JVI Journals.ASM.org

Endogenous Gammaretrovirus Acquisition in *Mus musculus* Subspecies Carrying Functional Variants of the XPR1 Virus Receptor

Devinka Bamunusinghe,[a] Qingping Liu,[a] Xiaoyu Lu,[a] Andrew Oler,[b] Christine A. Kozak[a]

Laboratory of Molecular Microbiology,[a] Bioinformatics and Computational Biosciences Branch,[b] National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, USA

The xenotropic and polytropic mouse leukemia viruses (X-MLVs and P-MLVs, respectively) have different host ranges but use the same functionally polymorphic receptor, XPR1, for entry. Endogenous retroviruses (ERVs) of these 2 gammaretrovirus subtypes are largely segregated in different house mouse subspecies, but both MLV types are found in the classical strains of laboratory mice, which are genetic mosaics of 3 wild mouse subspecies. To describe the subspecies origins of laboratory mouse XP-MLV ERVs and their coevolutionary trajectory with their XPR1 receptor, we screened the house mouse subspecies for known and novel *Xpr1* variants and for the individual full-length XP-MLV ERVs found in the sequenced C57BL mouse genome. The 12 X-MLV ERVs predate the origins of laboratory mice; they were all traced to Japanese wild mice and are embedded in the 5% of the laboratory mouse genome derived from the Asian *Mus musculus musculus* and, in one case, in the <1% derived from *M. m. castaneus*. While all 31 P-MLV ERVs map to the 95% of the laboratory mouse genome derived from P-MLV-infected *M. m. domesticus*, no C57BL P-MLV ERVs were found in wild *M. m. domesticus*. All *M. m. domesticus* mice carry the fully permissive XPR1 receptor allele, but all of the various restrictive XPR1 receptors, including the X-MLV-restricting laboratory mouse *Xpr1*^{n} and a novel *M. m. castaneus* allele, originated in X-MLV-infected Asian mice. Thus, P-MLV ERVs show more insertional polymorphism than X-MLVs, and these differences in ERV acquisition and fixation are linked to subspecies-specific and functionally distinct XPR1 receptor variants.

Gammaretroviruses of two distinctive host range groups that use the same XPR1 receptor have been isolated from laboratory mice. The xenotropic and polytropic mouse leukemia viruses (X-MLVs and P-MLVs, respectively) both are able to infect cells of nonrodent species, although P-MLVs but not X-MLVs can infect cells of their laboratory mouse strain hosts (1–3). These host range differences are due to polymorphisms in the viral envelope and in the XPR1 receptor (4).

MLVs can be transmitted in mice as infectious virus (5, 6) or can be inherited as endogenous retroviruses (ERVs) inserted into the host germ line. X-MLV and P-MLV (collectively, XP-MLV) ERVs are present in the common strains of laboratory mice; these strains carry 1 to 20 copies of X-MLVs and 10 to 30 copies of P-MLV ERVs (7, 8). The sequenced C57BL mouse genome has 3 distinct clades of X-MLV ERVs (*Xmv*s) (9), and P-MLV ERVs have been subgrouped as polytropic and modified polytropic murine viruses (*Pmv*s and *Mpmv*s, respectively) (10).

The XP-MLV ERVs are recent acquisitions in the *Mus* genome and are essentially restricted to the 4 house mouse lineages that diverged from the other *Mus* species about 0.5 MYA and that live in close association with humans (11). The ERV subtypes are largely segregated in different house mouse subspecies. The western European *Mus musculus domesticus* carries only P-MLVs, whereas *M. m. castaneus* of southeast Asia, *M. m. molossinus* of Japan, and *M. m. musculus* of eastern Europe and Asia all carry multiple copies of *Xmv*s and few copies of *Mpmv*s or *Pmv*s. The house mouse subspecies also carry XP-MLV ERV variants that appear to be recombinants (12).

XP-MLV infection is mediated by the XPR1 receptor, a transmembrane protein of unknown function that acts as a receptor for P-MLVs in mice and humans and for X-MLVs in humans but not most common strains of laboratory mice (13–15). The virus-infected *M. musculus* subspecies carry polymorphic variants of the XPR1 XP-MLV receptor with different sensitivities to XP-MLV subtypes (16). Four such variants of the *Xpr1* receptor have been found in species of *Mus*, 3 of which are restrictive and are found either in the mice that carry X-MLVs or in a species (*M. pahari*) that is sympatric with house mice that carry these viruses.

The classical inbred strains of laboratory mice have no wild mouse counterparts but were developed at the turn of the last century from fancy mice bred by hobbyists, and these strains have long been recognized as mosaics of several house mouse subspecies (17). The presence of multiple copies of both X- and P-MLV ERVs in the common inbred strains is consistent with this hybrid origin. Although a few of these common inbred strains carry the permissive XPR1 receptor of European *M. m. domesticus*, *Xpr1*^{sxv}, most of the laboratory strains are resistant to infection by X-MLVs, and this is due to a fifth XPR1 receptor variant, *Xpr1*^{n}, that has not been found in any wild mouse (16).

The infectious XP-MLVs represent a virus family that includes variants with atypical host range (18), and this diversity, together with corresponding polymorphisms in their XPR1 receptor, illustrates how genetic conflicts can produce diversifying selection in coevolving hosts and pathogens. To examine the coevolutionary pathways of these interacting entities, we examined the wild mouse origins and fixation of the individual XP-MLV ERVs found

Received 9 May 2013 **Accepted** 26 June 2013

Published ahead of print 3 July 2013

Address correspondence to Christine A. Kozak, ckozak@niaid.nih.gov.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.01264-13.

doi:10.1128/JVI.01264-13

TABLE 1 Inbred strains of laboratory mice grouped by related breeding history

Strain family[a]	Strains
Lathrop/Castle	A/HeJ, AKR/J, BALB/cJ, C3H/HeJ, C57BL/6J, C57BR/cdJ, C57L/J, C58/J, CBA/CaH-T6J, DBA/2J, F/St, I/LnJ, KK/H1J, LG/J, LT/SvEiJ, MA/MyJ, P/J, RF/J, YBR/EiJ, 129/J-GIX+
New Zealand	NZB/B1NJ, NZL/LtJ, NZO/H1LtJ
Swiss	FVB/NJ, NFS/N, NOD/ShiLt, NON/LtJ, NOR/LtJ, SIM.R, SJL/J, SWR/J
Other	BUB/BnJ, CE/J, LP/J, PL/J, RIIIS/J

[a] Collector Abbie Lathrop provided fancy mice to Castle that served as progenitors for the classical inbred strains. NZ mice were derived from outbred mice sent to the University of Otago. Swiss strains were developed from the outbred NIH Swiss colony or from other sources of Swiss mice.

in the sequenced C57BL mouse genome. This was done to establish whether these ERV insertions originated through invasion by infection or by introgression. We screened these same mice for known and novel XPR1 receptor variants. We describe different patterns of P-MLV and X-MLV acquisition and fixation that are consistent with the receptor variants carried by these wild mouse subspecies.

MATERIALS AND METHODS

Mice and DNAs. Sources of mice and DNAs are listed in Table 1, as well as in Table S1 in the supplemental material. DNAs for some strains of laboratory mice were isolated from mice maintained in our laboratory. Mice or DNAs were also purchased from The Jackson Laboratory (Bar Harbor, ME). For *M. musculus* samples, DNAs were isolated from 33 wild-trapped or wild-derived mice in our laboratory or were obtained from other sources for 54 mice. DNAs were isolated from livers of wild-derived mice obtained in 1984 to 1986 from the randomly bred colonies of M. Potter (NCI, Bethesda, MD). Many of these colonies were subsequently transferred to The Jackson Laboratory, where they were inbred, and DNAs were also obtained from some of these strains. Wild-trapped mice from Lake Casitas, CA, were obtained from S. Rasheed (University of Southern California, Los Angeles) in 1986. Additional DNAs from wild-caught or wild-derived mice were obtained from M. Nachmann (University of Arizona, Phoenix), from R. Abe in 1991 (Naval Medical Research Institute, Bethesda, MD), from S. Chattopadhyay and H. Morse III (NIAID, Bethesda, MD), and from RIKEN BioResource Center (Ibaraki, Japan). This mouse DNA panel includes 36 classical inbred strains, 14 DNAs isolated directly from wild-trapped mice, and 73 DNAs from wild-derived laboratory colonies and wild-derived inbred strains of *M. musculus* or *M. spretus*.

Primers and PCR. Primers to identify 3′ and 5′ cell-virus junction fragments and the empty preintegration loci were designed for 19 *Pmv*s, 12 *Xmv*s, and 12 *Mpmv*s (see Table S2 in the supplemental material). Junction fragments were generated using primers from flanking sequences and forward and reverse primers from the MLV long terminal repeat (LTR) sequence: 5′-CAGCTCGCTTCTCGCTTCTG. The *Pmv16* provirus was identified by an additional primer from the virus *env*, XT: 5′-TCAGGACAAGGGTGGTTTGAG. Because of limited DNA for many samples and the large number of samples from each subspecies, different subsets of 20 to 58 wild mouse DNAs were scored for each insertion. While most DNAs produced either the cell-virus junction fragment or the empty locus for each tested ERV insertion (Table 2), there were rare exceptions for which no primer set produced amplicons. These failures were generally subspecies specific, and analysis of the insertion site of one such example (*Mpmv9*) in the genomic sequence of *M. m. castaneus* identified large deletions overlapping the original primer sequences (data not shown).

TABLE 2 Distribution of individual C57BL XP-MLV ERVs in laboratory strains and house mouse (*M. musculus*) subspecies

C57BL ERV	No. positive for each ERV/no. tested[a]					
	Laboratory strains		*M. m. domesticus* (Americas, Eurasia)	*M. m. musculus* (Europe, China)	*M. m. castaneus* (southeast Asia)	*M. m. molossinus* (Japan, Korea)
	Classical	Swiss				
Xmv8	7/27	3/8	0/20	0/12	4/10	13/16
Xmv9	6/28	3/8	0/21	0/10	0/8	2/16
Xmv10	16/25	1/8	0/17	0/9	0/7	5/11
Xmv12	8/18	0/3	0/15	0/10	0/3	5/16
Xmv13	3/24	0/8	0/15	0/10	0/7	6/11
Xmv15	10/24	0/8	0/21	0/11	0/9	1/10
Xmv17	7/24	0/8	0/15	0/12	0/6	9/16
Xmv18	10/25	1/8	0/19	0/11	0/10	11/16
Xmv19	5/25	0/8	0/17	0/6	0/6	6/6
Xmv41	5/26	0/8	0/22	0/10	0/2	13/15
Xmv42	4/24	0/7	0/19	0/11	4/9	1/15
Xmv43	12/30	0/8	0/14	0/9	0/5	4/7
Pmv1	18/20	5/6	1/18	0/9	0/2	0/6
Pmv5	4/20	0/5	0/20	0/6	0/2	0/6
Pmv7	6/19	0/5	0/22	0/8		0/7
Pmv8	21/24	5/6	0/17	0/14	0/5	0/13
Pmv9	3/21	0/6	0/18	0/10	0/3	0/6
Pmv10	20/22	6/6	1/16	0/12	0/4	0/6
Pmv11	11/21	2/4	1/19	0/8	0/3	0/6
Pmv12	2/21	0/6	2/24	0/7	0/4	0/6
Pmv13	20/22	6/7	0/19	0/6	0/4	0/4
Pmv14	13/23	5/5	0/19		0/4	
Pmv15	3/18	1/4	0/18	0/8	0/5	0/5
Pmv16	1/26	0/7	0/25	0/6	0/3	0/5
Pmv18	3/20	2/5	0/16	0/9	0/4	0/3
Pmv19	5/19	1/5	0/15	0/7	0/3	0/2
Pmv20	20/24	6/8	0/18	0/11	0/5	0/9
Pmv21	4/26	0/6	0/20	0/7	0/2	0/6
Pmv22	3/21	0/5	0/16	0/8	0/3	0/7
Pmv23	4/22	5/7	0/19	0/9	0/4	0/7
Pmv24	10/26	1/7	0/20	0/7	0/2	0/7
Mpmv1	8/19	3/3	0/17		0/3	
Mpmv2	3/20	0/5	0/17	0/7	0/5	0/8
Mpmv4	6/23	4/7	0/21	0/7	0/3	1/4
Mpmv5	2/9	0/7	0/14	0/2	0/2	0/2
Mpmv6	9/18	2/8	2/20	0/6	0/3	0/7
Mpmv7	11/26	3/8	0/13	0/6	0/5	0/8
Mpmv8	5/14	1/3	0/18	0/7	0/2	0/3
Mpmv9	12/23	7/8	0/18	0/6	0/4	0/7
Mpmv10	26/26	3/4	0/26	2/5	0/1	0/5
Mpmv11	10/25	0/6	0/15	0/9	0/7	1/7
Mpmv12	9/22	0/6	0/24	0/7	0/5	0/7
Mpmv13	6/21	4/6	0/17	0/8	0/3	0/4

[a] Classical laboratory strains include strains derived from the Lathrop fancy mice, the NZ strains, and other strains not derived from Swiss mice. Blank entries represent DNA sets that failed to produce any PCR product (empty locus or junction fragments). *Xmv*, endogenous xenotropic MLV; *Pmv*, polytropic MLV; *Mpmv*, modified polytropic MLV.

Identification of the subspecies origins of XP-MLV integration sites and XPR1 variants. We used the Mouse Phylogeny Viewer at the University of North Carolina (http://msub.csbio.unc.edu) (19) to examine genomic segments containing each ERV integration site to provide insight into their wild mouse origins. This browser uses a set of diagnostic single-nucleotide polymorphisms (SNPs) to define the local subspecific origin along each autosome and the X chromosome for a set of 100 classical

laboratory strains and 98 wild-derived and wild-caught mice. This browser also provides information on regions of haplotype identity for the inbred strains and the SNP variants that define those regions. For this analysis, we used chromosome coordinates for each ERV and for *Xpr1*, determined for the NCBI37/mm9 reference assembly by BLAT searches (20) of flanking sequences using the UCSC genome browser (http://genome.ucsc.edu/).

Identification and characterization of a novel *Xpr1* variant. Sequence reads (BAM format) were obtained from D. Halligan and P. Keightley (University of Edinburgh, United Kingdom) for 10 wild *M. m. castaneus* animals trapped in 10 different locations in northwest India (see Table S1 in the supplemental material). Sequences were aligned for the region of *Xpr1* exon 13 using SAMtools (21), chromosome coordinates Chr1:157137324-157137461 according to the NCBI37/mm9 reference assembly. Files for each strain were converted to fastq format with bam2fastx (from the TopHat package [22]) and assembled with SOAPdenovo (http://soap.genomics.org.cn) (23) using the following parameters: SOAPdenovo63mer all -K 59. The resulting contigs and the corresponding sequence from GenBank accession number AF131097 were aligned with ClustalW using Jalview (24).

The QuikChange II XL site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used to introduce the *M. m. castaneus Xpr1* mutation F584I into two previously described *Xpr1* clones (25), one representing the full-length permissive *M. dunni Xpr1*^{sxv} and one with a mutation at a receptor critical site, K500E. Primers for mutagenesis were 5′-CCAAATCTCTAT TACTGCTACAACGATCAAGCCTCATGTTGGGG and its reverse complement.

The *Xpr1* clones were expressed in E36 Chinese hamster cells (26), which are poorly infected by X-MLVs and are resistant to other XP-MLVs. Pseudoviruses carrying the *lacZ* reporter gene were generated as previously described; these express the Env glycoproteins of 4 XP-MLVs with different host ranges: CAST-X X-MLV, AKR6 X-MLV, FrMCF P-MLV, and CasE#1 (18). Infection of stable transfected cells was scored by counting blue cells 24 h after staining for β-galactosidase activity (18). *P* values were determined from a 2-tailed Student's *t* test using GraphPad Prism 6: *, $P \leq 0.05$; **, $P \leq 0.02$; ***, $P \leq 0.0001$; ****, $P \leq 0.0001$.

Expression of *Xpr1* in the transfected cells was confirmed by Western analysis. Proteins were extracted with M-PER mammalian protein extraction reagent (Pierce, Rockford, IL). The expression vector inserts a V5 epitope at the C terminus of XPR1, and expression was detected using anti-V5 antibody (Invitrogen, Carlsbad, CA) followed by goat anti-mouse IgG conjugated with horseradish peroxidase (Invitrogen). The membrane was stripped and incubated with mouse anti-α-tubulin (Sigma, St. Louis, Mo) and the goat anti-mouse IgG.

Sequence analysis of wild mouse *Xpr1* exon 13. *Xpr1* exon 13 was PCR amplified from Asian mouse DNAs using primers Ex13F and Ex13R (16), cloned into TOPOI, and sequenced.

RESULTS

Identification of specific XP-MLV proviruses in laboratory and wild mice. Individual XP-MLV ERVs carried by the inbred strains of laboratory mice were previously identified using type-specific hybridization probes for 3 ERV subtypes: *Xmv*, *Pmv*, and *Mpmv* (10). Analysis of 7 inbred strains (C57BL, C57L, BALB/c, AKR, A, C3H, and DBA/2) and recombinant inbred strains derived from these 7 progenitor strains determined the strain distribution of each identifiable ERV and mapped their chromosome locations (8, 27, 28). We used information on the chromosome locations and envelope subtype of these XP-MLV ERVs (9) to extract the full-length sequences of 19 *Pmv*s, 12 *Mpmv*s, and 12 *Xmv*s, along with flanking cellular sequences from the C57BL genome (see Table S2 in the supplemental material).

Specific primers were designed to identify each ERV to describe their distribution in laboratory strains and house mouse subspecies (see Table S2 in the supplemental material). Cell-virus junction fragments diagnostic for each provirus were amplified by PCR using a flanking sequence primer and a primer from within the viral LTR. Empty (preintegration) sites were identified with primers flanking opposite ends of the provirus, as illustrated for *Mpmv11* in Fig. 1A.

The primer sets for all 43 ERVs produced the expected products and strain distribution patterns in the 7 test strains. A broader inbred strain distribution was then determined for up to 29 additional common inbred strains, a set that included 8 Swiss-derived strains and 21 classical inbred strains largely derived from the fancy mouse colonies obtained from Abbie Lathrop by William Castle (29) (Table 1). We also typed DNAs selected from a panel of 85 DNAs isolated from house mice from the eastern and western United States, South America, Eurasia, and north Africa (see Table S1 in the supplemental material).

***Xmv* orthologs are traceable to Japanese house mice.** PCR primers for 10 of 12 *Xmv*s identified either the cell-virus junction products or the empty locus. The empty preintegration locus was not identified for *Xmv8* and *Xmv9*, two linked ERVs located on chromosome 4 at Chr4:145.68 and Chr4:146.6, respectively. These *Xmv*s are inserted in a region of small local duplications; BLAT screens identified additional copies of the insertion sites for *Xmv8* (for example, Chr4:145.26 and Chr4:146.60) and for *Xmv9* (Chr4:145.45 and Chr4:145.7). Because of these additional copies, the sequences surrounding these *Xmv*s could not be used to design primers to detect the empty locus.

All 12 *Xmv*s were identified in multiple Lathrop-derived inbred strains (Table 2). The eight Swiss-derived strains generally lacked these 12 *Xmv*s, the notable exceptions being NOR, SIM.R, and FVB. These strains carry the linked *Xmv8* and *Xmv9* ERVs. These insertions map near the *Fv1* retrovirus restriction gene (Chr4:147.2) (30), and all 3 strains carry the *Fv1*^{b} allele at this locus, as does C57BL and other strains carrying *Xmv8* and *Xmv9* (31–33), suggesting a common origin for *Fv1*^{b} and these 2 *Xmv*s.

Not surprisingly, none of the C57BL *Xmv*s were identified in wild-derived or wild-caught *M. m. domesticus*, mice that lack X-MLV *env* genes according to southern analysis (Table 2) (11). Similarly, none of these insertions were identified in European *M. m. musculus*, despite the presence of multiple X-MLV *env* genes in these mice (11). All 12 *Xmv*s were, however, found in *M. m. molossinus*, and most were identified in multiple *molossinus* DNAs. *M. m. molossinus* is a natural hybrid of *M. m. castaneus* and *M. m. musculus* (34), and two *Xmv*s, *Xmv42* and *Xmv8*, were also found in *M. m. castaneus*, suggesting these ERVs predate the *molossinus* hybridization. The acquisition of these *Xmv*s by laboratory strains likely results from the inclusion of Japanese mice, like the Japanese waltzer mouse, in the fancy mouse colonies used to develop the common inbred strains (29), and the failure to detect these ERVs in European *M. m. musculus* confirms the previously observed distinction between Asian and European *M. m. musculus* (17).

C57BL *Pmv*s and *Mpmv*s are not found in house mouse subspecies. PCR analysis of 36 inbred mouse strains indicates that most of the 31 P-MLV ERVs (collectively termed *M/pmv*s) are carried by fewer than half of the inbred strains tested (Table 2), although several, like *Mpmv10*, *Pmv8*, *Pmv10*, and *Pmv13*, are carried by nearly all strains, suggesting that their acquisition predates laboratory strain origins.

Previous studies of *Mus* taxa have shown that P-MLV ERVs are largely restricted to the *M. m. domesticus* lineage of house mice (11), a subspecies which is globally distributed as a result of passive

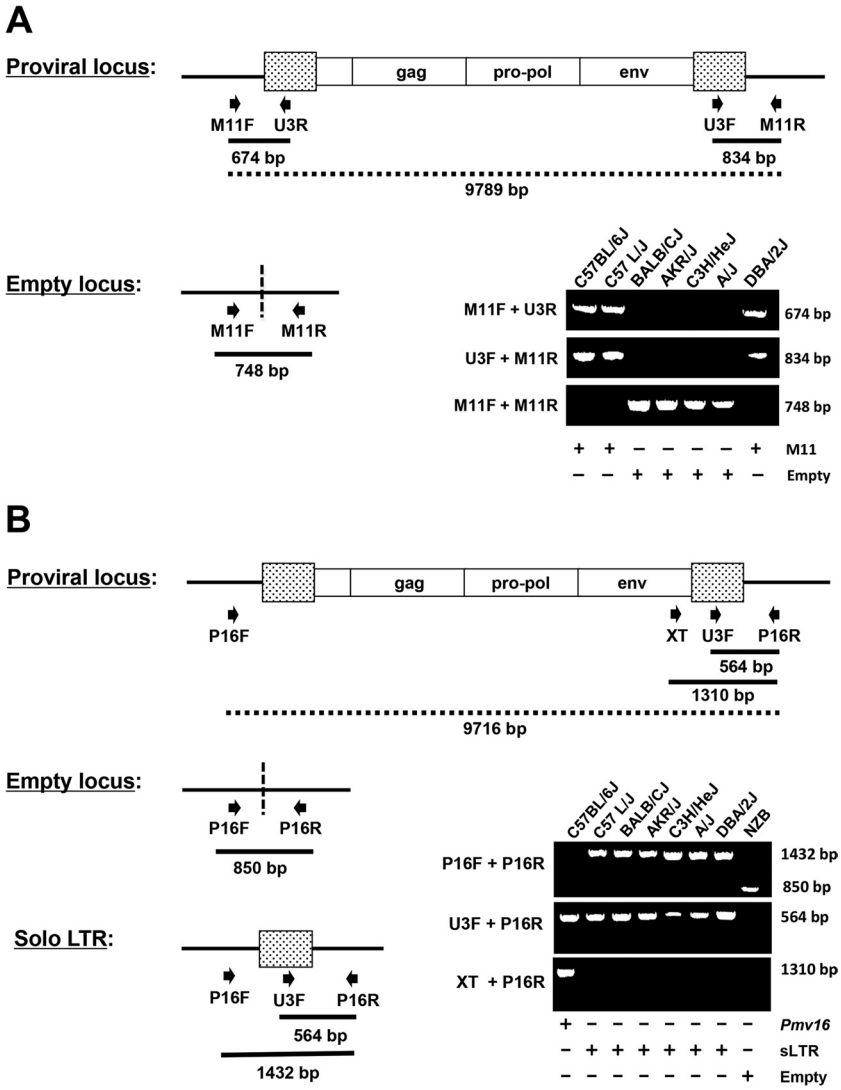

FIG 1 Detection of a representative intact provirus, *Mpmv11*, and a deleted provirus, *Pmv16*, in common inbred strains. (A) At the top is the *Mpmv11* provirus structure with primers designed from the LTR and from flanking sequences, followed by a diagram of the empty preintegration locus. At the bottom right are PCR tests of 7 inbred strains that had been typed for this ERV by Southern blotting (8). (B) The *Pmv16* genome, the empty locus, and the solo LTR are diagrammed with PCR primers and expected products, and PCR test results are provided for 8 strains.

transport by western Europeans. *M/pmv*s are also found in *M. spretus*, a species outside the house mouse lineage, but one that is sympatric with and interfertile with *M. m. domesticus*, suggesting its few *M/pmv*s have been acquired by interbreeding (35). Multiple P-MLV ERV *env* genes are also found in *M. m. musculus* from eastern Europe, likely reflecting the predominantly west-to-east introgression across the hybrid zone that separates the ranges of *M. m. domesticus* and *M. m. musculus* (36).

Not surprisingly, the *Pmv*s or *Mpmv*s were rarely identified by PCR in Asian subspecies which lack P-MLV *env* genes by Southern hybridization (Table 2) (11), but it was surprising that none of the 31 *M/pmv*s, including those present in nearly all laboratory strains, were reliably detected in the *M. m. domesticus* DNAs or in *M. spretus*, even though we sampled multiple geographically separated populations from 5 continents. This suggests a high level of insertional polymorphism and indicates that none of these ERVs were acquired at or soon after subspeciation.

Solo LTRs: evidence of provirus loss by homologous recombination. The mouse genome is replete with solo LTRs, resulting from the excision of the bulk of a retrovirus by homologous recombination, although for the recently acquired MLV ERV family, solo LTRs represent less than 20% of insertions (37). Solo LTRs at the 43 XP-MLV insertion sites should be identifiable by the presence of cell-LTR junction fragments together with an empty-locus PCR product of increased size (Fig. 1B). Within our set of 43 ERVs, we identified only one such deletion, of *Pmv16*. The larger-than-expected "empty" site for *Pmv16* contained a solo LTR with the expected U3-R-U5 configuration and the 190-bp U3 insertion typical of P-MLV ERVs (data not shown). The full-length *Pmv16* was found only in C57BL, and the solo LTR was found in half of the laboratory strains and none of the wild mice.

Subspecific origins of genome segments carrying XP-MLV ERVs. High-resolution genomic maps developed for 100 laboratory strains and 98 wild-derived and wild-trapped species (17)

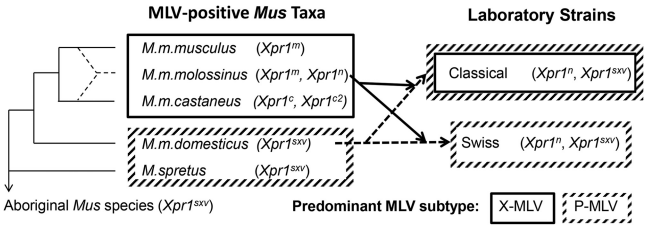

FIG 2 Distribution of XP-MLV ERVs and *Xpr1* receptor alleles in *Mus* taxa and in *M. musculus*-derived laboratory strains. A phylogenetic tree shows the house mouse (*M. musculus*) radiation. The inbred strain genomes are largely *M. m. domesticus* derived, but some of these strains acquired the restrictive $Xpr1^{n}$ receptor and X-MLV ERVs from Japanese fancy mice (*M. m. molossinus*).

have been used to infer the subspecific origins of DNA segments along each of the 19 mouse autosomes and the X Chr, and they showed that the genomes of the various inbred strains are 90 to 98% *M. m. domesticus*, 2 to 9% *M. m. musculus*, and <1% *M. m. castaneus*. This analysis also confirmed that Japanese *M. m. molossinus* is a natural hybrid of *M. m. musculus* and *M. m. castaneus* and determined that the *M. m. musculus* segments of the laboratory mouse genome derive from Japanese mouse (*M. m. molossinus*) progenitors (17) (Fig. 2). The smaller Asian mouse-derived segments are scattered over the laboratory mouse genome, and the pattern of this distribution differs from strain to strain. We used this resource to examine the insertion sites for each of the 43 C57BL ERVs.

Figure 3 represents 7- or 17-Mb segments surrounding ERV insertion sites for a few representative *M/pmv*s and for all 12 *Xmv*s for the 7 inbred strains initially typed for these ERVs. Not surprisingly, all of the 31 *Pmv*s and *Mpmv*s are found in segments of the genome identifiable as *M. m. domesticus*. Haplotype analysis of these 31 insertion sites further shows that the insertions are consistently present in segments of shared haplotype with C57BL (Fig. 4A shows representative data). This indicates that none of the 31 *M/pmv*s represent novel insertions acquired after the development of the inbred strains.

In contrast, 10 of the 12 *Xmv*s examined were each contained in genome segments clearly derived from *M. m. musculus* (Fig. 3). The 11th *Xmv*, *Xmv42*, had been identified in *M. m. castaneus* as well as *M. m. molossinus* DNAs by PCR, and this *Xmv* mapped to an *M. m. castaneus*-derived segment (Fig. 3).

The 12th *Xmv*, *Xmv15*, was not found in a recognized *M. m. musculus* segment (Fig. 3). This could be due to the failure to identify a small *M. m. musculus*-derived segment in this region of Chr 9, or it could have resulted from generation of a novel insertion in the fancy mouse progenitors of the laboratory mouse strains. The former possibility is suggested by the fact that *Xmv15* lies in a short (185-kb) region of SNPs shared with *M. m. molossinus* and *M. m. musculus* (Table 3).

Complications from pervasive intersubspecific introgression in wild-derived strains. Our PCR analysis identified several *M/pmv*s, as well as *Xmv*s, in rare *M. m. domesticus* samples. The exceptions involved wild-derived laboratory strains but no wild-trapped mice. This is an important distinction, because there is evidence of intersubspecific introgression in the wild-derived

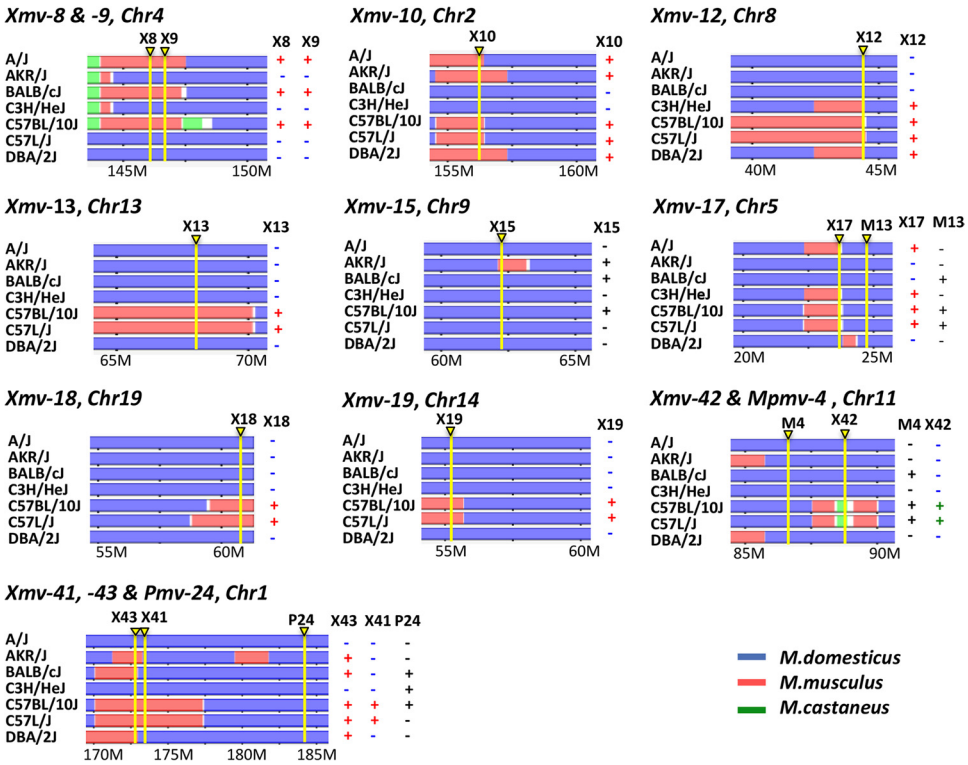

FIG 3 Subspecific origins of C57BL XP-MLV ERVs. Horizontal tracks for each of 7 inbred strains represent 7- or 17-Mb genomic segments surrounding the ERV insertion sites from the Mouse Phylogeny Viewer (19). ERV sites are indicated by yellow lines, and the 3 track colors represent blocks that originated from the 3 subspecies that contributed to the laboratory mouse genome. To the right of each panel are PCR typing data for each ERV, color coded for all *Xmv*s except *Xmv15*, to match track colors at the ERV site.

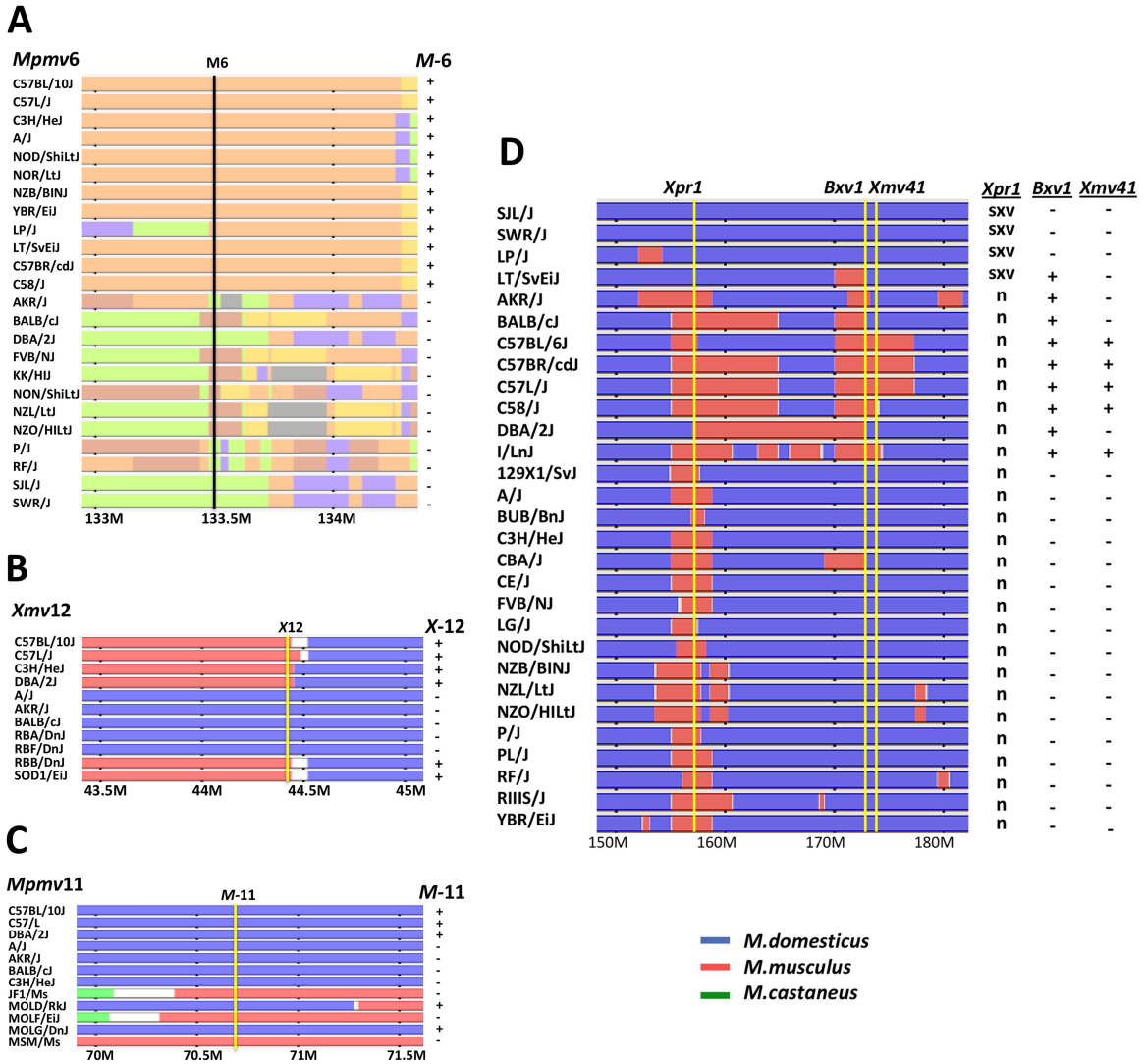

FIG 4 Subspecific origins of *Xpr1* alleles and XP-MLV ERVs and haplotype diversity at insertion sites. (A) Regions of haplotype similarity at a representative P-MLV ERV, *Mpmv6*. The black line marks the Chr 1 insertion site, and the different track colors represent different shared haplotypes. The *Mpmv6* insertion lies in a region of haplotype similarity (tan color) shared by the 12 positive strains as determined by Mouse Phylogeny Viewer (19). (B and C) The unexpected detection of *Xmv12* in *M. m. domesticus* strains RBB and SOD1 and *Mpmv6* in *M. m. molossinus* strains MOLD and MOLG is because their insertion sites in these strains are *M. m. musculus* and *M. m. domesticus* derived, respectively, due to genomic contamination. (D) Subspecific origins of the two laboratory mouse *Xpr1* alleles and two linked X-MLVs. In panels B to D, ERV sites are indicated by yellow lines, and track colors represent 3 *Mus musculus* subspecies. To the right are PCR typing data for each marker.

TABLE 3 Region of shared ancestry defined by SNPs around *Xmv15*

M. musculus subspecies	Mouse	SNP comparisons of 9 *Xmv15*$^+$ strains to representatives of 4 *M. musculus* subspecies[a]					
		Mismatched SNPs		Shared SNPs		Uniquely shared SNPs	
		No.	%	No.	%	No.	%
molossinus	JF1	2	5.0	38	95.0	19	90.5
musculus	BAG94	5	12.5	35	87.5	16	76.2
castaneus	IN13	12	33.3	24	66.7	9	47.4
domesticus	MWN1026	21	52.5	19	47.5	0	0.0

[a] Comparisons for the 40 SNPs in a region including *Xmv15* (Chr9:62112348-62297739); SNPs were from http://msub.csbio.unc.edu (19). The 9 *Xmv15*$^+$ inbred strains (C57BL, NZB, NZL, RF, LT, C57BR, C58, and AKR) share all 40 SNPs, and these were compared to representative DNAs of the 4 subspecies. The unique shared SNP subset excludes the 19 SNPs for which the 9 inbreds and 4 wild mice are identical.

strains that is suggestive of undocumented cross-breeding with common laboratory strains (17). In particular, multiple (4–12) C57BL ERVs were found in the *M. m. domesticus* wild-derived strains SOD1, RBB, RBA, and RBF, and all 4 strains contain substantial genomic regions that are nearly identical to classical inbred strains (17). The donor of these introgressed regions has been identified as C57BL for SOD1 and RBA, providing an explanation for the presence of several C57BL ERVs that were not found in any other *M. m. domesticus* mice. That these ERVs resulted from contamination is supported by examination of genomic insertion sites for the *Xmv*s that were detected in these mice by PCR. In all cases (Fig. 4B shows representative data), the *M. m. domesticus* wild-derived mice carrying *Xmv*s are *M. m. musculus* at these sites. Therefore, detection of these ERVs in clearly contaminated strains is not evidence of their wild mouse origins, so typing data from

these 4 strains was deleted from Table 2. Similarly, insertion sites for several *M/pmv* ERVs unexpectedly found in Asian wild-derived strains were found within contaminating *M. m. domesticus* segments (Fig. 4C). This also raises concerns about the few P-MLV ERVs detected in rare wild mouse DNAs (Table 2); all of these cases were wild-derived strains either known to be contaminated (PERA/C, *Pmv10* and *Mpmv10*; SF, *Mpmv6*; CALB, *Mpmv6* and *Pmv12*), known to be derived from contaminated lineages (MOLC, *Mpmv11* and *Mpmv4*), or not examined for subspecific origins and for possible contamination (BQC, *Pmv12*; VEJ and Viborg, *Mpmv10*; JJD, *Pmv1*; HAF, *Pmv11*). Thus, the rare positives for these ERVs are not convincing evidence of their wild mouse origins.

Wild mouse origin of *Xpr1^n*. Two of the 5 known alleles of the *Mus* XP-MLV *Xpr1* receptor are found in the classical and Swiss strains of laboratory mice (Fig. 2). The more common allele, *Xpr1^n*, supports entry of P-MLVs but not X-MLVs. The fully permissive wild mouse allele, *Xpr1^{sxv}*, is found in a few of the classical laboratory strains (33). In previous screenings of wild house mice, we identified *Xpr1^{sxv}* in all tested *M. m. domesticus* animals from Eurasia and the Americas as well as in aboriginal *Mus* species, but we did not find a single wild mouse carrying *Xpr1^n* (16).

To further investigate the wild mouse origins of the two laboratory mouse *Xpr1* alleles, we examined the subspecies origins of the chromosome segment carrying the *Xpr1* site on Chr 1 in the 29 strains typed for *Xpr1* (Fig. 4D). The 4 strains with the permissive *Xpr1^{sxv}* allele are *M. m. domesticus* at the Chr 1 *Xpr1* locus, consistent with the observation that *M. m. domesticus* carries this receptor variant. The site of the *Xpr1* locus, however, is *M. m. musculus* derived in the mice carrying *Xpr1^n*. This indicates that the restrictive *Xpr1^n* allele is derived from X-MLV-infected *M. m. musculus*. This same region of Chr 1 carries 2 *Xmv*s (*Bxv1*/*Xmv43* and *Xmv41*), and analysis of this larger strain set further illustrates the earlier point that the *Xmv*s map to *M. m. musculus*-derived genomic regions (Fig. 4D).

A novel XPR1 variant in *M. m. castaneus*. The house mouse lineages originated and diverged from an ancestral population on the Indian subcontinent, which carries many of the alleles found in peripheral Eurasian populations, but it is more genetically variable (38, 39). It is likely that the various XP-MLVs were acquired by this ancestral population along with the various *Mus Xpr1* receptors; founders of the *M. musculus* subspecies with virus and receptor variants then moved to their current ranges. To determine if this ancestral population carries additional *Xpr1* variants not found in the peripheral populations, we screened genomic sequences from wild-caught *M. m. castaneus* mice from northern India for *Xpr1* sequence variation.

The four *M. musculus Xpr1* alleles have receptor-critical replacement or deletion mutants in the putative fourth extracellular loop (ECL4) encoded by exon 13 and a substitution in ECL3 (exon 11) at position 500 (16). Alignment of the contigs from 10 sequenced *M. m. castaneus* mice to the ancestral *Xpr1* exon 13 sequence (AF131097) demonstrates that 7 mice are homozygous for the deletion of ECL4 amino acid residues TTFKP, a deletion that is diagnostic of the *Xpr1^c* receptor identified in *M. m. castaneus* animals trapped in Thailand (Fig. 5A). Only the deletion allele was detected in strain H24, so it is also likely to be homozygous for the deletion; however, as there is low sequence coverage for this strain in this region, it is difficult to rule out the possibility of heterozygosity. In all cases, this deletion is linked to a nearby C→T polymorphism resulting in a Thr→Ile replacement compared to the reference genome; this change lies in the 7th putative transmembrane domain. Mice H12 and H14 lack the *Xpr1^c* TTFKP deletion and are similar to the ancestral type, although they also have a T→A substitution within this segment, resulting in the altered protein sequence TTIKP. Strain H15 is heterozygous for these 2 alleles. In exon 11 (ECL3), all 10 mice carry the permissive residue K500 at this receptor-critical position.

In order to assess the effect of this novel F584I substitution on receptor function, we introduced it into two expression clones of the XPR1 receptor. One clone contained the fully permissive *Xpr1^{sxv}*, and the second carried the replacement mutation, K500E. *Mus* X-MLV receptor function requires either K500 or T582; both sites have to be mutated to restrict X-MLV entry, and both sites are mutated in *Xpr1^n* mice, which are X-MLV resistant (40). Thus, we introduced K500E into a second construct together with F584I to determine if the F584I mutation, like T582Δ of *Xpr1^n*, is unable to support infection in the presence of K500E.

The *Xpr1* constructs were expressed in XP-MLV-resistant E36 hamster cells and infected with 4 pseudoviruses with Env glycoproteins having different XPR1 receptor sensitivities. *Xpr1^{sxv}*-F584I showed the permissive parental phenotype, but in the presence of K500E, pseudovirus infection was significantly reduced for 2 of the 4 viruses, AKR6 X-MLV and CasE#1 ($P = 0.0011$ and $P < 0.0001$, respectively) (Fig. 5B). This indicates that residues at 584 contribute to the virus binding site, and that F584I can reduce receptor function in some receptor sequence contexts. We term this second *M. m. castaneus* receptor variant *Xpr1^{c2}*.

***Xpr1* variation in Asian *M. musculus*.** Three of the restrictive XPR1 genes identified in *Mus* have different deletions in the receptor-determining ECL4 of this membrane protein. In our previous screening of wild mouse DNAs (16), we typed most mice using a PCR-based assay designed to identify these 3 diagnostic deletions. Because our PCR screening may have missed replacement mutations like *Xpr1^{c2}*, we sequenced this receptor-determining segment of 31 DNAs from the 3 X-MLV-positive subspecies. No additional novel variants were identified, and virtually all sequences were true to subspecies type and/or their geographical distribution (Table 4). These sequenced DNAs included 12 Japanese mice classed as *M. m. molossinus*, 5 of which were determined to carry *Xpr1^n*; this confirms our earlier conclusion that *Xpr1^n* arose in Japanese mice.

DISCUSSION

The mammalian genome is populated with various families of ERVs that can alter gene function and genome structure and that are capable of forming stable associations with their hosts over evolutionary time. Understanding how these elements originate, spread, and diversify is fundamental to our understanding of evolution and the virus-host relationship. The present study was done to provide insights into the origins of the XP-MLVs that have become fixed in laboratory mice and to describe their subspecies and strain distribution in conjunction with the evolution of functional variants of their receptor.

Novel ERV families originate as the result of exogenous integration events. Active ERVs can produce additional ERVs, and the expansion of most ERVs is thought to occur soon after their introduction into the germ line as the result of recurrent reinfection and reinsertion of active elements in hosts that encode functional receptors. Host populations with active ERVs are characterized by

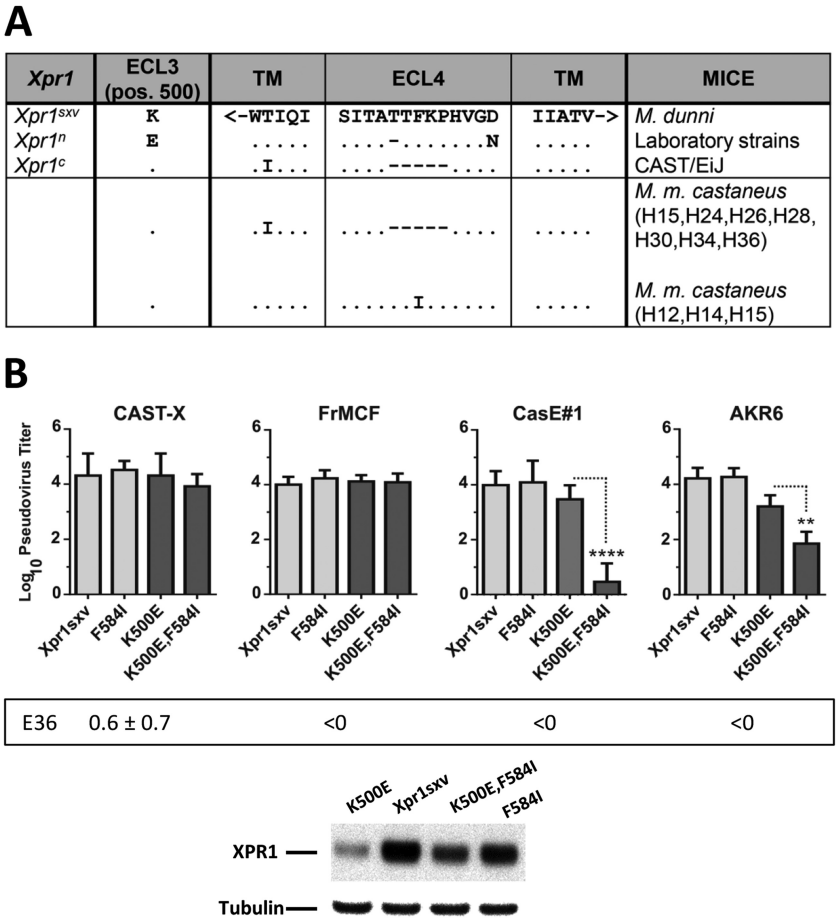

FIG 5 Identification of a novel *Xpr1* variant in wild-caught *M. m. castaneus*. (A) Sequence variation in the receptor determining ECLs of the *Xpr1* of 10 individual *M. m. castaneus* animals. At the top are sequences for the permissive allele, $Xpr1^{sxv}$, the common laboratory mouse allele, $Xpr1^{n}$, and the $Xpr1^{c}$ variant identified in inbred CAST/EiJ animals. Below are the two alleles identified in sequences of 10 wild-caught *M. m. castaneus* animals, $Xpr1^{c}$ and $Xpr1^{c2}$. (B) Infectivity of XP-MLV *lacZ* pseudoviruses in E36 hamster cells expressing XPR1 variants. Light bars represent $Xpr1^{sxv}$ with and without mutation F584I; the dark bars include the K500E mutation, which disables the ECL3 receptor determinant. Cells were infected with pseudoviruses expressing Env proteins of 4 XP-MLV host range variants. Infected cells were stained to detect the reporter gene 24 h after infection. Titers represent blue cells in 50 μl of virus stock ± standard errors of the means, and significant *P* values were determined from 3 independent tests. E36 cells are inefficiently infected by CAST-X. The Western blot shows expression of V5-tagged XPR1s and tubulin in transfected E36 cells.

TABLE 4 Distribution of *Xpr1* receptor alleles in subspecies of Asian house mice

Xpr1 allele and subspecies	Strain/line	Receptor-restricted MLVs
$Xpr1^{m}$		
M. m. castaneus	HMI	P-MLV, CasE#1, Cz524
M. m. molossinus	KOR5, MSM, MOLF, MOLD, MOLC, MOLG, MOM	
M. m. musculus	MBT, AKT, AST, NJL	
M. m. spp.	KJR, IAS-2, Shh1, BJN3, Chd	
$Xpr1^{c}$		
M. m. castaneus	CAST/EiJ, CAST/Ncr, CAST/Rp, JF1, MYS, BGR1, *bactrianus*	P-MLV, CasE#1
$Xpr1^{n}$		
M. m. molossinus	STM1, STM2, KOR7, AIZ, MAE	X-MLV, CasE#1, Cz524

some degree of insertional polymorphism. With time, ERVs accumulate mutations, and they become less active and less capable of generating new insertions.

Two distinct patterns of ERV acquisition and fixation were observed for X-MLVs and P-MLVs, and these patterns are generally consistent with their different tropisms and with the XPR1 receptors carried by their hosts (Fig. 2). As shown here, all of the C57BL X-MLV ERVs are found in Japanese wild mice, as previously demonstrated for one of them, *Bxv1/Xmv43* (33). This suggests limited ongoing retrotransposition, which is not surprising, because although many X-MLVs are active (4, 41), there is an entry block imposed by the laboratory mouse $Xpr1^{n}$ receptor, a receptor that we show here to have originated in Japanese wild mice, just as the *M. m. musculus*-derived genomic segments in which these ERVs are embedded have been determined to be Japanese (17).

In contrast, we failed to convincingly trace any of the P-MLV ERVs to wild mouse progenitors. This is surprising for 2 reasons. First, these ERVs are present in high copy numbers in all *M. m.*

domesticus mice (11), and we sampled *M. m. domesticus* mice trapped at multiple sites on 5 continents. Second, all 31 insertion sites were in regions of shared haplotypes, indicating that insertion predates the origins of laboratory mice. The unexpected failure to identify any of the 31 C57BL *M/pmv*s in *M. m. domesticus* suggests that these P-MLV ERVs either were fixed in *M. m. domesticus* subpopulations not sampled for this analysis or were acquired during the centuries-long development of the fancy mouse colonies that served as progenitors of the laboratory strains.

Insertional polymorphism is characteristic of recently acquired and active families of ERVs carried by hosts with permissive receptors, and P-MLV-positive wild and inbred mice carry the P-MLV permissive receptors, $Xpr1^{sxv}$ and $Xpr1^{n}$. However, this correlation is complicated by the fact that no infectious viruses corresponding to *Pmv*s or *Mpmv*s have been isolated; all infectious MLVs with P-MLV host range are recombinants between these P-MLV ERVs and infectious viruses of other host range subgroups (42, 43). The reason for the failure of P-MLV ERVs to produce infectious virus is not known. Although many of these ERVs have open reading frames for all coding regions, their LTRs carry a negative regulatory element and a 190-bp insertion of unknown functional significance that has not been found in any infectious MLV (44, 45). Thus, although functional defects in the *M/pmv* LTRs may be responsible for the absence of infectious *Pmv*- and *Mpmv*-like viruses, at least some of these LTRs have promoter activity (46). The failure to produce infectious virus suggests that the mechanism for P-MLV ERV retrotransposition is independent of reinfection, and in fact, P-MLV ERVs can amplify in the presence of ecotropic MLV (E-MLV) infection, because P-MLV genomes are preferentially packaged in E-MLV particles that then use the mCAT-1 receptor (47, 48). Such retrotransposition in the absence of infection would be expected to amplify ERVs with fatal defects, but our panel of 31 *M/pmv*s contains no obvious examples of this, and the E-MLVs needed to facilitate such retrotransposition have not been found in *M. m. domesticus* except in California, where *M. m. domesticus* has recently interbred with E-MLV-positive *M. m. castaneus* introduced from Asia (11). Further work on P-MLV ERV activity in inbred and wild mice is needed to resolve this issue.

MLVs have not been detected in species other than mice. This is surprising given the global distribution of MLV-infected mice, the ready transmissibility of MLVs in mice housed together (6), and the fact that the ubiquitously expressed XPR1 acts as an XP-MLV receptor in most mammals (16). The failure to find MLV ERVs in other mammals could reflect the fact that retroviral sequences in most sequenced genomes are incompletely annotated and the fact that few species in direct contact with Asian virus-producing mice have been sequenced. Also, virus transmission and the likelihood of endogenization can be influenced by multiple host factors. The factors that govern trans-species transmission are poorly understood, although some host restriction factors, like APOBEC3, are effective against XP-MLVs (49). Interestingly, while APOBEC3 is blocked by the MLV glycosylated Gag (50), this antagonist is not encoded by P-MLV ERVs, which, unlike *Xmv*s, show evidence of APOBEC3 editing (9, 51). Thus, the species specificity of XP-MLV ERVs may be influenced by factors in addition to receptor polymorphism.

Ever since the discovery of ERVs and the more recent recognition that ERVs and other transposable elements (TEs) make up nearly half of the mammalian genome, there has been speculation that they play critical but largely unrecognized roles in their hosts. While some specific ERVs have been recruited to serve specific host functions (e.g., *Fv1* and syncytins), the vast majority of individual ERVs and TEs have no such defined roles, although these elements clearly can participate in genomic remodeling and are important players in host gene regulatory networks, especially in early development and during stress responses (52, 53). The XP-MLV ERVs described here have already survived purging by natural selection but are not fixed. The insertional polymorphism that produces variation in the ERV content of mouse strains and species could influence host processes governed by individual ERVs or by coordinately regulated ERV subtypes. ERV insertions are estimated to cause 10% of spontaneous mutations in mice (54), and ERV polymorphisms have been associated with changes in gene expression, especially in genes with differential expression across strains (37, 55), suggesting that insertionally polymorphic ERVs like the XP-MLVs contribute to these strain-specific phenotypic differences.

Like the P-MLV and X-MLV ERVs, functional variants of the XPR1 receptor show a subspecies-specific distribution. The 3 major house mouse lineages generally carry distinctive and subspecies-specific *Xpr1* variants ($Xpr1^{sxv}$ in *M. m. domesticus*, $Xpr1^{m}$ in *M. m. molossinus* and *M. m. musculus*, and $Xpr1^{c}$ and $Xpr1^{c2}$ in *M. m. castaneus*), indicating a distribution that roughly coincides with the house mouse radiation and with the acquisition of the different MLV ERVs (Fig. 2). Here, we established that the restrictive allele of laboratory mice, $Xpr1^{n}$, originated in the Asian mice that also supplied the laboratory mouse *Xmv*s, and we identified a novel receptor variant in *M. m. castaneus* from northern India, bringing to five the number of restrictive alleles found in Asian mice. Because the house mouse subspecies are derived from an ancestral population in India which is more genetically variable than the peripheral populations, further characterization of these mice is likely to identify novel virus subtypes as well as receptor variants.

When retroviruses infect new hosts, coevolutionary pressures produce an "arms race" that results in adaptive mutations in pathogen and host, eventually allowing for some degree of peaceful coexistence. The restrictive alleles of *Xpr1* would be expected to interfere with XP-MLV spread and any deleterious consequences of infection, like disease induction. The presence of all 5 restrictive alleles in mice carrying X-MLVs or in mice sympatric with X-MLV-infected house mice (*M. pahari*) suggests that X-MLV infection has driven fixation of these restrictive receptors. On the other hand, for *M. m. domesticus*, the persistence of the permissive XPR1 receptor in these P-MLV-positive mice could have contributed to the observed insertional polymorphism but may also reflect the fact that horizontal spread of P-MLVs is not necessarily XPR1 receptor dependent, rendering XPR1 variation irrelevant for defense in this subspecies.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, Bethesda, MD.

We thank Alicia Buckler-White and Ronald Pliska for sequencing.

REFERENCES

1. **Levy JA.** 1973. Xenotropic viruses: murine leukemia viruses associated with NIH Swiss, NZB, and other mouse strains. Science **182:**1151–1153.
2. **Hartley JW, Wolford NK, Old LJ, Rowe WP.** 1977. New class of murine

leukemia-virus associated with development of spontaneous lymphomas. Proc. Natl. Acad. Sci. U. S. A. **74:**789–792.
3. **Fischinger PJ, Nomura S, Bolognesi DP.** 1975. A novel murine oncornavirus with dual eco- and xenotropic properties. Proc. Natl. Acad. Sci. U. S. A. **72:**5150–5155.
4. **Kozak CA.** 2010. The mouse "xenotropic" gammaretroviruses and their XPR1 receptor. Retrovirology **7:**101.
5. **Gardner MB, Chiri A, Dougherty MF, Casagrande J, Estes JD.** 1979. Congenital transmission of murine leukemia virus from wild mice prone to the development of lymphoma and paralysis. J. Natl. Cancer Inst. **62:** 63–70.
6. **Portis JL, McAtee FJ, Hayes SF.** 1987. Horizontal transmission of murine retroviruses. J. Virol. **61:**1037–1044.
7. **O'Neill RR, Khan AS, Hoggan MD, Hartley JW, Martin MA, Repaske R.** 1986. Specific hybridization probes demonstrate fewer xenotropic than mink cell focus-forming murine leukemia virus *env*-related sequences in DNAs from inbred laboratory mice. J. Virol. **58:**359–366.
8. **Frankel WN, Stoye JP, Taylor BA, Coffin JM.** 1990. A linkage map of endogenous murine leukemia proviruses. Genetics **124:**221–236.
9. **Jern P, Stoye JP, Coffin JM.** 2007. Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. PLoS Genet. **3:**e183. doi:10.1371/journal.pgen.0030183.
10. **Stoye JP, Coffin JM.** 1987. The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. J. Virol. **61:**2659–2669.
11. **Kozak CA, O'Neill RR.** 1987. Diverse wild mouse origins of xenotropic, mink cell focus-forming, and two types of ecotropic proviral genes. J. Virol. **61:**3082–3088.
12. **Tomonaga K, Coffin JM.** 1998. Structure and distribution of endogenous nonecotropic murine leukemia viruses in wild mice. J. Virol. **72:**8289–8300.
13. **Tailor CS, Nouri A, Lee CG, Kozak C, Kabat D.** 1999. Cloning and characterization of a cell surface receptor for xenotropic and polytropic murine leukemia viruses. Proc. Natl. Acad. Sci. U. S. A. **96:**927–932.
14. **Battini J-L, Rasko JEJ, Miller AD.** 1999. A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction. Proc. Natl. Acad. Sci. U. S. A. **96:** 1385–1390.
15. **Yang Y-L, Guo L, Xu S, Holland CA, Kitamura T, Hunter K, Cunningham JM.** 1999. Receptors for polytropic and xenotropic mouse leukaemia viruses encoded by a single gene at *Rmc1*. Nat. Genet. **21:**216–219.
16. **Yan Y, Liu Q, Wollenberg K, Martin C, Buckler-White A, Kozak CA.** 2010. Evolution of functional and sequence variants of the mammalian XPR1 receptor for mouse xenotropic gammaretroviruses and the human-derived XMRV. J. Virol. **84:**11970–11980.
17. **Yang H, Wang JR, Didion JP, Buus RJ, Bell TA, Welsh CE, Bonhomme F, Yu AH, Nachman MW, Pialek J, Tucker P, Boursot P, McMillan L, Churchill GA, de Villena FP.** 2011. Subspecific origin and haplotype diversity in the laboratory mouse. Nat. Genet. **43:**648–655.
18. **Yan Y, Liu Q, Kozak CA.** 2009. Six host range variants of the xenotropic/polytropic gammaretroviruses define determinants for entry in the XPR1 cell surface receptor. Retrovirology **6:**87.
19. **Wang JR, de Villena FP, McMillan L.** 2012. Comparative analysis and visualization of multiple collinear genomes. BMC Bioinformatics **13**(Suppl. 3):S13. doi:10.1186/1471-2105-13-S3-S13.
20. **Kent WJ.** 2002. BLAT–the BLAST-like alignment tool. Genome Res. **12:** 656–664.
21. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078–2079.
22. **Trapnell C, Pachter L, Salzberg SL.** 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:**1105–1111.
23. **Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J.** 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. **20:**265–272.
24. **Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ.** 2009. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics **25:**1189–1191.
25. **Yan Y, Knoper RC, Kozak CA.** 2007. Wild mouse variants of envelope genes of xenotropic/polytropic mouse gammaretroviruses and their XPR1 receptors elucidate receptor determinants of virus entry. J. Virol. **81:** 10550–10557.
26. **Gillin FD, Roufa DJ, Beaudet AL, Caskey CT.** 1972. 8-Azaguanine resistance in mammalian cells. I. Hypoxanthine-guanine phosphoribosyltransferase. Genetics **72:**239–252.
27. **Frankel WN, Stoye JP, Taylor BA, Coffin JM.** 1989. Genetic analysis of endogenous xenotropic murine leukemia viruses: association with two common mouse mutations and the viral restriction locus *Fv-1*. J. Virol. **63:**1763–1774.
28. **Frankel WN, Stoye JP, Taylor BA, Coffin JM.** 1989. Genetic identification of endogenous polytropic proviruses by using recombinant inbred mice. J. Virol. **63:**3810–3821.
29. **Morse HC, III.** 1978. Introduction, p 1–31. *In* Morse HC III (ed), Origins of inbred mice. Academic Press, New York, NY.
30. **Best S, LeTissier P, Towers G, Stoye JP.** 1996. Positional cloning of the mouse retrovirus restriction gene *Fv1*. Nature **382:**826–829.
31. **Taketo M, Schroeder AC, Mobraaten LE, Gunning KB, Hanten G, Fox RR, Roderick TH, Stewart CL, Lilly F, Hansen CT, Overbeek PA.** 1991. FVB/N: an inbred mouse strain preferable for transgenic analyses. Proc. Natl. Acad. Sci. U. S. A. **88:**2065–2069.
32. **Ware LM, Axelrad AA.** 1972. Inherited resistance to N- and B-tropic murine leukemia viruses *in vitro*: evidence that congenic mouse strains SIM and SIM.R differ at the *Fv-1* locus. Virology **50:**339–348.
33. **Baliji S, Liu Q, Kozak CA.** 2010. Common inbred strains of the laboratory mouse that are susceptible to infection by mouse xenotropic gammaretroviruses and the human-derived retrovirus XMRV. J. Virol. **84:**12841–12849.
34. **Yonekawa H, Moriwaki K, Gotoh O, Miyashita N, Matsushima Y, Shi L, Cho W, Zhen X, Tagashira Y.** 1988. Hybrid origin of Japanese mice "*Mus musculus molossinus*": evidence from restriction analysis of mitochondrial DNA. Mol. Biol. Evol. **5:**63–78.
35. **Jung YT, Wu T, Kozak CA.** 2003. Characterization of recombinant nonecotropic murine leukemia viruses from the wild mouse species *Mus spretus*. J. Virol. **77:**12773–12781.
36. **Teeter KC, Thibodeau LM, Gompert Z, Buerkle CA, Nachman MW, Tucker PK.** 2010. The variable genomic architecture of isolation between hybridizing species of house mice. Evolution **64:**472–485.
37. **Nellaker C, Keane TM, Yalcin B, Wong K, Agam A, Belgard TG, Flint J, Adams DJ, Frankel WN, Ponting CP.** 2012. The genomic landscape shaped by selection on transposable elements across 18 mouse strains. Genome Biol. **13:**R45. doi:10.1186/gb-2012-13-6-r45.
38. **Boursot P, Din W, Anand R, Darviche D, Dod B, VonDeimling F, Talwar GP, Bonhomme F.** 1996. Origin and radiation of the house mouse: mitochondrial DNA phylogeny. J. Evol. Biol. **9:**391–415.
39. **Din W, Anand R, Boursot P, Darviche D, Dod B, JouvinMarche E, Orth A, Talwar GP, Cazenave PA, Bonhomme F.** 1996. Origin and radiation of the house mouse: clues from nuclear genes. J. Evol. Biol. **9:**519–539.
40. **Marin M, Tailor CS, Nouri A, Kozak SL, Kabat D.** 1999. Polymorphisms of the cell surface receptor control mouse susceptibilities to xenotropic and polytropic leukemia viruses. J. Virol. **73:**9362–9368.
41. **Kozak CA, Hartley JW, Morse HC, III.** 1984. Laboratory and wild-derived mice with multiple loci for production of xenotropic murine leukemia virus. J. Virol. **51:**77–80.
42. **Elder JH, Gautsch JW, Jensen FC, Lerner RA, Hartley JW, Rowe WP.** 1977. Biochemical evidence that MCF murine leukemia viruses are envelope (*env*) gene recombinants. Proc. Natl. Acad. Sci. U. S. A. **74:**4676–4680.
43. **Chattopadhyay SK, Cloyd MW, Linemeyer DL, Lander MR, Rands E, Lowy DR.** 1982. Cellular origin and role of mink cell focus-forming viruses in murine thymic lymphomas. Nature **295:**25–31.
44. **Khan AS, Martin MA.** 1983. Endogenous murine leukemia proviral long terminal repeats contain a unique 190-base-pair insert. Proc. Natl. Acad. Sci. U. S. A. **80:**2699–2703.
45. **Flanagan JR, Krieg AM, Max EE, Khan AS.** 1989. Negative control region at the 5′ end of murine leukemia virus long terminal repeats. Mol. Cell. Biol. **9:**739–746.
46. **Levy DE, McKinnon RD, Brolaski MN, Gautsch JW, Wilson MC.** 1987. The 3′ long terminal repeat of a transcribed yet defective endogenous retroviral sequence is a competent promoter of transcription. J. Virol. **61:**1261–1265.
47. **Evans LH, Alamgir AS, Owens N, Weber N, Virtaneva K, Barbian K, Babar A, Malik F, Rosenke K.** 2009. Mobilization of endogenous retroviruses in mice after infection with an exogenous retrovirus. J. Virol. **83:** 2429–2435.
48. **Rosenke K, Lavignon M, Malik F, Kolokithas A, Hendrick D, Virtaneva K, Peterson K, Evans LH.** 2012. Profound amplification of pathogenic

murine polytropic retrovirus release from coinfected cells. J. Virol. **86:** 7241–7248.
49. **Groom HCT, Yap MW, Galão RP, Neil SJD, Bishop KN.** 2010. Susceptibility of xenotropic murine leukemia virus-related virus (XMRV) to retroviral restriction factors. Proc. Natl. Acad. Sci. U. S. A. **107:**5166–5171.
50. **Stavrou S, Nitta T, Kotla S, Ha D, Nagashima K, Rein AR, Fan H, Ross SR.** 2013. Murine leukemia virus glycosylated Gag blocks apolipoprotein B editing complex 3 and cytosolic sensor access to the reverse transcription complex. Proc. Natl. Acad. Sci. U. S. A. **110:**9078–9083.
51. **Nitta T, Lee S, Ha D, Arias M, Kozak CA, Fan H.** 2012. Moloney murine leukemia virus glyco-gag facilitates xenotropic murine leukemia virus-related virus replication through human APOBEC3-independent mechanisms. Retrovirology **9:**58.
52. **Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL.** 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature **487:**57–63.
53. **Wang T, Zeng J, Lowe CB, Sellers RG, Salama SR, Yang M, Burgess SM, Brachmann RK, Haussler D.** 2007. Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. Proc. Natl. Acad. Sci. U. S. A. **104:**18613–18618.
54. **Maksakova IA, Romanish MT, Gagnier L, Dunn CA, de Lagemaat LNV, Mager DL.** 2006. Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet. **2:**1–10. doi:10.1371/journal.pgen.0020002.
55. **Sanville B, Dolan MA, Wollenberg K, Yan Y, Martin C, Yeung ML, Strebel K, Buckler-White A, Kozak CA.** 2010. Adaptive evolution of *Mus Apobec3* includes retroviral insertion and positive selection at two clusters of residues flanking the substrate groove. PLoS Pathog. **6:**e1000974. doi:10.1371/journal.ppat.1000974.