol. Rev. 68, 173–186.

34 Butler, M., Goodwin, T., Simpson, M., Singh, M. and Poulter, R. (2001) Vertebrate LTR retrotransposons of the Tf1/sushi group. J. Mol. Evol. 52, 260–274.

35 Lynch, C. and Tristem, M. (2003) A co-opted gypsy-type LTR-retrotransposon is conserved in the genomes of humans, sheep, mice, and rats. Curr. Biol. 13, 1518–1523.

36 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.

37 Herniou, E., Martin, J., Miller, K., Cook, J., Wilkinson, M. and Tristem, M. (1998) Retroviral diversity and distribution in vertebrates. J. Virol. 72, 5955–5966.

38 Martin, J., Herniou, E., Cook, J., Waugh O'Neill, R. and Tristem, M. (1997) Human endogenous retrovirus type I-related viruses have an apparently widespread distribution within vertebrates. J. Virol. 71, 437–443.

39 Bailes, E., Gao, F., Bibollet-Ruche, F., Courgnaud, V., Peeters, M., Marx, P. A., Hahn, B. H. and Sharp, P. M. (2003) Hybrid origin of SIV in chimpanzees. Science 300, 1713.

40 Keele, B. F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M. L., Bibollet-Ruche, F., Chen, Y., Wain, L. V., Liegeois, F. et al. (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science 313, 523–526.

41 Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B. H., Wolinsky, S. and Bhattacharya, T. (2000) Timing the ancestor of the HIV-1 pandemic strains. Science 288, 1789–1796.

42 Santiago, M. L., Range, F., Keele, B. F., Li, Y., Bailes, E., Bibollet-Ruche, F., Fruteau, C., Noe, R., Peeters, M., Brookfield, J. F., Shaw, G. M., Sharp, P. M. and Hahn, B. H. (2005) Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. J. Virol. 79, 12515–12527.

43 Sharp, P. M., Bailes, E., Chaudhuri, R. R., Rodenburg, C. M., Santiago, M. O. and Hahn, B. H. (2001) The origins of acquired immune deficiency syndrome viruses: where and when? Philos. Trans. R. Soc. Lond. B Biol. Sci. 356, 867–876.

44 Sharp, P. M., Shaw, G. M. and Hahn, B. H. (2005) Simian immunodeficiency virus infection of chimpanzees. J. Virol. 79, 3891–3902.

45 Souquiere, S., Bibollet-Ruche, F., Robertson, D. L., Makuwa, M., Apetrei, C., Onanga, R., Kornfeld, C., Plantier, J. C., Gao, F., Abernethy, K. et al. (2001) Wild Mandrillus sphinx are carriers of two types of lentivirus. J. Virol. 75, 7086–7096.

46 Wain, L. V., Bailes, E., Bibollet-Ruche, F., Decker, J. M., Keele, B. F., Van Heuverswyn, F., Li, Y., Takehisa, J., Ngole, E. M., Shaw, G. M. et al. (2007) Adaptation of HIV-1 to its human host. Mol. Biol. Evol. 24, 1853–1860.

47 Van Heuverswyn, F., Li, Y., Neel, C., Bailes, E., Keele, B. F., Liu, W., Loul, S., Butel, C., Liegeois, F., Bienvenue, Y. et al. (2006) Human immunodeficiency viruses: SIV infection in wild gorillas. Nature 444, 164.

48 Wolfe, N. D., Heneine, W., Carr, J. K., Garcia, A. D., Shanmugam, V., Tamoufe, U., Torimiro, J. N., Prosser, A. T., Lebreton, M., Mpoudi-Ngole, E. et al. (2005) Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. Proc. Natl. Acad. Sci. USA 102, 7994–7999.

49 Wolfe, N. D., Switzer, W. M., Carr, J. K., Bhullar, V. B., Shanmugam, V., Tamoufe, U., Prosser, A. T., Torimiro, J. N., Wright, A., Mpoudi-Ngole, E. et al. (2004) Naturally acquired simian retrovirus infections in central African hunters. Lancet 363, 932–937.

50 Calattini, S., Chevalier, S. A., Duprez, R., Bassot, S., Froment, A., Mahieux, R. and Gessain, A. (2005) Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa. Retrovirology 2, 30.

51 Meertens, L. and Gessain, A. (2003) Divergent simian T-cell lymphotropic virus type 3 (STLV-3) in wild-caught Papio hamadryas papio from Senegal: widespread distribution of STLV-3 in Africa. J. Virol. 77, 782–789.

52 Nerrienet, E., Meertens, L., Kfutwah, A., Foupouapouognigni, Y., Ayouba, A. and Gessain, A. (2004) Simian T cell leukaemia virus type I subtype B in a wild-caught gorilla (Gorilla gorilla gorilla) and chimpanzee (Pan troglodytes vellerosus) from Cameroon. J. Gen. Virol. 85, 25–29.

53 Boneva, R. S., Switzer, W. M., Spira, T. J., Bhullar, V. B., Shanmugam, V., Cong, M. E., Lam, L., Heneine, W., Folks, T. M. and Chapman, L. E. (2007) Clinical and virological characterization of persistent human infection with simian foamy viruses. AIDS Res. Hum. Retroviruses 23, 1330–1337.

54 Switzer, W. M., Bhullar, V., Shanmugam, V., Cong, M. E., Parekh, B., Lerche, N. W., Yee, J. L., Ely, J. J., Boneva, R., Chapman, L. E. et al. (2004) Frequent simian foamy virus infection in persons occupationally exposed to nonhuman primates. J. Virol. 78, 2780–2789.

55 Jern, P., Sperber, G. O. and Blomberg, J. (2006) Divergent patterns of recent retroviral integrations in the human and chimpanzee genomes: probable transmissions between other primates and chimpanzees. J. Virol. 80, 1367–1375.

56 Medstrand, P. and Mager, D. L. (1998) Human-specific integrations of the HERV-K endogenous retrovirus family. J. Virol. 72, 9782–9787.

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

58 Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M. I., Kidd, K. K. and Lenz, J. (2001) Insertional polymorphisms of full-length endogenous retroviruses in humans. Curr. Biol. 11, 1531–1535.

59 Belshaw, R., Dawson, A. L., Woolven-Allen, J., Redding, J., Burt, A. and Tristem, M. (2005) Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J. Virol. 79, 12507–12514.

60 Dewannieux, M., Harper, F., Richaud, A., Letzelter, C., Ribet, D., Pierron, G. and Heidmann, T. (2006) Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. Genome Res. 16, 1548–1556.
61 Lee, Y. N. and Bieniasz, P. D. (2007) Reconstitution of an infectious human endogenous retrovirus. PLoS Pathog. 3, e10.
62 Jern, P., Lindeskog, M., Karlsson, D. and Blomberg, J. (2002) Full-length HERV-H elements with env SU open reading frames in the human genome. AIDS Res. Hum. Retroviruses 18, 671–676.
63 Jern, P., Sperber, G. O. and Blomberg, J. (2004) Definition and variation of human endogenous retrovirus H. Virology 327, 93–110.
64 Romano, C. M., de Melo, F. L., Corsini, M. A., Holmes, E. C. and Zanotto, P. M. (2007) Demographic histories of ERV-K in humans, chimpanzees and rhesus monkeys. PLoS ONE 2, e1026.
65 Mayer, J., Sauter, M., Racz, A., Scherer, D., Mueller-Lantzsch, N. and Meese, E. (1999) An almost-intact human endogenous retrovirus K on human chromosome 7. Nat. Genet. 21, 257–258.
66 Lavie, L., Medstrand, P., Schempp, W., Meese, E. and Mayer, J. (2004) Human endogenous retrovirus family HERV-K(HML-5): status, evolution, and reconstruction of an ancient betaretrovirus in the human genome. J. Virol. 78, 8788–8798.
67 Mayer, J., Ehlhardt, S., Seifert, M., Sauter, M., Muller-Lantzsch, N., Mehraein, Y., Zang, K. D. and Meese, E. (2004) Human endogenous retrovirus HERV-K(HML-2) proviruses with Rec protein coding capacity and transcriptional activity. Virology 322, 190–198.
68 Xiao, R., Kim, J., Choi, H., Park, K., Lee, H. and Park, C. (2008) Characterization of the Bovine Endogenous Retrovirus beta3 Genome. Mol. Cells 25, 142–147.
69 Bramble, D. M. and Lieberman, D. E. (2004) Endurance running and the evolution of Homo. Nature 432, 345–352.
70 Benit, L., Calteau, A. and Heidmann, T. (2003) Characterization of the low-copy HERV-Fc family: evidence for recent integrations in primates of elements with coding envelope genes. Virology 312, 159–168.
71 Tomonaga, K. and Coffin, J. M. (1998) Structure and distribution of endogenous nonecotropic murine leukemia viruses in wild mice. J. Virol. 72, 8289–8300.
72 Tomonaga, K. and Coffin, J. M. (1999) Structures of endogenous nonecotropic murine leukemia virus (MLV) long terminal repeats in wild mice: implication for evolution of MLVs. J. Virol. 73, 4327–4340.
73 Han, K., Konkel, M. K., Xing, J., Wang, H., Lee, J., Meyer, T. J., Huang, C. T., Sandifer, E., Hebert, K., Barnes, E. W. et al. (2007) Mobile DNA in Old World monkeys: a glimpse through the rhesus macaque genome. Science 316, 238–240.
74 Oliveira, N. M., Satija, H., Kouwenhoven, I. A. and Eiden, M. V. (2007) Changes in viral protein function that accompany retroviral endogenization. Proc. Natl. Acad. Sci. USA 104, 17506–17511.
75 Cohen, M. and Larsson, E. (1988) Human endogenous retroviruses. Bioessays 9, 191–196.
76 Mager, D. L. and Medstrand, P. (2003) Retroviral repeat sequences. In: Nature Encyclopedia of the Human Genome, Cooper, D. (ed.), Nature, London, UK.
77 Jurka, J. (2000) Repbase update: a database and an electronic journal of repetitive elements. Trends Genet. 16, 418–420.
78 Smit, A. F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr. Opin. Genet. Dev. 9, 657–663.
79 McCarthy, E. M. and McDonald, J. F. (2003) LTR_STRUC: a novel search and identification program for LTR retrotransposons. Bioinformatics 19, 362–367.
80 McClure, M. A., Richardson, H. S., Clinton, R. A., Hepp, C. M., Crowther, B. A. and Donaldson, E. F. (2005) Automated characterization of potentially active retroid agents in the human genome. Genomics 85, 512–523.
81 Sperber, G. O., Airola, T., Jern, P. and Blomberg, J. (2007) Automated recognition of retroviral sequences in genomic data – RetroTector. Nucleic Acids Res. 35, 4964–4976.
82 Hughes, J. F. and Coffin, J. M. (2004) Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. Proc. Natl. Acad. Sci. USA 101, 1668–1672.
83 Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. et al. (2001) Initial sequencing and analysis of the human genome. Nature 409, 860–921.
84 Jern, P., Sperber, G. O., Ahlsen, G. and Blomberg, J. (2005) Sequence variability, gene structure, and expression of full-length human endogenous retrovirus H. J. Virol. 79, 6325–6337.
85 Pringle, C. R. (1998) The universal system of virus taxonomy of the International Committee on Virus Taxonomy (ICTV), including new proposals ratified since publication of the Sixth ICTV Report in 1995. Arch. Virol. 143, 203–210.
86 de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U. and zur Hausen, H. (2004) Classification of papillomaviruses. Virology 324, 17–27.
87 Fauquet, C. M., Bisaro, D. M., Briddon, R. W., Brown, J. K., Harrison, B. D., Rybicki, E. P., Stenger, D. C. and Stanley, J. (2003) Revision of taxonomic criteria for species demarcation in the family Geminiviridae, and an updated list of begomovirus species. Arch. Virol. 148, 405–421.
88 Fauquet, C. M. and Fargette, D. (2005) International Committee on Taxonomy of Viruses and the 3,142 unassigned species. Virol. J. 2, 64.
89 Fauquet, C. M. and Mayo, M. A. (2001) The 7th ICTV report. Arch. Virol. 146, 189–194.
90 Fauquet, C. M. and Pringle, C. R. (2000) Abbreviations for bacterial and fungal virus species names. Arch. Virol. 145, 197–203.
91 Fauquet, C. M. and Stanley, J. (2005) Revising the way we conceive and name viruses below the species level: a review of geminivirus taxonomy calls for new standardized isolate descriptors. Arch. Virol. 150, 2151–2179.
92 van Regenmortel, M. H., Mayo, M. A., Fauquet, C. M. and Maniloff, J. (2000) Virus nomenclature: consensus versus chaos. Arch. Virol. 145, 2227–2232.
93 Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O. et al. (2007) A unified classification system for eukaryotic transposable elements. Nat. Rev. Genet. 8, 973–982.
94 Forsman, A., Yun, Z., Hu, L., Uzhameckis, D., Jern, P. and Blomberg, J. (2005) Development of broadly targeted human endogenous gammaretroviral pol-based real time PCRs Quantitation of RNA expression in human tissues. J. Virol. Methods 129, 16–30.
95 Hu, L., Hornung, D., Kurek, R., Ostman, H., Blomberg, J. and Bergqvist, A. (2006) Expression of human endogenous gammaretroviral sequences in endometriosis and ovarian cancer. AIDS Res. Hum. Retroviruses 22, 551–557.
96 Muradrasoli, S., Forsman, A., Hu, L., Blikstad, V. and Blomberg, J. (2006) Development of real-time PCRs for detection and quantitation of human MMTV-like (HML) sequences HML expression in human tissues. J. Virol. Methods 136, 83–92.
97 Lilly, F. (1967) Susceptibility to two strains of Friend leukemia virus in mice. Science 155, 461–462.
98 Stoye, J. P. (1998) Fv1, the mouse retrovirus resistance gene. Rev. Sci. Tech. 17, 269–277.
99 Arnaud, F., Murcia, P. R. and Palmarini, M. (2007) Mechanisms of late restriction induced by an endogenous retrovirus. J. Virol. 81, 11441–11451.
100 Murcia, P. R., Arnaud, F. and Palmarini, M. (2007) The transdominant endogenous retrovirus enJS56A1 associates with and blocks intracellular trafficking of Jaagsiekte sheep retrovirus Gag. J. Virol. 81, 1762–1772.

101 Gardner, M. B., Kozak, C. A. and O'Brien, S. J. (1991) The Lake Casitas wild mouse: evolving genetic resistance to retroviral disease. Trends Genet. 7, 22–27.
102 Crittenden, L. B., Smith, E. J., Weiss, R. A. and Sarma, P. S. (1974) Host gene control of endogenous avian leukosis virus production. Virology 57, 128–138.
103 Takaori-Kondo, A. (2006) APOBEC family proteins: novel antiviral innate immunity. Int. J. Hematol. 83, 213–216.
104 Cullen, B. R. (2006) Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. J. Virol. 80, 1067–1076.
105 Baumann, J. G. (2006) Intracellular restriction factors in mammalian cells – An ancient defense system finds a modern foe. Curr. HIV Res. 4, 141–168.
106 Ortiz, M., Bleiber, G., Martinez, R., Kaessmann, H. and Telenti, A. (2006) Patterns of evolution of host proteins involved in retroviral pathogenesis. Retrovirology 3, 11.
107 Perez, O. and Hope, T. J. (2006) Cellular restriction factors affecting the early stages of HIV replication. Curr. HIV/AIDS Rep. 3, 20–25.
108 Sokolskaja, E. and Luban, J. (2006) Cyclophilin, TRIM5, and innate immunity to HIV-1. Curr. Opin. Microbiol. 9, 404–408.
109 Towers, G. J. (2005) Control of viral infectivity by tripartite motif proteins. Hum. Gene Ther. 16, 1125–1132.
110 Campbell, E. M., Perez, O., Anderson, J. L. and Hope, T. J. (2008) Visualization of a proteasome-independent intermediate during restriction of HIV-1 by rhesus TRIM5alpha. J. Cell Biol. 180, 549–561.
111 Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R. and Hannon, G. J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089–1103.
112 Aravin, A. A., Hannon, G. J. and Brennecke, J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318, 761–764.
113 Schneider, P. M., Witzel-Schlomp, K., Rittner, C. and Zhang, L. (2001) The endogenous retroviral insertion in the human complement C4 gene modulates the expression of homologous genes by antisense inhibition. Immunogenetics 53, 1–9.
114 Schulte, A. M., Lai, S., Kurtz, A., Czubayko, F., Riegel, A. T. and Wellstein, A. (1996) Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. Proc. Natl. Acad. Sci. USA 93, 14759–14764.
115 Schulte, A. M., Malerczyk, C., Cabal-Manzano, R., Gajarsa, J. J., List, H. J., Riegel, A. T. and Wellstein, A. (2000) Influence of the human endogenous retrovirus-like element HERV-E.PTN on the expression of growth factor pleiotrophin: a critical role of a retroviral Sp1-binding site. Oncogene 19, 3988–3998.
116 Dunn, C. A. and Mager, D. L. (2005) Transcription of the human and rodent SPAM1/PH-20 genes initiates within an ancient endogenous retrovirus. BMC Genomics 6, 47.
117 Dunn, C. A., Medstrand, P. and Mager, D. L. (2003) An endogenous retroviral long terminal repeat is the dominant promoter for human beta1,3-galactosyltransferase 5 in the colon. Proc. Natl. Acad. Sci. USA 100, 12841–12846.
118 Dunn, C. A., van de Lagemaat, L. N., Baillie, G. J. and Mager, D. L. (2005) Endogenous retrovirus long terminal repeats as ready-to-use mobile promoters: the case of primate beta3-GAL-T5. Gene 364, 2–12.
119 Kowalski, P. E., Freeman, J. D. and Mager, D. L. (1999) Intergenic splicing between a HERV-H endogenous retrovirus and two adjacent human genes. Genomics 57, 371–379.
120 Kowalski, P. E., Freeman, J. D., Nelson, D. T. and Mager, D. L. (1997) Genomic structure and evolution of a novel gene (PLA2L) with duplicated phospholipase A2-like domains. Genomics 39, 38–46.
121 Kowalski, P. E. and Mager, D. L. (1998) A human endogenous retrovirus suppresses translation of an associated fusion transcript, PLA2L. J. Virol. 72, 6164–6168.
122 Landry, J. R., Medstrand, P. and Mager, D. L. (2001) Repetitive elements in the 5′ untranslated region of a human zinc-finger gene modulate transcription and translation efficiency. Genomics 76, 110–116.
123 Dunn, C. A., Romanish, M. T., Gutierrez, L. E., van de Lagemaat, L. N. and Mager, D. L. (2006) Transcription of two human genes from a bidirectional endogenous retrovirus promoter. Gene 366, 335–342.
124 Baust, C., Seifarth, W., Germaier, H., Hehlmann, R. and Leib, M.o.C. (2000) HERV-K-T47D-Related long terminal repeats mediate polyadenylation of cellular transcripts. Genomics 66, 98–103.
125 Renaudineau, Y., Hillion, S., Saraux, A., Mageed, R. A. and Youinou, P. (2005) An alternative exon 1 of the CD5 gene regulates CD5 expression in human B lymphocytes. Blood 106, 2781–2789.
126 Renaudineau, Y., Vallet, S., Le Dantec, C., Hillion, S., Saraux, A. and Youinou, P. (2005) Characterization of the human CD5 endogenous retrovirus-E in B lymphocytes. Genes Immun. 6, 663–671.
127 Hughes, J. F. and Coffin, J. M. (2001) Evidence for genomic rearrangements mediated by human endogenous retroviruses during primate evolution. Nat. Genet. 29, 487–489.
128 Hughes, J. F. and Coffin, J. M. (2005) Human endogenous retroviral elements as indicators of ectopic recombination events in the primate genome. Genetics 171, 1183–1194.
129 Belshaw, R., Katzourakis, A., Paces, J., Burt, A. and Tristem, M. (2005) High copy number in human endogenous retrovirus families is associated with copying mechanisms in addition to reinfection. Mol. Biol. Evol. 22, 814–817.
130 de Parseval, N. and Heidmann, T. (1998) Physiological knockout of the envelope gene of the single-copy ERV-3 human endogenous retrovirus in a fraction of the Caucasian population. J. Virol. 72, 3442–3445.
131 Cohen, J. (2007) Genomics. Venter's genome sheds new light on human variation. Science 317, 1311.
132 Levy, S., Sutton, G., Ng, P. C., Feuk, L., Halpern, A. L., Walenz, B. P., Axelrod, N., Huang, J., Kirkness, E. F., Denisov, G. et al. (2007) The diploid genome sequence of an individual human. PLoS Biol. 5, e254.
133 Coop, G., Wen, X., Ober, C., Pritchard, J. K. and Przeworski, M. (2008) High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. Science 319, 1395–1398.
134 Mayer, J., Stuhr, T., Reus, K., Maldener, E., Kitova, M., Asmus, F. and Meese, E. (2005) Haplotype analysis of the human endogenous retrovirus locus HERV-K(HML-2.HOM) and its evolutionary implications. J. Mol. Evol. 61, 706–715.
135 Katzourakis, A., Pereira, V. and Tristem, M. (2007) Effects of recombination rate on human endogenous retrovirus fixation and persistence. J. Virol. 81, 10712–10717.
136 Leupin, O., Attanasio, C., Marguerat, S., Tapernoux, M., Antonarakis, S. E. and Conrad, B. (2005) Transcriptional activation by bidirectional RNA polymerase II elongation over a silent promoter. EMBO Rep. 6, 956–960.
137 Mack, M., Bender, K. and Schneider, P. M. (2004) Detection of retroviral antisense transcripts and promoter activity of the HERV-K(C4) insertion in the MHC class III region. Immunogenetics 56, 321–332.
138 Rosok, O. and Sioud, M. (2005) Systematic search for natural antisense transcripts in eukaryotes (review). Int. J. Mol. Med. 15, 197–203.
139 Rosok, O. and Sioud, M. (2004) Systematic identification of sense-antisense transcripts in mammalian cells. Nat. Biotechnol. 22, 104–108.
140 Yelin, R., Dahary, D., Sorek, R., Levanon, E. Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R. et al. (2003) Widespread occurrence of antisense transcription in the human genome. Nat Biotechnol 21, 379–386.
141 Carmichael, G. G. (2003) Antisense starts making more sense. Nat Biotechnol 21, 371–372.

142 Pelisson, A., Sarot, E., Payen-Groschene, G. and Bucheton, A. (2007) A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the Drosophila ovary. J. Virol. 81, 1951–1960.
143 Buzdin, A. (2007) Human-specific endogenous retroviruses. ScientificWorldJournal 7, 1848–1868.
144 Werner, A. (2005) Natural antisense transcripts. RNA Biol. 2, 53–62.
145 Yu, W., Gius, D., Onyango, P., Muldoon-Jacobs, K., Karp, J., Feinberg, A. P. and Cui, H. (2008) Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature 451, 202–206.
146 Tsang, J., Zhu, J. and van Oudenaarden, A. (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. Mol. Cell 26, 753–767.
147 Mendes Soares, L. M. and Valcarcel, J. (2006) The expanding transcriptome: the genome as the 'Book of Sand'. EMBO J. 25, 923–931.
148 Wang, T., Zeng, J., Lowe, C. B., Sellers, R. G., Salama, S. R., Yang, M., Burgess, S. M., Brachmann, R. K. and Haussler, D. (2007) Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. Proc. Natl. Acad. Sci. USA 104, 18613–18618.
149 Desset, S., Buchon, N., Meignin, C., Coiffet, M. and Vaury, C. (2008) In Drosophila melanogaster the COM locus directs the somatic silencing of two retrotransposons through both Piwi-dependent and -independent pathways. PLoS ONE 3, e1526.
150 Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E. et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447, 799–816.
151 Villesen, P., Aagaard, L., Wiuf, C. and Pedersen, F. S. (2004) Identification of endogenous retroviral reading frames in the human genome. Retrovirology 1, 32.
152 de Parseval, N., Diop, G., Blaise, S., Helle, F., Vasilescu, A., Matsuda, F. and Heidmann, T. (2005) Comprehensive search for intra- and inter-specific sequence polymorphisms among coding envelope genes of retroviral origin found in the human genome: genes and pseudogenes. BMC Genomics 6, 117.
153 de Parseval, N., Lazar, V., Casella, J. F., Benit, L. and Heidmann, T. (2003) Survey of human genes of retroviral origin: identification and transcriptome of the genes with coding capacity for complete envelope proteins. J. Virol. 77, 10414–10422.
154 Britten, R. J. (2004) Coding sequences of functioning human genes derived entirely from mobile element sequences. Proc. Natl. Acad. Sci. USA 101, 16825–16830.
155 Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C. et al. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220, 868–871.
156 Gallo, R. C. (2005) History of the discoveries of the first human retroviruses: HTLV-1 and HTLV-2. Oncogene 24, 5926–5930.
157 Abramczuk, J. W., Pezen, D. S., Leonard, J. M., Monell-Torrens, E., Belcher, J. H., Martin, M. A. and Notkins, A. L. (1992) Transgenic mice carrying intact HIV provirus: biological effects and organization of a transgene. J. Acquir. Immune Defic. Syndr. 5, 196–203.
158 Katzourakis, A., Tristem, M., Pybus, O. G. and Gifford, R. J. (2007) Discovery and analysis of the first endogenous lentivirus. Proc. Natl. Acad. Sci. USA 104, 6261–6265.
159 Goedert, J. J., Sauter, M. E., Jacobson, L. P., Vessella, R. L., Hilgartner, M. W., Leitman, S. F., Fraser, M. C. and Mueller-Lantzsch, N. G. (1999) High prevalence of antibodies against HERV-K10 in patients with testicular cancer but not with AIDS. Cancer Epidemiol. Biomarkers Prev. 8, 293–296.
160 Gotzinger, N., Sauter, M., Roemer, K. and Mueller-Lantzsch, N. (1996) Regulation of human endogenous retrovirus-K Gag expression in teratocarcinoma cell lines and human tumours. J. Gen. Virol. 77, 2983–2990.
161 Herbst, H., Sauter, M., Kuhler-Obbarius, C., Loning, T. and Mueller-Lantzsch, N. (1998) Human endogenous retrovirus (HERV)-K transcripts in germ cell and trophoblastic tumours. APMIS 106, 216–220.
162 Herbst, H., Sauter, M. and Mueller-Lantzsch, N. (1996) Expression of human endogenous retrovirus K elements in germ cell and trophoblastic tumors. Am. J. Pathol. 149, 1727–1735.
163 Sauter, M., Roemer, K., Best, B., Afting, M., Schommer, S., Seitz, G., Hartmann, M. and Mueller-Lantzsch, N. (1996) Specificity of antibodies directed against Env protein of human endogenous retroviruses in patients with germ cell tumors. Cancer Res. 56, 4362–4365.
164 Sauter, M., Schommer, S., Kremmer, E., Remberger, K., Dolken, G., Lemm, I., Buck, M., Best, B., Neumann-Haefelin, D. and Mueller-Lantzsch, N. (1995) Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas. J. Virol. 69, 414–421.
165 Galli, U. M., Sauter, M., Lecher, B., Maurer, S., Herbst, H., Roemer, K. and Mueller-Lantzsch, N. (2005) Human endogenous retrovirus rec interferes with germ cell development in mice and may cause carcinoma in situ, the predecessor lesion of germ cell tumors. Oncogene 24, 3223–3228.
166 Boese, A., Sauter, M., Galli, U., Best, B., Herbst, H., Mayer, J., Kremmer, E., Roemer, K. and Mueller-Lantzsch, N. (2000) Human endogenous retrovirus protein cORF supports cell transformation and associates with the promyelocytic leukemia zinc finger protein. Oncogene 19, 4328–4336.
167 Blomberg, J., Uzhameckis, D. and Jern, P. (2003) Evolutionary aspects of human endogenous retroviral sequences (HERVs) and disease. In: Retroviruses and Primate Genome Evolution, Sverdlov, E. (ed.), Eurekah Bioscience, Austin, TX.
168 Perron, H., Garson, J. A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., Mallet, F., Tuke, P. W., Voisset, C., Blond, J. L. et al. (1997) Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc. Natl. Acad. Sci. USA 94, 7583–7588.
169 Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F., Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E., Mallet, F., Mandrand, B. et al. (1999) Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles. Virology 260, 1–9.
170 Haahr, S., Sommerlund, M., Christensen, T., Jensen, A. W., Hansen, H. J. and Moller-Larsen, A. (1994) A putative new retrovirus associated with multiple sclerosis and the possible involvement of Epstein-Barr virus in this disease. Ann. N. Y. Acad. Sci. 724, 148–156.
171 Christensen, T., Tonjes, R. R., zur Megede, J., Boller, K. and Moller-Larsen, A. (1999) Reverse transcriptase activity and particle production in B lymphoblastoid cell lines established from lymphocytes of patients with multiple sclerosis. AIDS Res. Hum. Retroviruses 15, 285–291.
172 Christensen, T., Dissing Sorensen, P., Riemann, H., Hansen, H. J., Munch, M., Haahr, S. and Moller-Larsen, A. (2000) Molecular characterization of HERV-H variants associated with multiple sclerosis. Acta Neurol. Scand. 101, 229–238.
173 Christensen, T., Dissing Sorensen, P., Riemann, H., Hansen, H. J. and Moller-Larsen, A. (1998) Expression of sequence variants of endogenous retrovirus RGH in particle form in multiple sclerosis. Lancet 352, 1033.
174 La Mantia, G., Maglione, D., Pengue, G., Di Cristofano, A., Simeone, A., Lanfrancone, L. and Lania, L. (1991) Identification and characterization of novel human endogenous retroviral sequences prefentially expressed in undifferentiated embryonal carcinoma cells. Nucleic Acids Res. 19, 1513–1520.

175 Antony, J. M., Zhu, Y., Izad, M., Warren, K. G., Vodjgani, M., Mallet, F. and Power, C. (2007) Comparative expression of human endogenous retrovirus-W genes in multiple sclerosis. AIDS Res. Hum. Retroviruses 23, 1251–1256.
176 Mattson, M. P. and Taub, D. D. (2004) Ancient viral protein enrages astrocytes in multiple sclerosis. Nat. Neurosci. 7, 1021–1023.
177 Hirschl, S., Schanab, O., Seppele, H., Waltenberger, A., Humer, J., Wolff, K., Pehamberger, H. and Muster, T. (2007) Sequence variability of retroviral particles derived from human melanoma cells melanoma-associated retrovirus. Virus Res. 123, 211–215.
178 Buscher, K., Hahn, S., Hofmann, M., Trefzer, U., Ozel, M., Sterry, W., Lower, J., Lower, R., Kurth, R. and Denner, J. (2006) Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines. Melanoma Res. 16, 223–234.
179 Schiavetti, F., Thonnard, J., Colau, D., Boon, T. and Coulie, P. G. (2002) A human endogenous retroviral sequence encoding an antigen recognized on melanoma by cytolytic T lymphocytes. Cancer Res. 62, 5510–5516.
180 Humer, J., Waltenberger, A., Grassauer, A., Kurz, M., Valencak, J., Rapberger, R., Hahn, S., Lower, R., Wolff, K., Bergmann, M. et al. (2006) Identification of a melanoma marker derived from melanoma-associated endogenous retroviruses. Cancer Res. 66, 1658–1663.
181 Takahashi, Y., Harashima, N., Kajigaya, S., Yokoyama, H., Cherkasova, E., McCoy, J. P., Hanada, K., Mena, O., Kurlander, R., Abdul, T. et al. (2008) Regression of human kidney cancer following allogeneic stem cell transplantation is associated with recognition of an HERV-E antigen by T cells. J. Clin. Invest. 118, 1099–1109.
182 Wang, X. S., Zhang, Z., Wang, H. C., Cai, J. L., Xu, Q. W., Li, M. Q., Chen, Y. C., Qian, X. P., Lu, T. J., Yu, L. Z. et al. (2006) Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma. Clin. Cancer Res. 12, 4851–4858.
183 Foerster, J., Nolte, I., Junge, J., Bruinenberg, M., Schweiger, S., Spaar, K., van der Steege, G., Ehlert, C., Mulder, M., Kalscheuer, V. et al. (2005) Haplotype sharing analysis identifies a retroviral dUTPase as candidate susceptibility gene for psoriasis. J. Invest. Dermatol. 124, 99–102.
184 Mayer, J. and Meese, E. U. (2003) Presence of dUTPase in the various human endogenous retrovirus K (HERV-K) families. J. Mol. Evol. 57, 642–649.
185 Sousa, M. M., Krokan, H. E. and Slupphaug, G. (2007) DNA-uracil and human pathology. Mol. Aspects Med. 28, 276–306.