

Murine endogenous retroviruses

C. Stocking^{a,*} and C. A. Kozak^{b,*}

^a Heinrich-Pette-Institute, Martinistrasse 52, 20251 Hamburg (Germany),
e-mail: stocking@hpi.uni-hamburg.de

^b Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases,
National Institutes of Health, 4 Center Drive MSC 0460 Bethesda, Maryland 20892-0460 (USA),
e-mail: ckozak@niaid.nih.gov

Online First 27 September 2008

Abstract. Up to 10% of the mouse genome is comprised of endogenous retrovirus (ERV) sequences, and most represent the remains of ancient germ line infections. Our knowledge of the three distinct classes of ERVs is inversely correlated with their copy number, and their characterization has benefited from the availability of divergent wild mouse species and subspecies, and from ongoing analysis of the *Mus* genome sequence. In contrast to human ERVs, which are nearly all extinct, active mouse ERVs can still be

found in all three ERV classes. The distribution and diversity of ERVs has been shaped by host-virus interactions over the course of evolution, but ERVs have also been pivotal in shaping the mouse genome by altering host genes through insertional mutagenesis, by adding novel regulatory and coding sequences, and by their co-option by host cells as retroviral resistance genes. We review mechanisms by which an adaptive coexistence has evolved. (Part of a Multi-author Review)

Keywords. Virus restriction factors, retroviral receptors, Env-receptor interactions, murine ERV classification.

Introduction

Close to 40% of the mouse genome is made up of fossils of transposable elements, of which approximately 10% represent endogenous retrovirus (ERV) sequences [1]. These ERV sequences have resulted from both ancient and modern infections of exogenous retroviruses, which have successfully colonized the germ line of their host [2]. Once a retrovirus becomes endogenous, the provirus survives as part of the host genome rather than as an autonomous infectious agent and thus is subjected to selection pressures acting on the host genome. Although extinction is probably the fate for most ERV lineages, it is clear that certain endogenous viruses are still capable of expression and replication even after millions of years within the host genome. Mouse genomes, in particular, have many active ERVs, which contrasts strikingly with the human genome, where all

ERVs are nearly extinct, with the possible exception of HERV-K [3]. Thus, there must be other, poorly defined determinants that influence whether ERVs remain active or become extinct.

The aim of this review is to summarize our current knowledge of the various ERV lineages within the mouse genome and to provide examples that illustrate how elements have evolved together with their host, not only to confer benefits to the host, but also to permit continued coexistence, by modulating viral load and pathogenicity. Such studies in the mouse are particularly valuable because of the key role played by the inbred laboratory mouse strains in studies on mammalian genetics, development, pathology and retrovirology, and also because of the availability of wild mouse species that cover 12 million years of evolutionary history [4]. Emphasis is placed on changes in Env-receptor interactions, which are well studied for the infectious mouse retroviruses, and which have been exploited by host and virus to evolve and evade entry-related restrictions. These receptor interactions may also contribute to horizontal trans-

* Corresponding authors.

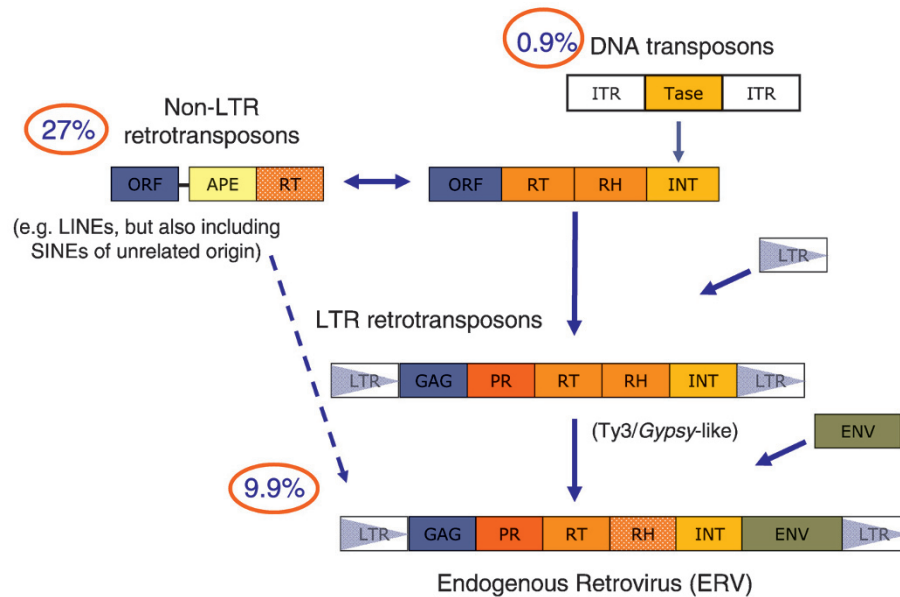


Figure 1. Fossils of transposable elements make up a large proportion of the mouse genome. The percentage of the mouse genome sequences that are derived from one of two types of transposable elements (DNA transposons and retrotransposons) is shown. The retrotransposons are further divided into non-LTR and LTR retrotransposons. LTR retrotransposons in the mouse belong to the ERV superfamily, which is made up of three families. Arrows and fill colors denote potential evolutionary relationships and/or recombination events. Abbreviations used in the figure are defined in the text, with the exception of APE (apurinic endonuclease), found in LINE elements. The figure is adapted from [10] and [5].

mission to other species, where ERVs can begin a new cycle of unchecked copy number increase until new mechanisms evolve to limit their spread.

Origin and classification of ERVs

The human and mouse sequencing projects provided definitive evidence that a substantial proportion of both genomes was comprised of transposable elements (46 and 37.5%, respectively). Using the unified classification system proposed by Wicker et al. [5], the transposable elements can be divided into two main classes: retrotransposons and DNA transposons, the former constituting 95% of the transposable element sequences present in the mouse genome (Fig. 1). Whereas DNA transposons amplify without an RNA intermediate, retrotransposons rely on an RNA transcript that is 'retrotranscribed' by a reverse transcriptase (RT) before integration into the genome. Retrotransposons can be further divided into five orders, three of which are found in mammals: LINES (long-interspersed nucleotide elements), SINES (short-interspersed nucleotide elements), and LTRs (long-terminal direct repeats). LINES are autonomous and encode at least an RT and a nuclease in their *pol* ORF [open reading frame] for transposition. Copies of LINE (superfamily L1) form the single largest fraction of interspersed repeat sequence in both human and mouse, with about 4800 full-length copies in mouse, of which 3000 are predicted to be active [6]. The SINE order is also classified within the class I retrotransposons, but is distinct in origin. SINES originate from accidental retrotransposition of various

polymerase III transcripts and rely on LINES for trans-acting transposition functions such as RT [7]. Whereas only a single SINE family (Alu) is active in the human lineage, the mouse lineage has been exposed to four distinct SINES (B1, B2, ID, B4), originally derived from tRNA and 7SL genes [7]. Together they occupy about 27.4% of the mouse genome [1].

The third order is the LTR retrotransposons. LTR retrotransposons are the predominant order of retrotransposons in plants and are generally less abundant in animals; nevertheless, close to 10% of the mouse and human genomes are derived from this order of transposable elements. LTR retrotransposons have a proposed chimeric origin, arising from fusion(s) between a DNA transposon and a non-LTR retrotransposon [8] (Fig. 1); the DNA transposon providing integrase (transposase, Tase) and the requirement for a short inverted terminal repeat at the ends of the element, and the non-LTR retrotransposon contributing the RT and RH (ribonuclease H) enzymatic functions, but also a *gag*-like ORF domain that provided the basis of the capsid (CA) and nuclear capsid (NC) proteins. Although several other components would be required to complete the formation of a fully functional LTR retrotransposon (e.g. including the evolution of LTRs as the means to overcome the problem of replicating the ends of any DNA molecule), the only additional protein domain required is a aspartic proteinase (PR) domain, which may have originated from the host's pepsin gene family [9].

In mammals, all LTR retrotransposons are derivatives of the vertebrate-specific ERV superfamily. Retro-

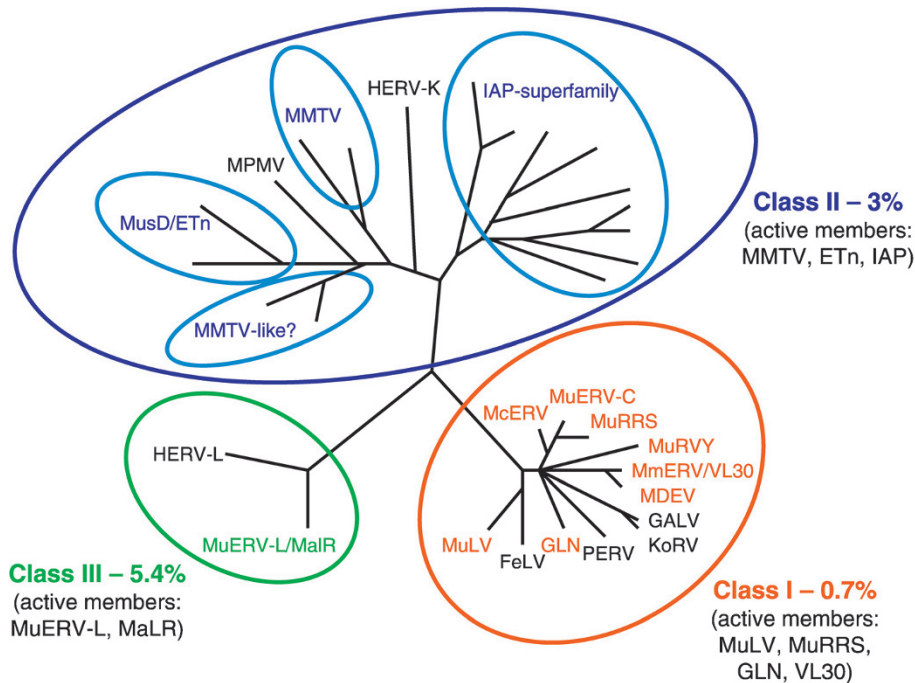


Figure 2. Phylogenetic analysis of ERV RT domains [19] demonstrates the three classes of mouse retrotransposons. RT sequences of ERVs from host species other than mouse are included for comparison and are in black letters. The close relationship of Class I ERVs to XRVs and ERVs from other species is clearly shown. Four distinct clades or superfamilies are defined for the Class II ERVs, one of which (MMTV-like) is poorly characterized. Non-autonomous elements, such as the abundant VL30 s (Class I), ETNs (Class II), and MaLRs (Class III) are listed with their presumed parental ERVs, as they do not contain RT domains. The phylogenetic tree was obtained from <http://genomebiology.com/2004/5/3/R14> and modified to include our own analysis of Class I ERVs.

viruses are generally thought to have evolved from *Gypsy*-like LTR retrotransposons, which adopted a viral lifestyle through acquisition of an envelope protein (Env) (Fig. 1). Most ERVs show clear homology to one another and to modern exogenous retroviruses (XRV) (albeit to a lesser extent), especially across the RT gene, which is relatively refractory to nonsynonymous substitution [11, 12]. In addition, shared characteristics such as translational strategy, number of zinc finger proteins in the NC of gag, the presence and location of dUTPase (preventing incorporation of uracil), presence of a GPY/F motif in the C-terminal end of IN, and accessory genes can be used to classify ERVs [11]. There has been a growing tendency to group ERVs into classes according to their similarity to XRVs, which have been classified into seven genera (alpha-, beta-, gamma-, delta, and epsilonretrovirus, lentivirus, and spumavirus), the latter belonging to a distinct subfamily [13]. Using this system of classification, ERVs clustering with gamma- and epsilonretrovirus are termed Class I, those that cluster with lentivirus, alpha-, beta-, and deltaretroviruses are termed Class II, and those that cluster with spumaviruses are termed Class III [11, 14–16]. Notably, intermediates between these different families have been identified, indicating an evolutionary continuum.

Distribution and classes of ERVs in the mouse genome

Unveiling of the mouse genome sequence in 2002 allowed the first comprehensive effort to catalogue the diversity of ERVs in the mouse genome [1]. Subsequently, several data-mining programs have been used to both identify novel ERV families, as well as validate earlier genetic analysis [1, 17–20]. As in the human genome, the three different classes of ERVs can be readily distinguished, and together make up close to 10% of their host's genome (Fig. 2). However, a markedly dissimilar evolutionary history in human and mouse has been noted, both in the distribution and number of ERV families within the different classes, but also in the fact that ERVs are nearly extinct in human, whereas in mouse there are many active members [1, 21]. Although this reduced activity of ERVs in humans reflects in part an unexplained drop in the overall rate of transposition in the human but not mouse genome over the past 40 million years [1, 22], many other factors are clearly involved. Before addressing some of these mechanisms by which ERV activity is maintained or extinguished, an overview of the three major ERV classes found in the mouse genome is warranted.

Class III ERVs, which show closest (although distant) homology to the spuma-like genus of retrovirus, make up 5.4% of the mouse genome [1] (Fig. 2). These are probably the most ancient ERVs, accounting for 80% of recognized LTR element copies predating the

human-mouse speciation [1, 14]. Interestingly, the human genome contains about the same percentage of class III ERVs as the mouse genome, although in contrast to mouse, no active forms have been detected. This discrepancy most likely reflects a higher nucleotide substitution rate in the mouse, making it more difficult to recognize ancient repeat sequences, and thus giving an underestimate of their actual density [1].

Two types of transposon elements constitute the Class III ERVs: murine ERV-L elements and the non-autonomous MaLRs (mammalian apparent LTR retrotransposons), the latter being the most common retroviral elements in the mouse genome, making up 4.8% of the genome. Based on the 51% similarity between their LTRs, these two types are thought to have a recent common ancestor [19]. The ERV-L family was first identified in humans, but was subsequently found to be in all placental mammals, thus was present in the mammalian common ancestor more than 70 million years ago (Mya) [14]. Accordingly, MuERV-L sequences are found in all characterized species within the *Mus* subgenus, as well as in mice of the other three subgenera *Nannomys*, *Pyromys*, and *Coelomys* (Fig. 3). Despite its age, this family has maintained some of its elements in an active state in the mouse, as demonstrated by recent amplifications in this species [14, 23]. It has recently also been shown that MuERV-L sequences are responsible for epsilon virus-like particles observed in the early mouse embryo [24], consistent with several reports showing high levels of expression during early embryonic development [25, 26]. Interestingly, the ERV-L family members have *gag* and *pol* genes but no detectable *env*, and an overall length of 6400 bp. Like spumaretroviruses, MuERV-L lacks the Cys-His region in the NC, important for binding to nucleic acids, but does contain the highly conserved major homology region (MHR) in CA, which is lacking in the related XRV family. Notably, both HERV-L and MuERV-L encode dUTPase, in contrast to spumaretrovirus but similar to betaretrovirus and non-primate lentivirus. However, the location of this gene within the viral genome is distinct, suggesting that its acquisition was an independent event [11, 27]. There is little evidence of MuERV-L horizontal transfer between species, an observation that is consistent with its lack of *env* [14]. The non-autonomous MaLRs are all internally deleted, containing only non-coding repetitive DNA [38]. Nevertheless they have typical LTRs, a primer binding site and a polypurine tract. In the mouse genome there are an estimated 380 000 copies of MaLR elements (including solitary LTRs) [1], which belong to one of two types: MT (mouse transposon) and ORR₁ (origin-region repeat) MaLRs [38]. They are closely related to

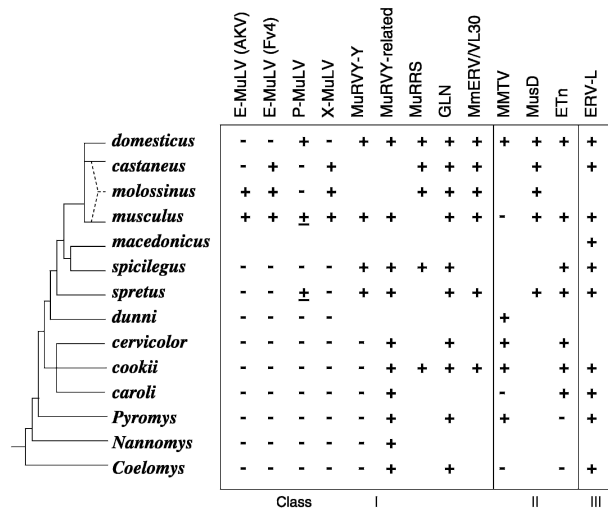


Figure 3. Schematic representation of the evolution of *Mus* indicating the species distribution of different ERV families. This evolutionary tree is based largely on the synthetic tree of Boursot and Guenet [4, 28]. All species from *domesticus* through *caroli* are subgenus *Mus*. The most recent node of the tree represents the house mouse *Mus musculus* complex; *M. molossinus* is a natural hybrid of *castaneus* and *musculus*. Species tested for specific ERV sequences by Southern-blot analysis are indicated; cases with a very few detectable copies are underlined [14, 29–37].

the THE-1 (or MstII) family in the human genome. The MT lineage is the most prevalent type of ERV in the mouse genome and has a mean length of 1980 bp. In contrast, members of the ORR₁ lineage have a mean length of approximately 2460 bp and are about 10-fold less frequent in the genome. Both member types are still active in the mouse, but died out some 50 Mya in human [38].

Class II ERV elements make up 3.14% of the mouse genome, a 10-fold higher proportion in mouse than in humans [1]. There is no evidence for a Class II ERV that predates the human-mouse speciation. The Class II ERVs include elements that resemble the betaretrovirus XRV genus, typified by the mouse mammary tumor virus (MMTV), which is morphologically classified as a type-B virus, and the Mason-Pfizer Monkey Virus (MPMV), a type-D virus (Fig. 2). Both types are distinguished by their mode of assembly, which initiates in the cytoplasm before transport to the plasma membrane [39]. In addition, betaretroviruses share a translational strategy involving two –1 frameshifts to translate *pro-pol*-encoded proteins, the presence of a UTPase coding region upstream of PR, and two zinc-finger motifs in NC [11]. More than 30 endogenous MMTVs are scattered over the laboratory mouse genome, although only two or three full-length, active alleles have been identified [19, 40, 41]. Notably, the human MMTV-like (HML)-2 HERV-K subfamily may be the only ERV family with active elements in the human genome [42, 43].

Related to the MMTV, but more closely to MPMV, is a much larger group of elements known as MusD, with their deleted variants denoted ETn (early transposons) [35, 44]. MusD contains *gag-pro-pol* genes, but lacks an *env* gene. Furthermore, its *gag*-encoded matrix (MA) protein lacks a functional sequence for myristylation and plasma membrane targeting, thereby restricting virions to intracellular compartments, perhaps facilitating its capacity as an intracellular mutagen [45]. Interestingly, due to their high expression levels, ETn genomes are the actual primary mutagen in this family, their retrotransposition being mediated by MusD-encoded viral proteins [46]. ETns, similar to MaLRs, are flanked by LTRs but contain mainly non-retroviral, non-coding sequences of unknown origin. MusD/ETns are widely distributed among the different species of the subgenus *Mus* but are not found in *Rattus*, or in the *Pyromys* or *Coelomys* subgenera of *Mus*, which diverged between 5 and 10 Mya [36] (Fig. 3).

The final well-characterized clade within the Class II ERVs is related to IAP [intracisternal A-type particles], which are present in mice at approximately 1000 copies per cell [47]. Phylogenetic analysis of related RT sequences in the mouse genome indicate that this clade of elements can be grouped into seven to eight families (defined by sharing < 90% identity at the amino acid level within RT) [19, 47]. They assemble on the endoplasmic reticulum and bud into the cisternae but are not released from the cell. Although originally thought to lack an *env* gene, two subsets of IAP-related proviruses that encode Env proteins have been identified, with a single provirus being infectious [48, 49].

Class I ERVs comprise the smallest class, making up 0.68% of the mouse genome, but its members are the best-studied ERVs. They are closely related to the gammaretrovirus genus of XRVs, typified by the type-C murine leukemia viruses (MuLV). In contrast to the type-B and -D retroviruses, type-C viruses assemble their genomes at the plasma membrane, with the concurrent formation of the immature viral core and the acquisition of the envelope. Class I ERVs share a similar translation strategy (read-through of the termination codon at the end of *gag*), the presence of one zinc-finger in NC, and a conserved GPY/F motif in the C-terminus of IN. Although these sequences are fourfold more common in the human than the mouse genome, no active Class I ERVs are known in humans. In contrast, the mouse genome contains many active members of this class, which have been the subject of intense research over the last four decades.

Early work of Todaro et al. based on DNA hybridization and antigen screening defined two major

subclasses of type-C retroviruses [50, 51]. Strikingly, these two major subclasses have stood the test of time and technology. The subclass C-I of Todaro includes XRVs and ERVs from several primate sources, including gibbon ape and woolley monkey, as well as isolates from wild Asian mice, such as *Mus cervicolor* and *Mus caroli*. In contrast, the subclass C-II is composed of MuLV ERV and XRV isolates from different inbred laboratory strains and species of *Mus musculus* and includes the related ERV family from cats (FeLV). Sporadic reports over the last two decades have described related but distinct ERV isolates from various mouse strains that belong to one of these two subclasses, but only with the availability of the mouse genome sequence have there been more rigorous attempts to further classify these Class I ERVs. Our attempt is shown in Table 1, and is pictorially demonstrated by a phylogenetic analysis based on RT analysis of different ERV and XRV isolates, which also shows the two major clades (subclasses) of Class I ERVs, first recognized by Todaro (Fig. 2).

The best-characterized ERVs are the members that belong to the subclass C-II MuLV family. This is a homogeneous family that has recently (< 1.5 Mya) entered the *Mus* genome, as indicated by the insertional polymorphism among inbred mouse strains and *Mus* subspecies and by the presence of infectious members [34, 60, 61] (Fig. 3). MuLVs are generally divided into distinct groups on the basis of the host ranges specified by their *env* gene (Table 1). The xenotropic (X)-MuLVs are present in about 20 copies per mouse genome, whereas the polytropic (P)-MuLVs, which can be subdivided into two closely related subgroups, are present in about 40 copies per genome [62]. Ecotropic (E)-MuLVs are found in only a few copies in diverse inbred strains [63]. Analysis of wild mouse species indicates that ERVs of all three subgroups entered the *Mus* germline recently (Fig. 3) [34]. Of the three groups of ERVs, only X-MuLV and E-MuLV have infectious members, although infectious P-MuLVs are generated following recombination with E-MuLVs [64, 65]. Some wild mouse populations carry infectious amphotropic MuLVs (A-MuLVs) or a novel E-MuLV (HoMuLV) in the absence of germline ERVs [66, 67].

Somewhat surprisingly, most of the Class I ERVs in the mouse genome actually belong to the lesser-characterized subclass C-I family. At least six distinct families can be recognized (as defined by < 90% homology in RT). The first family recognized in this group were the MuRRSs (murine retroviral-related sequences) [37]. This is a family of approximately 30–50 copies with an approximate size of 5.5 kb, with recognizable but highly mutated *gag*, *pol*, and *env*

Table 1. Seven families of Class I ERVs and related XRVs with their *env* variants.

Families	Prototype viruses in different <i>Mus</i> species ¹	Host range ²	Receptor	Related XRVs or ERVs in other species
Subclass C-I				
γ -A	MuRRS	no <i>env</i>	N.A.	GALV, KoRV, PERV
γ -B	MURV-Y	not active	unknown	
γ -C	MmERV MDEV (<i>Mus dunni</i>)	n.t. multitropic	unknown unknown [52]	
γ -D	McERV(<i>Mus caroli</i>)	monotropic	PLLp [53]	
γ -E	MuERV-C	not active	unknown	
γ -F	GLN-2	ecotropic	unknown [18]	
Subclass C-II				
γ -G	P-MuLV X-MuLV AKR-MuLV CasBrE ⁴ Ho-MuLV (<i>Mus spicilegus</i>) ⁵ 4070A-MuLV ⁵ 10A1-MuLV ⁵ HMEV (<i>Mus spicilegus</i>) M813 (<i>Mus cervicolor</i>)	polytropic ³ xenotropic ecotropic amphotropic 10A1 ecotropic	XPR1[54] XPR1[54] mCAT-1[55] PiT2 [56] PiT1, PiT2 [57] mSMIT [58, 59]	FeLV, BaEV, XMRV

¹ Prototype ERVs listed have been characterized in inbred mouse strains, unless noted.

² Host range refers to the historical nomenclature used to define 'host' susceptibility to virus infections, generally tested using cell lines of various origins. Ecotropic refers to the ability of the virus to infect cells of murine origin only; but not all ecotropic viruses use the same receptor. Xenotropic refers to infectivity of cells of non-murine origin only. Poly-, amphi-, 10A1-, and multitropic all refer to the ability of the virus to infect cells from all (or most) species tested, but through different cellular receptors. Monotropic refers to the fact that the virus can infect cells of many species, but only distinct cell types, which reflects the limited expression pattern of the receptor.

³ Active forms have been found only after recombination events with XRV.

⁴ Isolated from Lake Casitas (CA) wild mice of unclear origin. Similar ERVs have been identified in *M. musculus* and *M. castaneus*.

⁵ These prototypes have only been found as XRVs.

Abbreviations used: N.A., not applicable; n.t., not tested

genes. Many of these copies share identical deletions within the *pol* gene, suggesting that they were mobilized by a nondefective helper virus. In addition to these mutated proviral genomes, there are at least 500 copies of solo LTRs, probably arising through homologous recombination between the two LTRs. Two additional families (MuRV-Y and MuERV-C) are closely related, but are predicted to have been seeded by a separate 'founder' retrovirus [68]. Similar to MuRRS, MuERV-C family members are highly mutated but with detectable *gag*, *pol*, and *env* genes, the latter being often extensively deleted. At least 20 copies can be detected in the mouse genome, 10 of which are clustered on the X chromosome. This clustering of integrations suggests that these have been amplified by a non-retroviral mechanism. A similar mechanism is probably responsible for the close to 500 copies of MuRV-Y (murine repeated virus on the Y chromosome), which are almost exclusively found on the Y chromosome, as part of a 25-kb amplicon [69, 70]. In contrast to the MuRRS and MuERV-C members characterized so far, MuRV-Y members have incurred fewer mutations/deletions, but no active forms are known [71]. MuRRS and MuRV-Y are thought to have entered the mouse

genome within the past 9 Mya, whereas the multiple Y chromosome MuRV-Y copies are found only in the most recently derived European *Mus* species, reflecting a more recent amplification [37, 70] (Fig. 3).

In contrast to the families described above, three additional families have been recently recognized for which active ERVs are known. One is the GLN family, named for its glutamine tRNA primer-binding site, in contrast to proline tRNA found in many other Class I ERVs [72]. Of the over 80 copies found in the mouse genome, one copy was recently found both to be intact and capable of releasing virus particles [18]. In view of the fact that related sequences have been found in both *Mus* subspecies and other rodent (but not nonrodent) genomes, it must have entered a common ancestor prior to the *Mus/Rattus* split (circa 16–23 Mya) [18, 72] (Fig. 3).

The first member of the next family was first identified in the *Mus dunni* (also termed *M. terricolor*) genome as an active ERV (MDEV) [73], but it was not until the sequence of the mouse genome was 'mined' for novel ERVs that its counterpart within the C57BL/6 laboratory mouse genome was identified (MmERV) [20]. Among the more than 50+ full-length copies in the mouse genome, at least one provirus has intact

reading frames for all three genes, although its activity has not been assessed. Interestingly, both MmERV and MDEV have LTRs that are similar to VL30 elements, which are nonautonomous retrotransposons first identified over 30 years ago [74, 75]. It is thus probable that VL30s are derived from internally deleted MmERV elements, which utilize various Class I ERVs for transposition. Similar to MmERV and MDEV, VL30 elements generally (but not exclusively) depend on glycine tRNA for priming the reverse transcription. There are approximately 100–200 copies of VL30 elements in the mouse genome. In contrast to other ERV sequences analyzed, where there is generally a 20-fold higher frequency of solitary LTRs than proviral genomes, for VL30s there are at least 20-fold more full-sized elements than solo LTRs [76]. This probably reflects the high transposition activity of these elements (coupled with increased mutagenesis capacity), which is facilitated by their high expression levels and permissive packaging signals. VL30 (and presumably MmERV) sequences are present in all inbred mice and various *Mus* species, although the number of copies varies (Fig. 3). Finally, the last characterized family is actually the first C-I family member to be identified in mice [51]. Recent molecular cloning and analysis of this isolate (dubbed McERV) from *Mus caroli* fibroblasts has shown that it is closely related but distinct from the MmERV/MDEV family ([77], unpublished observation). Hybridization and sequence analysis has shown that although it is found at a high copy number in *Mus caroli* cells (ca. 50), highly related proviruses are also found in the inbred C57BL/6 laboratory mouse genome (3 intact but mutated copies). Notably, proviruses with at least two different *env* genes have been characterized ([77], unpublished observation).

The major lesson learned from mouse ERVs:
‘It takes all the running you can do, to keep in the same place.’ – Lewis Carroll

ERVs are clearly detrimental to their hosts, but are also important functional components of the genome, so opposing mechanisms must be in place to mitigate their effects as well as maintain their presence and function. Two key characteristics of retroviruses are responsible for the production of ERVs: a life cycle that requires integration into the genome of the host cell, and the unique ability to infect germ cells (albeit a rare event). The observation that distinct *Mus* species contain multiple copies of highly related proviruses prompted the hypothesis that endogenization of a retrovirus occurs in bursts, in which a relatively rapid amplification of ERVs occurs shortly after the initial

colonization event, but then declines over time [2]. These initial bursts can be attributed to two factors: 1) The evolution of mechanisms in the host that suppress further infection and thereby prevent the deleterious consequences of ERV activity, as discussed below; and 2) the absence of any strong selection pressure to maintain ERV sequence integrity during host replication, resulting in a decrease in fitness (i.e. replicative ability) of the viral lineage over time. Because the accumulation of mutations from one ERV generation to the next is irreversible, eventually all endogenous viruses become defective, fail to replicate or express any gene products, and the lineage becomes extinct. While this protects the host from the consequences of ERV insertions like disease and mutation, it is becoming clear that these protections are counteracted by mechanisms that work to maintain active ERVs to support their function as an important source of somatic and genomic diversity.

Mechanisms supporting the maintenance of active ERVs and de novo endogenization

In the face of host resistance factors and the accumulation of inactivating mutations, ERVs have evolved mechanisms that either postpone inactivation or re-establish replication fitness. Two important mechanisms by which replication fitness is prolonged have been revealed by the analysis of mouse ERVs: 1) the switch from an extra- to intracellular life cycle that does not require receptor interaction and 2) *trans* complementation by other ERVs or XRVs. Two successful ERVs that have adapted quite efficiently to an intracellular life cycle are IAPs and MusD, members of the Class II ERV family. The transition from an infectious IAPE to intracellular-restricted IAP has been shown to coincide with loss of a functional *env* gene and the simultaneous gain of an endoplasmic reticulum signal and loss of myristylation signals in the MA protein that determine the site of virus assembly [49, 78, 79]. Similarly, the MA proteins of MusD also have lost plasma-membrane targeting signals [45]; *env* containing ancestors of MusD have not been identified. Examples of the second common mechanism, *trans* complementation, are clearly observed in the highly repetitive and non-autonomous MaLR and VL30 elements found in the mouse genome. Such degenerate ERV sequences that do not code for functional *pol* sequences can continue to replicate provided that their *cis*-acting elements, permitting packaging and DNA integration, remain intact and the proteins required for replication are supplied by infecting XRVs or by functional, or partially functional, ERVs within the same cell. Importantly, this does not increase the fitness of ERVs – but merely allows defective viruses to

continue replicating. It is only through new rounds of endogenization that active ERV proviral genomes can be maintained.

New ERVs may arise within genomes by at least two different mechanisms: retrotransposition from a pre-existing endogenous retrovirus (intraspecific transmission) or infection and integration via an exogenous source virus (horizontal transmission). For defective ERVs, replicative capacity (i.e. full fitness) can be restored via recombination at the RNA level either with XRV or with another expressed ERV. The frequency of these recombinational events is clearly facilitated by the retroviral diploid RNA genome and replication strategy. Analysis of ERV genomes in the mouse genome has revealed multiple examples in which exchanges between different ERV lineages have been observed. Indeed many of the ERV lineages were first discovered due to incorporation of their sequences into other better-characterized ERV lineages, such as the acquisition of LTR sequences from either GLN or MuRRS by VL30 or MuLV elements [72, 80]. Notably, LTR and *env* sequences have been found to be quite variable in different ERV lineages. Examples of LTR exchanges have been investigated in detail for VL30 elements [81]. Such exchanges would be expected to increase expression rates (and thus increase the chance of *de novo* endogenization) in particular cell types. Indeed, it is tempting to speculate that the high expression of several mouse ERVs (e.g. IAP and ERV-L) during embryogenesis may increase the likelihood of germ cell infections [26, 82].

The acquisition of novel *env* sequences is best documented for the Class I ERVs [83]. As mentioned above and depicted in Table 1, members of both MuLV and McERV families have altered host ranges due to genetic mutations or recombination events within their *env* genes. The acquisition of novel sequences within the variable regions of the SU domain may contribute to overcoming host defence mechanisms that block cell entry (see below) or may result in usage of novel receptors. The latter would be expected to either facilitate infection of different cell types within the host or may increase the host spectrum of the retrovirus, and thereby facilitate infection of other species. It should be noted, however, that there are clearly limits on receptor switches. All six known gammaretrovirus receptors are transporters with multiple transmembrane domains. This common structural feature suggests constraints that may define how these related viruses interact with their receptors to bind, to produce the conformational changes that allow for Env interactions and cell fusion, and to penetrate the cell.

The importance of cross-species transmission in maintaining viable ERV genomes or supporting *de novo* endogenization has long been recognized based on comparative analysis of host and viral genomes [84–86]. Conceivably, cross-species infections may allow rapid expansion of the XRV in its new host, which may not have developed fine-tuned defence mechanisms against the invading parasite (see [87], this issue). Sequences related to Class I ERVs have been identified in a variety of terrestrial vertebrates suggesting several zoonotic transmissions [86], and several members of this group are also clearly related to infectious retroviruses, such as GALV, KoRV, and FeLV, found in other species (Fig. 2 and Table 1). Although the absence of active ERV elements in the human genome may reflect both changes in host ecology (e.g. decreased exposure to XRV from other animals) plus a strong retroviral defence mechanism, it should be noted that chronic cross-species infection of an X-MuLV in humans (XMRV) has recently been reported [88, 89].

Host mechanisms that suppress retrovirus infection and favor adaptive coexistence

Mice exposed to infectious retroviruses or to unchecked ERV amplification or expression are subject to virus-induced disease or virus-induced genetic mutations. Mice are protected by the innate and acquired immune systems, but have also evolved numerous constitutively expressed antiviral factors that target various stages of the retroviral life cycle, defined largely in studies with the Class I MuLVs. The factors responsible for this intrinsic immunity block virus entry, interfere with specific post-entry stages of virus replication, integration, assembly or release, or mediate endogenization and replicative silencing (Fig. 4). These host resistance factors also produce the selective pressures that favor the outgrowth of virus variants able to circumvent those blocks. This results in a ratchet-like pattern of sequential mutations in both host and virus that generate substantial polymorphism in the critical regions of the responsible genes. This coevolution of virus and host also results in a form of adaptive coexistence in which retroviruses are rarely pathogenic in their natural hosts.

The best-characterized mechanisms of resistance to mouse retroviruses rely on inhibition of virus entry. Entry can be blocked by mutations in the receptor gene that alter their function or by host resistance genes that interfere with receptor function. To date, seven receptors for mouse retroviruses have been identified (Table 1), of which six represent receptors for Class I gammaretroviruses and four are used by different host range subclasses of MuLVs. Studies on two of these receptors have identified residues critical

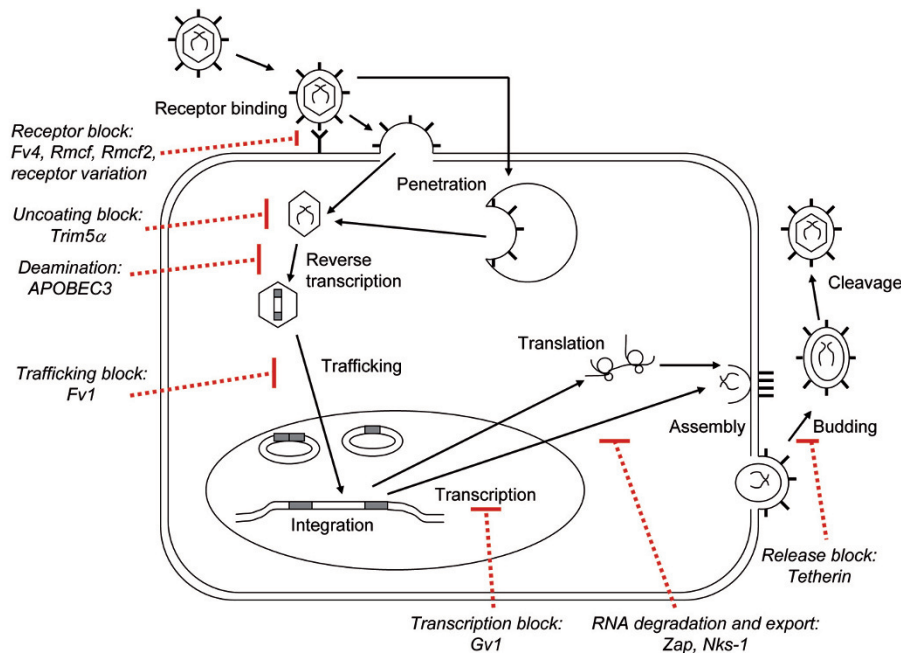


Figure 4. Retrovirus lifecycle and blocks to replication by host restriction factors.

for virus entry in the E-MuLV CAT1 receptor [90, 91] and in the X/P-MuLV receptor XPR1 [92] (Fig. 5). In both cases, these receptor regions are hypervariable and glycosylated, but it is not clear if the identified critical amino acids actually mediate virus attachment and entry or if they are negative regulators that protect nearby highly conserved receptor determinants [54]. Among mouse species, there are two naturally occurring functional variants of the CAT-1 receptor [55, 93], and four variants of XPR1 [28, 92]. These functional subtypes are distinguished by their ability to mediate entry of different virus isolates (Fig. 5). The evolution of XPR1 can be examined in relation to the appearance and spread of X-MuLVs and P-MuLVs in *Mus* species [28, 34]. The XPR1 variant with the broadest susceptibility phenotype is widely distributed among the older Asian *Mus* species. It was not, however, until the appearance of the house mouse (*M. musculus*) about 1 Mya that mice were exposed to and began to acquire ERVs of X-MuLVs. This acquisition is also coincident with the appearance of restrictive XPR1 variants in *M. m. castaneus* (cXPR1) and *M. m. domesticus* (nXPR1). The *M. m. domesticus* mice, like the laboratory strains they gave rise to, are not infectible by X-MuLVs, but the appearance of this restrictive nXPR1 variant undoubtedly contributed to the evolution of P-MuLVs as a new host range variant able to use this restrictive receptor. This coevolutionary relationship based on the patterns of receptor evolution and the pattern of endogenous X/P-MuLV distribution in wild mouse species is consistent with recent phylogenetic analysis

in the sequenced mouse genome. This analysis confirms the close evolutionary relationship of these virus types and suggests X-MuLV is the most recent common ancestor [17].

In addition to receptor polymorphisms, other mechanisms can inhibit infection at the level of virus entry. First, variable receptor glycosylation can result in resistance. Glycosylation has been found to be responsible for resistance to E-MuLV in rodent cells and X-MuLVs in hamster cells [95–98] (Fig. 5). Second, Env-producing ERVs can interfere with exogenous viruses that use the same receptor. Such interfering Env genes have been found for E-MuLVs (*Fv4*) [99] and X/P-MuLVs (*Rmcf*, *Rmcf2*) [100–102]. All three of these resistance genes are associated with ERVs that have intact *env* genes. Two of these genes evolved in wild mouse populations exposed to infectious virus; *M. castaneus* carries *Fv4* and *Rmcf2* along with infectious E-MuLV and X-MuLV [103]. Other host factors may also inhibit gammaretrovirus infection by targeting Env or affecting other factors needed for binding or entry. For example, mice produce a serum factor, LVIF, which inactivates viruses of the P/X-MuLV host range class but not those of other Env host range groups [104]. Also, although no mouse retrovirus is known to use a coreceptor, the efficiency of entry can be modulated by other factors [105]. For example, neurotropic Friend virus infects brain capillaries more efficiently because of the presence of heparin sulfate [106].

Viruses have evolved mechanisms to evade these entry restrictions. The major determinant of receptor

A

	ECL3	MoMuLV	AKV-MuLV	Spl574
mCAT-1	VKGSIKNWQLTEK---NFS---CNNNDTNVKY-GEGGF ^M	+++	+++	+
dCAT-1	...V.....G.....	++	+++	+++
haCAT-1EDFL.R.SPL.G.....H.....	-*	-*	-
ratCAT-1E.....K.SPL.G.....	++	ND	ND
xcCAT-1K.....G.D.....	++	ND	ND
huCAT-1	...V.....EDFG.T.GPL.L...KEGKPGV....	-	-	-

B

	ECL3	ECL4	X-MuLV	P-MuLV	CasE#1MuLV
dXPR1	// SKGLLPNDPQEPEFC// <u>HKEQN</u> HSDTV	SITATTFKPHVGD	+++	+++	+++
nXPR1E.....N	-	+++	-
cXPR1	+++	-	-
pXPR1G.....P.YKV.....	+++	-	+++
haXPR1	<u>N.S</u>L.....G.....A.Q.....	(+)*	-	-
huXPR1NSE.SGI.....RG.....MV.S...LM...T...	+++	++	+++

Figure 5. Amino acid sequences of the extracellular loops (ECLs) of CAT-1 and XPR1 that contain MuLV receptor determinants. (A) Sequences represent NIH 3T3 (mCAT-1), *M. dunni* (dCAT-1), hamster (haCAT-1), rat (ratCAT-1), rat XC cell (xcCAT-1), and human (huCAT-1) [55, 93, 94]. (B) Sequences represent *M. dunni* (dXPR1), NIH 3T3 (nXPR1), *M. castaneus* (cXPR1), *M. pahari* (pXPR1), hamster (haXPR1), and human (huXPR1) [28, 92]. Shown are residues 416–430 and 499–508 in ECL3. Relative susceptibilities are given on the right for 3 E-MLVs for CAT-1 and for 3 X/P-MLVs for XPR1; asterisks identify infections that can be enhanced by inhibitors of N-linked glycosylation. Critical residues in CAT1 and XPR1 are underlined; N-linked glycosylation sites are shaded. ND, not done.

recognition (and thus virus host range) in Class I ERVs is within the variable region of the surface (SU) domain of the Env protein [107]. Mutations within this region can lead to the generation of retroviruses that can use alternative receptors, multiple receptors or even multiple receptor determinants on the same protein; this contributes to virus survival by limiting the impact of host escape mutations. Thus, the closely related phosphate transporters PiT1 or PiT2 both function as gammaretrovirus receptors, and at least one MuLV, mouse 10A1 MuLV, can use either for entry [108]. Also, the XPR1 receptor carries two largely independent determinants in different extracellular loops that can mediate X-MuLV entry and undefined determinants for P-MuLV entry [28, 92, 109]. Several mechanisms also allow viruses to bypass the need for their cognate receptor. In viremic animals, recombination between infectious E-MuLV or A-MuLV and endogenous P-MuLV *env* genes can generate novel recombinants with P-MuLV host range [110, 111]. These P-MuLVs are also often pseudotyped by E-MuLVs in infected animals. It has also been shown that some viruses can use alternative receptors in the presence of the soluble Env glycoprotein for that receptor. P-MuLVs can be transactivated in this way by E-MuLV Env [112].

In addition to resistance mechanisms that block virus entry, there are three major genes known to inhibit early post entry replication of retroviruses: *FvI*, *TRIM5 α* , and *APOBEC3* (Fig. 4). Alleles at the *FvI* locus control the relative sensitivities of mouse cells to different subgroups of MuLVs [113]. There are at least five allelic variants of *FvI*, four of which produce different patterns of resistance to mouse-tropic virus-

es, described as N-, B-, NR, or NB-tropic. The *FvI* sequence is related to the *gag* gene of MuERV-L [114]. The mechanism of restriction is unknown, but *FvI* generally blocks virus replication at or just after reverse transcription to limit proviral integrations [115]. *FvI* targets the viral capsid; the major determinant that distinguishes N- and B-tropic viruses is at CA position 110 [116], but other targets in this same CA region have been identified in studies on NR- and NB-tropism [117]. *FvI* restriction is found only in laboratory mice and laboratory mouse-related *Mus* species [118, 119], but African pygmy mice carry an unusual post-entry resistance to some ecotropic MuLVs that targets some of the same amino acid residues as *FvI* [96].

While cells of nonrodent species lack *FvI*, they carry another restriction gene, *TRIM5 α* , that blocks retrovirus replication before reverse transcription. *TRIM5 α* is a member of the tripartite interaction motif family, and no mouse ortholog has been identified. Primate *TRIM5 α* , however, restricts MuLV infection by targeting amino acids at two of the capsid sites also targeted by *FvI*, 110 and 117 [120]. *TRIM5 α* may thus function to limit transspecies transmission by the broadly infectious MuLVs. The evolution of two polymorphic host genes (*FvI* and *TRIM5 α*) targeting the same hypervariable capsid segment also suggests that, like the viral Env and their receptors, virus capsid and host genes that target capsid represent a second major battleground in the coevolution of virus and host.

A third host gene involved in post-entry restriction to retroviral infection is *APOBEC3G*. This enzyme catalyzes C-to-U deamination during reverse tran-

scription, resulting in G-to-A mutations in the resulting provirus [121]. Although most extensively studied in humans where it inhibits HIV as well as MuLVs, the mouse counterpart, mA3, also blocks HIV-1 and the mouse retroviruses MMTV, IAP, and MusD [122–124]. Although mA3 has now been identified as Friend MuLV resistance gene *Rfv3* [125] and a recent study characterized MuLV ERVs in C57BL/6 mice and found that a significant number of the observed mutations in endogenous P-MuLVs, but not X-MuLVs, represented G-to-A mutations and also showed a gradient of G-to-A mutations consistent with mA3 activity [17]. This suggests that mA3 editing may have contributed to endogenization and silencing of these ERVs at the time of integration.

A few mouse genes have been identified that inhibit late stages of retrovirus replication. One such gene, *Gv1*, is known to broadly restrict transcription of endogenous MuLVs [126]. This restriction was shown to be due to a single host gene that was mapped to chromosome 13 [127], but it has not been characterized further. Transcriptional silencing can also be effected by TRIM28 which targets the MuLV primer binding site [128]. Another late-acting gene, *Nks1*, encodes an mRNA export factor that suppresses sense-oriented intronic IAP inserts [129]. Also, host genomes can silence ERVs by methylation; failure to methylate new or specific ERVs has been attributed to *Dnmt* genes [130].

Another factor involved in intrinsic immunity to late stages of retroviral infection is ZAP, a zinc-finger antiviral protein that inhibits translation of MuLVs as well as alphaviruses and filoviruses [131]. ZAP binds viral RNA and leads to its degradation. ZAP is found in multiple species, and sequence comparisons confirm that it has evolved under positive selection. Finally, recent studies have identified a novel factor in human cells that inhibits release of many enveloped viruses, including MuLV [132]. Tetherin binds newly formed virions to cell surfaces, is constitutively expressed by some cells, and can be antagonized by the HIV-1 accessory protein Vpu.

The role of ERVs in shaping the mouse genome

The data from the sequenced mouse genome together with a century of mouse genetic analysis has established that ERVs have played and continue to play an important role in shaping the mouse genome. These ERVs modify the mouse phenome by altering host gene expression or by contributing novel protein-coding sequences. The great majority of ERV-induced changes are insertional mutations that disrupt host protein-coding genes or alter gene expression by affecting splicing or by providing novel signals for initiation, regulation or termination of

transcription. ERV insertions in somatic cells have long been studied for their ability to induce neoplastic diseases by activating oncogenes. Insertional mutations also accumulate in the germ line, and a recent review identified 63 examples of ERV-induced mouse mutations and argued that 10–12% of all mutations in the mouse are caused by ERVs [133]. The majority of these germ line mutations are due to IAP and ETn/MusD insertions, which is not surprising as these ERV families are present in high copy numbers and contain many active members. The largest number of these mutational insertions is found in introns, and occasional reversions have been reported which are accompanied by provirus deletion. In addition to insertional mutagenesis, ERV sequences may promote rearrangement of DNA by way of non-allelic homologous recombination between elements [2].

The first example of ERV-induced insertional mutagenesis was identified in the oldest inbred strain of the laboratory mouse, DBA, which was bred to carry three visible mutations all affecting coat color: dilute (*d*), brown (*b*), and agouti (*a*). The single ecotropic MuLV ERV in this strain was linked to the *d* locus, and it was found that this ERV was lost in rare color revertants [134]. The *d* locus encodes an unconventional myosin heavy chain gene, Myosin Va, that controls transport of melanosomes into the dendritic processes of melanocytes from where they enter the hair shaft. The ERV integration site is within an intron of this gene and results in abnormally spliced transcripts [135]; this is associated with the failure to produce pigment-filled dendritic processes and an abnormal distribution of pigment in the hair shaft leading to lightened coat color.

The importance of ERV insertions for gene regulation is well recognized. An early and dramatic example was the demonstration that the mouse gene for sex-limited protein requires androgen for its expression because of a hormone-responsive enhancer introduced by ERV insertion [136]. More recent work has provided evidence that MuERV-L and MaLR elements may provide essential functions as ‘early’ promoters in full-grown oocytes and during embryogenesis [26]. The temporal and spatial regulation of ERV elements are controlled not only by the differential expression of transcription factors that regulate LTR expression, but by epigenetic modifications. ERVs are often preferential sites of methylation and can provide methylation-sensitive promoters for host genes [137]; this can result in variable expression, as is the case for the viable yellow agouti allele that uses an IAP LTR as promoter [138, 139]. Furthermore, it is increasingly appreciated that the stochastic nature of ERV expression modulated by epigenetic modifications dur-

ing embryogenesis may also impinge on neighboring genes by RNA interference [140, 141].

In addition to altering gene regulation or expression, ERV integrations also insert virus protein-coding genes into host chromosomes, and some of these genes have been co-opted or domesticated by the host for cellular functions. Such genes are rare, and the majority of the recognizable co-opted virus genes serve to protect against or modulate further retroviral infection. The largest set of these genes are the minor lymphocyte-stimulating (*Mls*) genes. These *Mls* genes contain integrated MMTVs that express an LTR-encoded *sag* gene. The *sag* product is a transmembrane glycoprotein that functions as a superantigen to stimulate T cell proliferation and the deletion of specific T cell subsets. Studies on transgenic mice expressing an exogenous MMTV indicated that Sag-induced T-cell depletion prevents infection by exogenous MMTV of the same *sag* gene specificity [142]. Subsequent analysis of the 3 endogenous MMTVs of BALB/c mice indicated that *Mtv*-encoded Sag is needed for exogenous MMTV infection, further suggesting that interfering with *sag* function alters susceptibility to MMTV infection [143].

Another set of co-opted retroviral genes serve an antiviral function through production of MuLV Env glycoproteins. As discussed above, the products of these genes, *Fv4*, *Rmcf*, and *Rmcf2*, interfere with exogenous infection by MuLVs of the same host range subgroup [99, 101, 102]. It is thought that these Env glycoproteins and their receptors associate in the endoplasmic reticulum to prevent processing and transport to the plasma membrane. Of the three identified interfering Env genes, two were found in a single Asian species, *M. castaneus*, which harbors infectious viruses of the corresponding host range types [103]. Comparable interfering ERV *env* genes have been found in chickens, sheep, and cats, suggesting that this is a common host resistance mechanism. The oldest known retroviral resistance gene, *Fv1*, is also a co-opted ERV sequence, [144]. The *Fv1* sequence is responsible for resistance to MuLVs and is found in all *Mus* species [145]. However, the older species carry a non-restrictive null allele, while the restrictive alleles are found only in more recently diverged *M. musculus* subspecies, animals that are exposed to infectious MuLVs [119].

ERVs have also contributed coding sequences for at least two genes with functions other than virus resistance. The mouse has two syncytin genes, SynA and SynB, that are responsible for the cell fusions that produce the trophoblast during development of the placenta. These genes are ERV *env* genes [146]. Interestingly, humans also carry two syncytin genes with similar function, but the human and mouse genes

are not at orthologous positions in the germ line, suggesting that ERVs were independently co-opted to serve similar functions in these different mammalian lineages.

Conclusions

Mouse XRVs were first identified in 1951 with the discovery that extracts from tumors or embryos could induce leukemia [147]. It was subsequently determined that uninfected mouse embryo cells could produce infectious virus [148]. Liquid DNA hybridization showed that cells carried viral sequences [149], suggesting that these XRVs were the products of ERVs rather than latent exogenous virus infections. Since that time, the number of identified ERV copies has progressively grown along with the number of distinct families, aided by the publication of the mouse genome sequence and increasing use of wild mouse species in addition to inbred laboratory mouse strains. The realization that a staggering proportion of the genome is ERV-related and that some ERVs serve key host functions has revitalized interest in the evolutionary forces that have shaped this host-virus co-dependency. While it is clear that newly emerging XRVs are likely to be mutagenic and pathogenic in their new hosts, long-term associations have clearly produced coevolutionary adaptations that tend to mitigate these deleterious effects. While some domesticated ERVs move toward extinction, others remain active over long evolutionary intervals, suggesting that all this activity is neither necessarily damaging nor entirely superfluous for the host. While ERVs and the transposable elements from which they evolved obviously contribute to genetic diversity on an evolutionary timescale, it can also be argued that these elements may have important roles in development, as shown for example by L1 contributions to neuronal diversity in the developing brain [150] or by the developmentally regulated high level of expression of MuERV-L and MaLR ERVs in early embryogenesis [26, 151]. Because work on murine ERVs and XRVs initially focused on the most pathogenic groups and was limited to studies on the inbred laboratory strains, we are only beginning to identify and characterize the larger and older families present in the mouse germ line, and to describe what may be common and widespread interspecies transmissions of Class I ERVs/XRVs. Further work on these elements should produce novel insights into the evolutionary past and refine our understanding of the present interdependence of these genomic parasites and their hosts.

Acknowledgements. This work was supported in part by the Intramural Research Program of the NIH, NIAID. The Heinrich-Pette-Institut is supported by the Freie und Hansestadt Hamburg and the Bundesministerium für Gesundheit und soziale Sicherung.

- 1 Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562.
- 2 Boeke, J. and Stoye, J. (1997) Retrotransposons, endogenous retroviruses, and the evolution of retroelements. In: *Retroviruses*, pp. 343–436. Coffin, J., Hughes, S. and Varmus, H. (eds.), Cold Spring Harbor Laboratory Press, New York.
- 3 Stoye, J. (2001) Endogenous retroviruses: still active after all these years? *Curr. Biol.* 11, R914–R916.
- 4 Guenet, J.-L. and Bonhomme, F. (2003) Wild mice: an ever-increasing contribution to a popular mammalian model. *Trends Genet.* 19, 24–31.
- 5 Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J., 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.* 9, 411–412.
- 6 DeBerardinis, R., Goodier, J., Ostertag, E. and Kazazian, H. J. (1998) Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat. Genet.* 20, 288–290.
- 7 Kramerov, D. and Vassetzky, N. (2005) Short retroposons in eukaryotic genomes. *Int. Rev. Cytol.* 247, 165–221.
- 8 Malik, H. and Eickbush, T. (2001) Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. *Genome Res.* 11, 1187–1197.
- 9 Doolittle, R., Feng, D., Johnson, M. and McClure, M. (1989) Origins and evolutionary relationships of retroviruses. *Q. Rev. Biol.* 64, 1–30.
- 10 Biémont, C. and Vieira, C. (2006) Genetics: junk DNA as an evolutionary force. *Nature* 443, 521–524.
- 11 Jern, P., Sperber, G. and Blomberg, J. (2005) Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. *Retrovirology* 2, 50.
- 12 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.
- 13 Faquet, C., Mayo, M., Maniloff, J., Desselberger, U. and Ball, L. (2004) *Virus Taxonomy: The Eighth Report of the International Committee on Taxonomy of Viruses*. Academic Press, San Diego, CA.
- 14 Bénéit, L., Lallemand, J.-B., Casella, J.-F., Philippe, H. and Heidmann, T. (1999) ERV-L elements: a family of endogenous retrovirus-like elements active throughout the evolution of mammals. *J. Virol.* 73, 3301–3308.
- 15 Wilkinson, D., Mager, D. and Leong, J. (1994) Endogenous human retroviruses. In: *The Retroviridae*, vol. 1, Levy, J. (ed.), Plenum Press, New York.
- 16 Gifford, R., Kabat, P., Martin, J., Lynch, C. and Tristem, M. (2005) Evolution and distribution of class II-related endogenous retroviruses. *J. Virol.* 79, 6478–6486.
- 17 Jern, P., Stoye, J. P. and Coffin, J. M. (2007) Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. *PLoS Genet.* 3, 2014–2022.
- 18 Ribet, D., Harper, F., Pierron, G. and Heidmann, T. (2008) The GLN family of mouse endogenous retroviruses comprises an element competent for infectious viral particle formation. *J. Virol.* 82, 4413–4419.
- 19 McCarthy, E. and McDonald, J. (2004) Long-terminal repeat retrotransposons of *Mus musculus*. *Genome Biol.* 5, R14.
- 20 Bromham, L., Clark, F. and McKee, J. (2001) Discovery of a novel murine type C retrovirus by data mining. *J. Virol.* 75, 3053–3057.
- 21 International Human Genome Sequencing Consortium (2001) Initial sequence and analysis of the human genome. *Nature* 409, 860–921.
- 22 Smit, A. F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.* 9, 657–663.
- 23 Costas, J. (2003) Molecular characterization of the recent intragenomic spread of the murine endogenous retrovirus MuERV-L. *J. Mol. Evol.* 56, 181–186.
- 24 Ribet, D., Louvet-Vallée, S., Harper, F., de Parseval, N., Dewannieux, M., Heidmann, O., Pierron, G., Maro, B. and Heidmann, T. (2008) Murine endogenous retrovirus MuERV-L is the progenitor of the ‘orphan’ epsilon viruslike particles of the early mouse embryo. *J. Virol.* 82, 1622–1625.
- 25 Svoboda, P., Stein, P., Anger, M., Bernstein, G., Hannon, G. and Schultz, R. (2004) RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. *Dev. Biol.* 269, 276–285.
- 26 Peaston, A., Esikov, A., Graber, J., de Vries, W., Holbrook, A., Solter, D. and Knowles, K. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* 7, 597–606.
- 27 Baldo, A. M. and McClure, M. A. (1999) Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. *J. Virol.* 73, 7710–7721.
- 28 Yan, Y., Knoper, R. C. and Kozak, C. A. (2007) Wild mouse variants of envelope genes of xenotropic/polytropic mouse gammaretroviruses and their receptors elucidate receptor determinants of virus entry. *J. Virol.* 81, 10550–10557.
- 29 Obata, M. M. and Khan, A. S. (1988) Structure, distribution, and expression of an ancient murine endogenous retrovirus-like DNA family. *J. Virol.* 62, 4381–4386.
- 30 Eicher, E. M., Hutchinson, K. W., Phillips, S. J., Tucker, P. K. and Lee, B. K. (1989) A repeated segment on the mouse Y chromosome is composed of retroviral-related, Y-enriched and Y-specific sequences. *Genetics* 122, 181–192.
- 31 Callahan, R., Drohan, W., Gallahan, D., D’Hoostelaere, L. and Potter, M. (1982) Novel class of mouse mammary tumor virus-related DNA sequences found in all species of *Mus*, including mice lacking the virus proviral genome. *Proc. Natl. Acad. Sci. USA* 79, 4113–4117.
- 32 Courtney, M. G., Elder, P. K., Steffen, D. L. and Getz, M. J. (1982) Evidence for an early evolutionary origin and locus polymorphism of mouse VL30 DNA sequences. *J. Virol.* 43, 511–518.
- 33 Inaguma, Y., Miyashita, N., Moriwaki, K., Huai, W. C., Mei-Lei, J., Xinqiao, H. and Ikeda, H. (1991) Acquisition of two endogenous ecotropic murine leukemia viruses in distinct Asian wild mouse populations. *J. Virol.* 65, 1796–1802.
- 34 Kozak, C. and O’Neil, R. (1987) Diverse wild mouse origins of xenotropic, mink-cell focus-forming, and two types of ecotropic proviral genes. *J. Virol.* 61, 3082–3388.
- 35 Mager, D. and Freeman, J. (2000) Novel mouse type D endogenous proviruses and ETn elements share long terminal repeat and internal sequences. *J. Virol.* 74, 7221–7229.
- 36 Sonigo, P., Wain-Hobson, S., Bougueleret, L., Tiollais, P., Jacob, F. and Brület, P. (1987) Nucleotide sequence and evolution of *ETn* elements. *Proc. Natl. Acad. Sci. USA* 84, 3768–3771.
- 37 Schmidt, M., Wirth, T., Kröger, B. and Horak, I. (1985) Structures and genomic organization of a new family of murine retrovirus-related DNA sequences (MuRRS). *Nucleic Acids Res.* 13, 3461–3470.
- 38 Smit, A. F. A. (1993) Identification of a new, abundant superfamily of mammalian LTR-transposons. *Nucleic Acids Res.* 21, 1863–1872.
- 39 Vogt, V. (1997) Retroviral virions and genomes. In: *Retroviruses*, pp. 343–436. Coffin, J., Hughes, S. and Varmus, H. (eds.), Cold Spring Harbor Laboratory Press, New York.
- 40 Kozak, C. A. and Ruscetti, S. (1994) Retroviruses in rodents. In: *The Retroviridae*, vol. 3, Levy, J. (ed.), Plenum Press, New York.

- 41 Michalides, R., van Nie, R., Nusse, R., Hynes, N. and Groner, B. (1981) Mammary tumor induction loci in GR and DBA/f mice contain one provirus of the mouse mammary tumor virus. *Cell* 23, 165–173.
- 42 Medstrand, P. and Mager, D. (1998) Human-specific integrations of the HERV-K endogenous retrovirus family. *J. Virol.* 72, 9782–9787.
- 43 Belshaw, R., Dawson, A., 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.
- 44 Brûlet, P., Kaghad, M., Xu, Y., Croissant, O. and Jacob, F. (1983) Early differential tissue expression of transposon-like repetitive DNA sequences of the mouse. *Proc. Natl. Acad. Sci. USA* 80, 5641–5645.
- 45 Ribet, D., Harper, F., Dewannieux, M., Pierron, G. and Heidmann, T. (2007) Murine MusD retrotransposon: structure and molecular evolution of an 'intracellularized' retrovirus. *J. Virol.* 81, 1888–1898.
- 46 Baust, C., Gagnier, L., Baillie, G., Harris, M., Juriloff, D. and Mager, D. (2003) Structure and expression of mobile ETnII retroelements and their coding-competent MusD relatives in the mouse. *J. Virol.* 77, 11448–11458.
- 47 Kuff, E. and Lueders, K. (1988) The intracisternal A-particle gene family: structure and functional aspects. *Adv. Cancer Res.* 51, 183.
- 48 Reuss, F. (1992) Expression of intracisternal A-particle-related retroviral element-encoded envelope proteins detected in cell lines. *J. Virol.* 66, 1915–1923.
- 49 Ribet, D., Harper, F., Dupressoir, A., Dewannieux, M., Pierron, G. and Heidmann, T. (2008) An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. *Genome Res.* 18, 597–609.
- 50 Benveniste, R., Callahan, R., Sherr, C., Chapman, V. and Todaro, G. (1977) Two distinct endogenous type C viruses isolated from the asian rodent *Mus cervicolor*: conservation of virogene sequences in related rodent species. *J. Virol.* 21, 849–862.
- 51 Lieber, M., Sherr, C., Todaro, G., Benveniste, R., Callahan, R. and Coon, H. (1975) Isolation from the asian mouse *Mus caroli* of an endogenous type C virus related to infectious primate type C viruses. *Proc. Natl. Acad. Sci. USA* 72, 2315–2319.
- 52 Volgamot, G., Bonham, L. and Miller, A. (1998) Sequence analysis of *Mus dunni* endogenous virus reveals a hybrid VL/30 gibbon ape leukemia virus-like structure and a distinct envelope. *J. Virol.* 72, 7459–7466.
- 53 Miller, A., Bergholz, U., Ziegler, M. and Stocking, C. (2008) Identification of the myelin protein plasmolipin as the cell-entry receptor for *Mus caroli* endogenous retrovirus. *J. Virol.* 82, 6862–6868.
- 54 Taylor, C. S., Nouri, A., Lee, C. G., Kozak, C. A. and Kabat, D. (1999) A cell surface receptor for xenotropic and polytropic murine leukemia viruses and characterized. *Proc. Natl. Acad. Sci. USA* 96, 927–932.
- 55 Albritton, L., Tseng, L., Scadden, D. and Cunningham, J. (1989) A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell* 57, 659–666.
- 56 Miller, D., Edwards, R. and Miller, A. (1994) Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *Proc. Natl. Acad. Sci. USA* 91, 78–82.
- 57 Miller, D. and Miller, A. (1994) A family of retroviruses that utilize related phosphate transporters for cell entry. *J. Virol.* 68, 8270–8276.
- 58 Hein, S., Prassolov, V., Zhang, Y., Ivanov, D., Löhler, J., Ross, S. and Stocking, C. (2003) Sodium-dependent myo-inositol transporter 1 is a cellular receptor for *Mus cervicolor* M813 murine leukemia virus. *J. Virol.* 77, 5926–5932.
- 59 Tipper, C., Bencsics, C. and Coffin, J. (2005) Characterization of *Hortulanus* endogenous murine leukemia virus, and endogenous provirus that encodes an infectious murine leukemia virus of a novel subgroup. *J. Virol.* 79, 8316–8329.
- 60 Kozak, C., Hartley, J. and Morse 3rd, H. (1984) Laboratory and wild-derived mice with multiple loci for production of xenotropic murine leukemia virus. *J. Virol.* 51, 77–80.
- 61 Levy, J. and Pincus, T. (1970) Demonstration of biological activity of a murine leukemia virus of New Zealand black mice. *Science* 170, 326–327.
- 62 Frankel, W. N., Stoye, J. P., Taylor, B. A. and Coffin, J. M. (1990) A linkage map of endogenous murine leukemia proviruses. *Genetics* 124, 221–236.
- 63 Jenkins, N. A., Copeland, N. G., Taylor, B. A. and Lee, B. K. (1982) Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus*. *J. Virol.* 43, 26–36.
- 64 Alamgir, A., Owens, N., Lavignon, M., Malik, F. and Evans, L. (2005) Precise identification of endogenous proviruses of NFS/N mice participating in recombination with Moloney ecotropic murine leukemia virus (MuLV) to generate polytropic MuLVs. *J. Virol.* 79, 4664–4671.
- 65 Khan, A. (1984) Nucleotide sequence analysis establishes the role of endogenous murine leukemia virus DNA segments in formation of recombinant mink cell focus-forming murine leukemia viruses. *J. Virol.* 50, 864–871.
- 66 Voytek, P. and Kozak, C. (1988) HoMuLV: a novel pathogenic ecotropic virus isolated from the European mouse, *Mus hortulanus*. *Virology* 165, 469–475.
- 67 O'Neill, R., Hartley, J., Repaske, R. and Kozak, C. (1987) Amphotropic proviral envelope sequences are absent from the *Mus* germ line. *J. Virol.* 61, 2225–2231.
- 68 Zhao, Y., Jacobs, C., Wang, L. and Hardies, S. (1999) MuERVC: a new family of murine retrovirus-related repetitive sequences and its relationship to previously known families. *Mamm. Genome* 10, 477–448.
- 69 Phillips, S., Birkenmeier, E., Callahan, R. and Eicher, E. (1982) Male and female mouse DNAs can be discriminated using retroviral probes. *Nature* 297, 241–243.
- 70 Hutchison, K. and Eicher, E. M. (1989) An amplified endogenous retroviral sequence on the murine Y chromosome related to the murine leukemia viruses and viruslike 30S sequences. *J. Virol.* 63, 4043–4046.
- 71 Fennelly, J., Harper, K., Laval, S., Wright, E. and Plumb, M. (1996) Co-amplification of tail-to-tail copies of MuRVY and IAPE retroviral genomes on the *Mus musculus* Y chromosome. *Mamm. Genome* 7, 31–36.
- 72 Itin, A. and Keshet, E. (1986) A novel retroviruslike family in the mouse DNA. *J. Virol.* 59, 301–307.
- 73 Wolgamot, G., Bonham, L. and Miller, A. (1998) Sequence analysis of *Mus dunni* endogenous virus reveals a hybrid VL30/gibbon ape leukemia virus-like structure and a distinct envelope. *J. Virol.* 72, 7459–7466.
- 74 Howk, R., Troxler, D., Lowy, D., Duesberg, P. and Scolnick, E. (1978) Identification of a 30S RNA with properties of a defective type C virus in murine cells. *J. Virol.* 25, 115–123.
- 75 French, N. S. and Norton, J. D. (1997) Structure and functional properties of mouse VL30 retrotransposons. *Biochim. Biophys. Acta* 1352, 33–47.
- 76 Rotman, G., Itin, A. and Keshet, E. (1984) 'Solo' large terminal repeats (LTR) of an endogenous retrovirus-like gene family (VL30) in the mouse genome. *Nucleic Acids Res.* 12, 2273–2282.
- 77 Reference removed in proof.
- 78 Rhee, S. and Hunter, E. (1990) A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a type C retrovirus. *Cell* 63, 77–86.
- 79 Fehrmann, F., Jung, M., Zimmermann, R. and Kräusslich, H. (2003) Transport of the intracisternal A-type particle Gag polyprotein to the endoplasmic reticulum is mediated by the signal recognition particle. *J. Virol.* 77, 6293–6304.

- 80 Khan, A. and Martin, M. (1983) Endogenous murine leukemia proviral long terminal repeats contain a unique 190-base-pair insert. *Proc. Natl. Acad. Sci. USA* 80, 2699–2703.
- 81 Itin, A. and Keshet, E. (1986) Diverse long terminal repeats are associated with murine retroviruslike (VL30) elements. *Mol. Cell. Biol.* 6, 1276–1282.
- 82 Poznanski, A. and Calarco, P. (1991) The expression of intracisternal A particle genes in the preimplantation mouse embryo. *Dev. Biol.* 143, 271–281.
- 83 Tailor, C., Lavillette, D., Marin, M. and Kabat, D. (2003) Cell surface receptors for gammaretroviruses. *Curr. Top. Microbio. Immunol.* 281, 29–106.
- 84 Benveniste, R. and Todaro, G. (1974) Evolution of C-type viral genes: inheritance of exogenously acquired viral genes. *Nature* 252, 456–459.
- 85 Benveniste, R. and Todaro, G. (1973) Homology between type-C viruses of various species as determined by molecular hybridization. *Proc. Natl. Acad. Sci. USA* 70, 3316.
- 86 Martin, J., Herniou, E., Cook, J., O'Neill, R. W. and Tristem, M. (1999) Interclass transmission and phyletic host tracking in murine leukemia virus-related retroviruses. *J. Virol.* 73, 2442–2449.
- 87 Young, P. (2008) Biology and evolution of the endogenous Koala retrovirus. *Cell. Mol. Life Sci.*, in press.
- 88 Urisman, A., Molinaro, R., Fischer, N., Plummer, S., Casey, G., Klein, E., Malathi, K., Magi-Galluzzi, C., Tubbs, R., Ganem, D., Silverman, R. et al. (2006) Identification of a novel gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. *PLoS Pathog.* 2, e25.
- 89 Dong, B., Kim, S., Hong, S., Das Gupta, J., Malathi, K., Klein, E., Ganem, D., Derisi, J., Chow, S. and Silverman, R. (2007) An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors. *Proc. Natl. Acad. Sci. USA* 104, 1655–1660.
- 90 Albritton, L. M., Kim, J. W., Tseng, L. and Cunningham, J. M. (1993) Envelope-binding domain in the cationic amino acid transporter determined the host range of ecotropic murine retroviruses. *J. Virol.* 67, 2091–2096.
- 91 Yoshimoto, T., E. Yoshimoto and Meruelo, D. (1993) Identification of amino acid residues critical for infection with ecotropic murine leukemia retrovirus. *J. Virol.* 67, 1310–1314.
- 92 Marin, M., Tailor, C. S., Nouri, A., Kozak, S. L. and Kabat, D. (1999) Polymorphisms of the cell surface receptor control mouse susceptibilities to xenotropic and polytropic leukemia viruses. *J. Virol.* 73, 9362–9368.
- 93 Eiden, M., Farell, K., Warsowe, J., Mahan, L. and Wilson, C. (1993) Characterization of a naturally occurring ecotropic receptor that does not facilitate entry of all ecotropic murine retroviruses. *J. Virol.* 67, 4056–4061.
- 94 Kubo, Y., Ono, T., Ogura, M., Ishimoto, A. and Amanuma, H. (2002) A glycosylation-defective variant of the ecotropic murine retrovirus receptor is expressed in rat XC cells. *Virology* 303, 338–344.
- 95 Eiden, M. V., Farrell, K. and Wilson, C. A. (1994) Glycosylation-dependent inactivation of the ecotropic murine leukemia virus receptor. *J. Virol.* 68, 626–631.
- 96 Yan, Y., Jung, Y. T., Wu, T. and Kozak, C. A. (2008) Role of receptor polymorphism and glycosylation in syncytium induction and host range variation of ecotropic mouse gammaretroviruses. *Retrovirology* 5, 2.
- 97 Miller, D. G. and Miller, A. D. (1992) Tunicamycin treatment of CHO cells abrogates multiple blocks to retrovirus infection, one of which is due to a secreted inhibitor. *J. Virol.* 66, 78–84.
- 98 Tavoloni, N. and Rudenholz, A. (1997) Variable efficiency of murine leukemia retroviral vector on mammalian cells: role of cellular glycosylation. *Virology* 229, 49–56.
- 99 Ikeda, H., Laigret, F., Martin, M. A. and Repaske, R. (1985) Characterization of a molecularly cloned retroviral sequence associated with *Fv-4* resistance. *J. Virol.* 55, 768–777.
- 100 Bassin, R. H., Ruscetti, S., Ali, I., Haapala, D. K. and Rein, A. (1982) Normal DBA/2 mouse cells synthesize a glycoprotein which interferes with MCF virus infection. *Virology* 123, 139–151.
- 101 Jung, Y. T., Lyu, M. S., Buckler-White, A. and Kozak, C. A. (2002) Characterization of a polytropic murine leukemia virus proviral sequence associated with the virus resistance gene *Rmcf* of DBA/2 mice. *J. Virol.* 76, 8218–8224.
- 102 Wu, T., Yan, Y. and Kozak, C. A. (2005) *Rmcf2*, a xenotropic provirus in the Asian mouse species *Mus castaneus*, blocks infection by polytropic mouse gammaretroviruses. *J. Virol.* 79, 9677–9684.
- 103 Chattopadhyay, S. K., Oliff, A. I., Linemeyer, D. L., Lander, M. R. and Lowy, D. R. (1981) Genomes of murine leukemia viruses isolated from wild mice. *J. Virol.* 39, 777–791.
- 104 Wu, T., Lee, C. G., Buckler-White, A. and Kozak, C. A. (2002) Genetic control of a mouse serum lipoprotein factor that inactivates murine leukemia viruses: evaluation of apolipoprotein F as a candidate. *J. Virol.* 76, 279–286.
- 105 Tailor, C., Nouri, A. and Kabat, D. (2000) Cellular and species resistance to murine amphotropic, gibbon ape, and feline subgroup C leukemia viruses is strongly influenced by receptor expression levels and by receptor masking mechanisms. *J. Virol.* 74, 9797–9801.
- 106 Jinno-Oue, A., Oue, M. and Ruscetti, S. (2001) A unique heparin-binding domain in the envelope protein of the neuropathogenic PVC-211 murine leukemia virus may contribute to its brain capillary endothelial cell tropism. *J. Virol.* 75, 12439–12445.
- 107 Battini, J.-L., Heard, J. and Danos, O. (1992) Receptor choice determinants in the envelope glycoprotein of amphotropic, xenotropic, and polytropic murine leukemia viruses. *J. Virol.* 66, 1468–1475.
- 108 Wilson, C. A., Farrell, K. B. and Eiden, M. V. (1994) Properties of a unique form of the murine amphotropic leukemia virus receptor expressed on hamster cells. *J. Virol.* 68, 7697–7703.
- 109 Van Hoeven, N. S. and Miller, A. D. (2005) Use of different but overlapping determinants in a retrovirus receptor accounts for non-reciprocal interference between xenotropic and polytropic murine leukemia viruses. *Retrovirology* 2, 76.
- 110 Fischinger, P. J., Nomura, S. and Bolognesi, D. P. (1975) A novel murine oncornavirus with dual eco- and xenotropic properties. *Proc. Natl. Acad. Sci. USA* 72, 5150–5155.
- 111 Hartley, J. W., Wolford, N. K., Old, L. J. and Rowe, W. P. (1977) A new class of murine leukemia virus associated with development of spontaneous lymphomas. *Proc. Natl. Acad. Sci. USA* 74, 789–792.
- 112 Wensel, D. L., Li, W. and Cunningham, J. M. (2003) A virus-virus interaction circumvents the virus receptor requirement for infection by pathogenic retroviruses. *J. Virol.* 77, 3460–3469.
- 113 Pincus, T., Hartley, J. W. and Rowe, W. P. (1971) A major genetic locus affecting resistance to infection with murine leukemia viruses. I. Tissue culture studies of naturally occurring viruses. *J. Exp. Med.* 133, 1219–1233.
- 114 Best, S., LeTossier, P., Towers, G. and Stoye, J. P. (1996) Positional cloning of the mouse retrovirus restriction gene *FvI*. *Nature* 382, 826–829.
- 115 Sveda, M. M. and Soeiro, R. (1976) Host restriction of Friend leukemia virus: synthesis and integration of the provirus. *Proc. Natl. Acad. Sci. USA* 73, 2356–2360.
- 116 Kozak, C. A. and Chakraborti, A. (1996) Single amino acid changes in the murine leukemia virus capsid protein gene define the target of *FvI* resistance. *Virology* 225, 300–305.
- 117 Stevens, A., Bock, M., Ellis, S., LeTossier, P., Bishop, K. N., Yap, M. W., Taylor, W. and Stoye, J. P. (2004) Retroviral capsid determinants of *FvI* NB and NR tropism. *J. Virol.* 78, 9592–9598.
- 118 Yan, Y. and Kozak, C. A. (2008) Novel post-entry resistance to AKV ecotropic mouse gammaretroviruses in the African pygmy mouse, *Mus minutoides*. *J. Virol.* Epub ahead of print.

- 119 Kozak, C. A. (1985) Analysis of wild-derived mice for the *Fv-1* and *Fv-2* murine leukemia virus restriction loci: a novel wild mouse *Fv-1* allele responsible for lack of host range restriction. *J. Virol.* 55, 281–285.
- 120 Yap, M. W., Nisole, S., Lynch, C. and Stoye, J. P. (2004) *Trim5a* protein restricts both HIV-1 and murine leukemia virus. *Proc. Natl. Acad. Sci. USA* 101, 10786–10791.
- 121 Goila-Gaur, R. and Strebel, K. (2008) HIV-1 vif, APOBEC, and intrinsic immunity. *Retrovirology* 5, 51.
- 122 Bishop, K. N., Holmes, R. K., Sheehy, A. M., Davidson, N. O., Cho, S.-J. and Malim, M. H. (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.* 14, 1392–1396.
- 123 Okeoma, C. M., Lovsin, N., Peterlin, M., B. and Ross, S. R. (2007) APOBEC3 inhibits mouse mammary tumor virus replication in vivo. *Nature* 445, 927–930.
- 124 Esnault, C., Neidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A. J., Heidmann, T. and Schwartz, O. (2005) APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature* 433, 430–433.
- 125 Santiago, M. L., Montano, M., Benitez, R. J., Yonemoto, W., Chesebro, B., Hasenkrug, K. J., and Greene, W. C. (2008) *Apoec3* encodes *Rfv3*, a gene influencing neutralizing antibody control of retrovirus infection. *Science* 321, 1343–1346.
- 126 Levy, D. E., Lerner, R. A. and Wilson, M. C. (1985) The *Gv-1* locus coordinately regulates the expression of multiple endogenous murine retroviruses. *Cell* 44, 289–299.
- 127 Oliver, P. L. and Stoye, J. P. (1999) Genetic analysis of *Gv1*, a gene controlling transcription of endogenous murine polytropic proviruses. *J. Virol.* 73, 8227–8234.
- 128 Wolf, D., Hug, K., and Goff, S. P. (2008) TRIM28 mediates primer binding site-targeted silencing of Lys1,2 tRNA-utilizing retroviruses in embryonic cells. *Proc. Natl. Acad. Sci. USA* 105, 12521–12526.
- 129 Floyd, J. A., Gold, D. A., Concepcion, D., Poon, T. H., Wang, X., Keithley, E., Chen, D., Ward, E. J., Chinn, S. B., Friedman, R. A. et al. (2003) A natural allele of *Nxf1* suppresses retrovirus insertional mutations. *Nat. Genet.* 35, 221–228.
- 130 Howard, G., Eiges, R., Gaudet, F., Jaenisch, R. and Eden, A. (2008) Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene* 27, 404–408.
- 131 Gao, G., Guo, X. and Goff, S. P. (2002) Inhibition of retroviral RNA production by ZAP, a CCG-type zinc finger protein. *Science* 297, 1703–1706.
- 132 Neil, S. J.D., Zang, T. and Bieniasz, P. (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451, 425–430.
- 133 Maksakova, I. A., Romanish, M. T., Gagnier, L., Dunn, C. A., van de Lagemaat, L. N. and Mager, D. L. (2006) Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet.* 2, e2.
- 134 Jenkins, N. A., Copeland, N. G., Taylor, B. A. and Lee, B. K. (1981) *Dilute* (d) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. *Nature* 293, 370–374.
- 135 Separack, P. K., Mercer, J. A., Stobel, M. C., Copeland, N. G. and Jenkins, N. A. (1995) Retroviral sequences located within an intron of the *dilute* gene alter *dilute* expression in a tissue-specific manner. *EMBO J.* 14, 2326–2332.
- 136 Stavenhagen, J. B. and Robins, D. M. (1988) An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene. *Cell* 55, 247–254.
- 137 Maksakova, I. and Mager, D. (2005) Transcriptional regulation of early transposon elements, an active family of mouse long terminal repeat retrotransposons. *J. Virol.* 79, 13865–13874.
- 138 Morgan, H., Sutherland, H., Martin, D. and Whitelaw, E. (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23, 314–318.
- 139 Gaudet, F., Rideout, W., 3rd, Meissner, A., Dausman, J., Leonhardt, H. and Jaenisch, R. (2004) Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing. *Mol. Cell. Biol.* 24, 1640–1648.
- 140 Jähner, D. and Jaenisch, R. (1985) Retrovirus-induced de novo methylation of flanking host sequences correlates with gene inactivity. *Nature* 315, 594–597.
- 141 Whitelaw, E. and Martin, D. (2001) Retrotransposons as epigenetic mediators of phenotypic variation in mammals. *Nat. Genet.* 27, 361–365.
- 142 Golovkina, T. V., Chervonsky, A., Dudley, J. P. and Ross, S. R. (1992) Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. *Cell* 69, 637–645.
- 143 Bhadra, S., Lozano, M. M., Payne, S. M. and Dudley, J. P. (2006) Endogenous MMTV proviruses induce susceptibility to both viral and bacterial pathogens. *PLoS Pathogens* 2, 1134–1143.
- 144 Bénit, L., De Parseval, N., Casella, J., Callebaut, I., Cordonnier, A. and Heidmann, T. (1997) Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the *Fv1* restriction gene. *J. Virol.* 71, 5652–5657.
- 145 Qi, C.-F., Bonhomme, F., Buckler-White, A., Buckler, C., Orth, A., Lander, M. R., Chattopadhyay, S. K. and Morse III, H. C. (1998) Molecular phylogeny of *Fv1*. *Mamm. Genome* 9, 1049–1055.
- 146 Dupressoir, A., Marceau, G., Vernochet, C., Bénit, L., Kanellopoulos, C. and Sapin, V. (2005) Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc. Natl. Acad. Sci. USA* 102, 725–730.
- 147 Gross, L. (1951) 'Spontaneous' leukemia developing in C3H mice following inoculation in infancy, with Ak-leukemic extracts, or AK-embryos. *Proc. Soc. Exp. Biol. Med.* 76, 27–32.
- 148 Aaronson, S. A., Hartley, J. W. and Todaro, G. J. (1969) Mouse leukemia virus: 'spontaneous' release by mouse embryo cells after long-term in vitro cultivation. *Proc. Natl. Acad. Sci. USA* 64, 87–94.
- 149 Harel, L., Harel, J. and Huppert, J. (1967) Partial homology between RNA from Rauscher mouse leukemia virus and cellular DNA. *Biochem. Biophys. Res. Commun.* 28, 44–49.
- 150 Muotri, A. R., Chu, V. T., Marchetto, M. C., Deng, W., Moran, J. V. and Gage, F. H. (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435, 903–910.
- 151 Peaston, A., Knowles, B. and Hutchison, K. (2007) Genome plasticity in the mouse oocyte and early embryo. *Biochem. Soc. Trans.* 35, 618–622.