JOURNAL OF VIROLOGY, Nov. 2010, p. 11970–11980 Vol. 84, No. 22
0022-538X/10/$12.00 doi:10.1128/JVI.01549-10

Evolution of Functional and Sequence Variants of the Mammalian XPR1 Receptor for Mouse Xenotropic Gammaretroviruses and the Human-Derived Retrovirus XMRV▽†

Yuhe Yan,[1]‡ Qingping Liu,[1] Kurt Wollenberg,[2] Carrie Martin,[1] Alicia Buckler-White,[1] and Christine A. Kozak[1]*

Laboratory of Molecular Microbiology[1] and Bioinformatics and Computational Biosciences Branch, Office of Cyber Infrastructure and Computational Biology,[2] National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 23 July 2010/Accepted 1 September 2010

Genetic conflicts between retroviruses and their receptors result in the evolution of novel host entry restrictions and novel virus envelopes, and such variants can influence trans-species transmission. We screened rodents and other mammals for sequence variation in the *Xpr1* receptor for the mouse xenotropic or polytropic mouse leukemia viruses (X-MLVs or P-MLVs, respectively) of the gammaretrovirus family and for susceptibility to mouse-derived X/P-MLVs and to XMRV (xenotropic murine leukemia virus-related virus), an X-MLV-like virus isolated from humans with prostate cancer and chronic fatigue syndrome. We identified multiple distinct susceptibility phenotypes; these include the four known *Xpr1* variants in *Mus* and a novel fifth *Xpr1* gene found in *Mus molossinus* and *Mus musculus*. We describe the geographic and species distribution of the *Mus Xpr1* variants but failed to find the X-MLV-restrictive laboratory mouse allele in any wild mouse. We used mutagenesis and phylogenetic analysis to evaluate the functional contributions made by constrained, variable, and deleted residues. Rodent *Xpr1* is under positive selection, indicating a history of host-pathogen conflicts; several codons under selection have known roles in virus entry. All non-*Mus* mammals are susceptible to mouse X-MLVs, but some restrict other members of the X/P-MLV family, and the resistance of hamster and gerbil cells to XMRV indicates that XMRV has unique receptor requirements. We show that the hypervariable fourth extracellular XPR1 loop (ECL4) contains three evolutionarily constrained residues that do not contribute to receptor function, we identify two novel residues important for virus entry (I579 and T583), and we describe a unique pattern of ECL4 variation in the three virus-restrictive *Xpr1* variants found in MLV-infected house mice; these mice carry different deletions in ECL4, suggesting either that these sites or loop size affects receptor function.

The XPR1 receptor mediates entry for the mouse leukemia viruses (MLVs) with xenotropic and polytropic host ranges (X-MLVs and P-MLVs, respectively). X-MLVs and P-MLVs can be isolated from laboratory mice and are capable of infecting cells of nonrodent species; these viruses are distinguished by the ability of P-MLVs, but not X-MLVs, to infect cells of the laboratory mouse and by the cytopathic and leukemogenic properties of P-MLVs, also termed MCF MLVs (mink cell focus-inducing MLVs) (11, 16, 24). XPR1 is also the receptor for several wild mouse isolates with an atypical host range (6, 48, 49) and for the recently described virus XMRV (xenotropic murine leukemia virus-related virus) (8), isolated from human patients with prostate cancer or chronic fatigue syndrome (27, 37, 43). Studies on the XPR1 receptor have identified residues critical for virus entry and described functionally distinct variants of XPR1 in human and rodent species that differ in their abilities to mediate entry of various virus isolates (18, 29, 31, 48, 49).

In *Mus*, four receptor variants of *Xpr1* are found in different taxonomic groups. $Xpr1^{n}$ was originally described in strains of the laboratory mouse (1, 41, 51), which are largely derived from *Mus domesticus* (50). $Xpr1^{c}$ was identified in the Asian species *Mus castaneus* (29, 31); $Xpr1^{p}$ is in the Asian species *Mus pahari* (48); and $Xpr1^{sxv}$ was found in several Eurasian species (18, 31). These variants are distinguished by their differential susceptibilities to prototype X-MLVs and P-MLVs as well as to two wild mouse isolates, CasE#1 and Cz524 (49); only $Xpr1^{sxv}$ encodes a receptor that is fully permissive for all isolates. The host range differences of these various virus isolates are due to sequence polymorphisms in both receptor and viral envelope genes.

The various mouse X/P-MLV isolates and the humanized XMRV define six different tropism patterns based on infectivity on rodent cells carrying *Xpr1* variants (49). These tropisms distinguish the two wild mouse isolates, CasE#1 and Cz524, and identify two P-MLV host range subgroups and two X-MLV/XMRV subgroups. Specific XPR1 residues responsible for entry of these viruses have been identified by analysis of rodent *Xpr1* variants and mutants. These receptor determinants lie in two of the four predicted extracellular loops (ECLs) of *Xpr1*, ECL3 and ECL4 (31, 44, 48, 49). Two critical amino acids have been defined for X-MLV entry:

* Corresponding author. Mailing address: NIH, NIAID, LMM, Bldg. 4, Room 329, 4 Center Drive MSC 0460, Bethesda, MD 20892-0460. Phone: (301) 496-0972. Fax: (301) 480-6477. E-mail: ckozak@niaid.nih.gov.

‡ Present address: Yisheng US Biopharma, Inc., 11 Firstfield Rd., Suite A, Gaithersburg, MD 20878.

† Supplemental material for this article may be found at http://jvi.asm.org/.

▽ Published ahead of print on 15 September 2010.

K500 in ECL3 and T582 in ECL4 (31). These two receptor determinants independently produce X-MLV receptors but are not functionally equivalent, as the Δ582T insertion into *Xpr1*^{n} generates a receptor for CasE#1, but the E500K substitution does not (48). Sensitivity to the six tropism subgroups is further modulated by specific substitutions at ECL3 residues 500, 507, and 508 (49). The sequence variations that distinguish the rodent XPR1 receptors can result in subtle differences in the efficiency of virus infection or complete resistance to specific X/P-MLVs.

The characterization of host genes that effect and/or block entry has obvious importance for a broader understanding of how viruses spread in natural populations and are transmitted to new hosts and how those populations adapt to retrovirus infections. The four house mouse species carry endogenous retroviruses (ERVs) for X-MLVs and P-MLVs (XMVs and PMVs, respectively) (3, 20, 42), and three of these species harbor infectious X-MLVs (4, 19, 48, 49). Restrictive variants of the XPR1 receptor have evolved in these virus-infected mice, along with the virus envelope (*env*) variants that define the tropism subgroups. We thus sought to examine the evolution of *Xpr1* in rodent species, and we extended this functional and sequence analysis to nonrodent species for two reasons. First, identification of XMRV in several human patient cohorts (27, 37, 43), the recent detection of P-MLV-related sequences in patients and blood donors (26), and the multiple instances of transspecies transmission of mouse gammaretroviruses (33) support an effort to describe factors that mediate or modulate virus entry in these species. Second, analysis of nonrodent species with novel patterns of virus restriction may uncover different or additional entry determinants. In the present study, we characterized 49 mice of different species or from different geographic locations and 24 other mammalian species for sequence and functional variants of the *Xpr1* receptor. We identified a novel 5th functional *Xpr1* variation in *Mus*, showed that restrictive XPR1 receptors in the three MLV-infected house mouse lineages have different deletions in ECL4, demonstrated that XPR1 is under positive selection, identified novel virus restriction phenotypes in nonrodent species, and demonstrated that XMRV relies on unique entry determinants.

MATERIALS AND METHODS

Viruses and cells. CAST-X is a xenotropic MLV isolated in our laboratory from the spleen of a CAST/EiJ mouse (48). Cz524 is a novel MLV isolated from the spleen of a CZECHII/EiJ mouse 2 months after inoculation with Moloney MLV (MoMLV) (49). The human xenotropic murine leukemia virus-related virus, XMRV (8), was kindly provided by R. Silverman (Cleveland Clinic, Cleveland, OH). AKR6, CasE#1, FrMCF, HIX MoMCF, and amphotropic A-MLV 4070A were originally obtained from J. Hartley (NIAID, Bethesda, MD).

Susceptibility to X/P-MLVs was tested in the cells from the following sources or cell lines: *Mus dunni* (21), ferret MA139 and goat lung cells obtained from J. Hartley (NIAID, Bethesda, MD), human 293 (CRL-1573), mink Mv-1-Lu (ATCC CCL64), Rat-2 (CRL-1764), gerbil GeLu (CCL-100), guinea pig JH4 clone 1 (CCL-158), dog MDCK (CCL-34), bat Tb-1-Lu (CCL-88), African green monkey kidney COS-1 (CRL-1650), cat CRFK (CCL-94), rabbit SIRC (CCL-60), buffalo IMR30 (obtained from M. Eiden; NIMH, Bethesda, MD), armadillo DNI.Tr (CRL-6009), Chinese hamster lines E36 (14) and Lec8 (CRL-1737), and E36 hamster cells transfected with *Xpr1* variants (48, 49). Embryo fibroblasts homozygous for *Xpr1*^{c} were described previously (48); tail fibroblasts were prepared from wild mouse-derived strains MOLD/EiJ, MOLG/EiJ, and CZECHII/EiJ as described previously (22).

Genomic DNA and RNA. DNA and RNA were isolated from animals, from cell lines that were developed from wild mice and wild mouse-derived breeding colonies or inbred strains (see Table S1 in the supplemental material), and from the mammalian cells listed above. Many wild-derived mice were obtained from M. Potter (NCI, Bethesda, MD). Mice trapped in California were provided by S. Rasheed (University of Southern California, Los Angeles). CAST/Rp mice were obtained from R. Elliott (Roswell Park Cancer Institute, Buffalo, NY). Cells from some wild mouse species were obtained from J. Rodgers (Baylor College of Medicine, Houston, TX). Mice or DNA samples from *M. castaneus* (CAST/EiJ), *M. domesticus* (PERA/EiJ, LEWES/EiJ, and WSB/EiJ), *M. musculus* (CZECHI/EiJ and CZECHII/EiJ), and various inbred lines derived from *Mus molossinus* were obtained from The Jackson Laboratory (Bar Harbor, ME). DNA samples from *M. musculus* (VEJ, and BRNO, Viborg, Belgrade) were obtained from S. Chattopadhyay and H. Morse (NIAID, Bethesda, MD). An African pygmy mouse sample was provided by H. Leathers (Philadelphia, PA). DNA samples from wild-trapped European *M. domesticus* mice were provided by M. Nachman (University of Arizona, Tucson).

A set of DNAs from African pygmy mouse DNA and other *Muridae* were obtained from Y. Cole and P. D'Eustachio (Departments of Biochemistry and Medicine, NYU, New York); the pygmy mice were classed into five species of subgenus *Nannomys* on the basis of skeletal features by J. T. Marshall (Smithsonian Natural History Museum, Washington, DC), and the other wild-trapped African *Muridae* (*Arvicanthis niloticus*, *Lemniscomys striatus*, *Hybomys univittatus*, *Uranomys ruddi*, *Lophuromys sikapusi*, *Lophuromys flavopunctatus*, *Mastomys natalensis*, *Mastomys fumatus*, *Mastomys stella*, and *Grammomys dolichurus*) were typed using standards at the NY Natural History Museum.

Mouse taxa are identified by genus and species, or where needed for clarification, by genus, subgenus, and species.

Pseudotype assay. Viral pseudotypes carrying the LacZ reporter were generated for the various X/P-MLVs by infecting GP2-293 cells transfected with pCL-MFG-LacZ as described previously (49). Cells were tested for susceptibility by infection with appropriate dilutions of these pseudotype virus stocks in the presence of 4 to 8 μg/ml Polybrene. One day after infection, cells were fixed with 0.4% glutaraldehyde and assayed for β-galactosidase activity, using as a substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal; 2 mg/ml) (ICN Biomedicals, Aurora, Ohio). Infectious titers were expressed as the number of blue cells per 50 μl of virus supernatant.

Cloning and sequencing of *Xpr1* genes. The full-length *Xpr1* genes from *Mus spicilegus*, *Mus shortridgei*, *M. domesticus* (WSB), *M. musculus* (CZECHI/EiJ), *Mus minutoides*, goat, cat, ferret, rabbit, gerbil, buffalo, armadillo, dog, African green monkey (AGMK), and bat cells were amplified by reverse transcription (RT)-PCR, cloned into pCR2.1-TOPO, and sequenced. Primers for amplification were 5′-ATGAAGTTCGCCGAGCACCTCTC and 5′-AGTGTTAGCTTC GTCATCTGTGTC.

Xpr1 exon13 sequences, containing ECL4, were amplified from DNA of the rodents and other mammals listed below in Fig. 1 and 2 by using primers Ex13F, 5′-GCCTATTACTACTGTGCC and Ex13R, 5′-CGGAAAACCTCAAGG GGG. *Xpr1* exons 10 to 12, which contain ECL3, were amplified from many of the same DNAs using the following two sets of primer pairs: Ex10-11F (5′-C GAGTATTTACTGCTCCCTTCC) and Ex10-11R (5′-GCAGGGATGGCA AAGTCCAGG) and Ex12F (5′-GGAGTTCCTGGAGAGAGAGTGG) and Ex12R (5′-GACCCCAGTCCATCTTGACATCC). All fragments were cloned into pCR2.1-TOPO before sequencing.

Generation of *Xpr1* mutants. Five novel mutant variants of the *Xpr1* gene were generated by mutagenesis PCR with the QuikChange II XL site-directed mutagenesis kit (Stratagene, La Jolla, CA), using as templates previously described clones of *Xpr1*^{n} and *Xpr1*^{sxv} (48). Mutations were produced using the following primers and their reverse complements: T583K (5′-CCCAACATGAGGCTTA AACTTTGTAGCAGTAATAGAG), I579Δ (5′-GACTATCCAAATCTCTAC TGCTACAACGTTTAAGC), S578A (5′-GGACTATCCAAATCGCCATTAC TGCTACAACG), T580A (5′-ATGAGGCTTAAACGTTGTAGCGGCAATA GAGATTTGGATAGTC), and G589A (5′-CGTTTAAGCCTCATGTTGCCG ACATCATTGCTACTG). All mutants were confirmed by sequencing.

The recombinant plasmids were transfected into E36 Chinese hamster cells. Stable transfectants were selected with Geneticin (830 μg/ml), and the expression of the *Xpr1* variants was confirmed by Western analysis. Proteins were extracted from transfected cells with M-PER mammalian protein extraction reagent (Pierce, Rockford, IL). The expression vector used for XPR1 inserts a V5 epitope at the C terminus; XPR1 expression was detected in Western blots using anti-V5 antibody (Invitrogen, Carlsbad, CA) followed by goat anti-mouse IgG conjugated with horseradish peroxidase (HRP) (Invitrogen). The membrane was then stripped and incubated with mouse anti-α-tubulin (Sigma, St. Louis, MO) and goat anti-mouse IgG conjugated with HRP (Invitrogen).

Selection analysis of lineages and codons. Two data sets were analyzed: one for the rodent ECL3-4 sequences and one for the rodent ECL4 sequences. DNA sequences were aligned using MUSCLE (9) and improved manually. Kimura two-parameter distance-based neighbor-joining phylogenies were generated for each set by PHYLIP (version 3.68) (10) and were corrected for closer correspondence to the appropriate consensus phylogeny of murid rodents (38) and mice (28, 45).

The codeml program of the PAML4 package (52) was used for maximum likelihood analysis of codon evolution (2). Both lineage-specific and codon-specific analyses were performed. In the lineage-specific selection analyses, the free-ratio model (codon model = 1) was used to calculate branch-specific rates of nonsynoymous/synonymous evolutionary changes (dN/dS). In this model, each branch is assumed to have a specific dN/dS ratio. The likelihood of the phylogeny under this model was tested against the likelihood of the phylogeny under the model of one uniform dN/dS ratio across all branches (codon model 0) using a likelihood ratio test (LRT). The significance of the LRT value was assessed using a chi-square distribution with 37 degrees of freedom for the ECL3/4 analysis and 53 degrees of freedom for the ECL4 analysis.

Selection acting on *Xpr1* codons was analyzed using two models of equilibrium codon frequencies and four models of codon selection. The two codon frequency models used were the F3x4 model (codon frequencies estimated from the nucleotide frequencies in the data at each codon site) and the F61 (codon table) model (frequencies of each of the 61 non-stop codons estimated from the data). The codon selection models were two neutral/negative selection models (M1 and M7) which were compared against corresponding positive selection models which included a category for $dN/dS > 1$ (M2 and M8, respectively). The significance of this additional codon selection category was assessed using LRTs of the phylogeny likelihoods under the neutral and positive selection models. Significance of the test statistics was calculated using a chi-square distribution with 2 degrees of freedom. The posterior probabilities of individual codons experiencing $dN/dS > 1$ was calculated using the naïve empirical Bayes (NEB) and Bayes empirical Bayes (BEB) algorithms (53).

Nucleotide sequence accession numbers. The sequences of the *Xpr1* genes have been deposited in GenBank under accession no. HQ022826 to HQ022857.

RESULTS

Functional and sequence variation of *Xpr1* in *Mus*. Mouse XPR1 is encoded by a gene on chromosome 1 with 15 exons over 140 kb. We screened 45 DNAs from species in all four *Mus* subgenera and mice trapped in different geographic locations for sequence variations in exon 13 (Fig. 1 and 2; see Table S1 in the supplemental material). Exon 13 encodes the *Xpr1* ECL4, a highly variable segment that contains receptor determinants (31). We also tested cells from available taxa for virus susceptibility.

ECL4 has only 13 amino acid residues, but deletions or replacement mutations in *Mus Xpr1* genes affect 8 codon sites (Fig. 1). These mutations define six receptor variants, four of which represent previously described alleles that differ in their ability to mediate entry of the various members of the X/P-MLV family of viruses (18, 29, 31, 48). One of the two novel sequence variants, here termed $Xpr1^{m}$, has a virus restrictive phenotype. A sixth sequence variant in the African pygmy mouse species *M. tenellus* was also identified (Fig. 1), but it was not characterized for receptor function. The species and geographic distribution of these sequence variants in *Mus* and the relationship between *Xpr1* variation and exposure to X/P-MLV infection can be described as follows.

(i) $Xpr1^{sxv}$. The largest number of *Xpr1* sequences, 27, had an $Xpr1^{sxv}$-like ECL4 (Fig. 2). The *sxv* receptor variant is fully permissive for all host range groups of X/P-MLVs (18, 49), and the 27 mice that carry this sequence include species in 3 of the 4 *Mus* subgenera (Fig. 2 and 3). We also identified this ECL4 sequence in other *Murinae*: rat and five wild-trapped mice in four species of two African genera, *Mastomys* and *Grammomys*

	Allele	ECL4 Sequence	Reference
Mus			
M. terricolor (dunni)	$Xpr1^{sxv}$	S I T A T T F K P H V G D	AF131097
Laboratory strains	$Xpr1^{n}$	S I T A - T F K P H V G N	AF131096
M. castaneus	$Xpr1^{c}$	S I T A - - - - - H V G D	AF131102
M. molossinus, M. musculus	$Xpr1^{m}$	S - T A T K F K P H V G D	HQ022828
M. pahari, M. shortridgei	$Xpr1^{p}$	S I T V T T F K P H V G D	EF606903, HQ022854
M. tenellus		S I T A T T F R P H V G D	HQ022830
Other Rodents			
Rat		S I T A T T F K P H V G D	NM_001105992
A. niloticus, L. striatus		S I T A T T L K P H V G D	HQ022826,HQ022827
H. univittatus		S M T A T A F K P R V G D	HQ022834
U. ruddi		S I T V T S L Q P N V G D	HQ022833
L. sikapusi, L. flavopunctatus		S I T H T A L Q P H V G D	HQ022831,HQ022832
Gerbil		S I T V T A L Q P H V G D	HQ022852
Chinese hamster		S I T A T A F Q P H V G D	AF131099
Kangaroo rat		S I T S T T K L S N V G D	ENSDORG00000001768
Guinea pig		S I T T M P T L P H S G D	ENSCPOG00000011532
Deer mouse		S I T T M P T L P H S G D	HQ022841
Other Mammals			
Human		S I T S T T L L P H S G D	AF115389
Monkey		S I T S T T L L P H S G D	HQ022844
Rabbit		S I T S T T L L P H S G D	HQ022850
Elephant		S I T S T T L L P H S G D	ENSLAFG00000004438
Pig		S I T T T T S L P H S G D	AK352504
Buffalo		S I T S T T S L P H S G D	HQ022857
Bat		S V T S T T L M P H T G D	HQ022842
Mink		S I T S M T L L P H S G D	AF131100
Ferret		S I T S M T L L P H S G D	HQ022851
Cat		S I T S M T L L P H S G D	HQ022843
Dog		S S T S M T S L P H S G D	HQ022847
Goat		S I T S M T S L P H S G D	HQ022848
Cow		S I T S M T S L P H S G D	XM_616684
Horse		S I T S M T L L P H S G D	XM_001488444
Armadillo		S I T S M T L L P H S G D	HQ022846
Platypus		S Y T T M D I F P Y A G D	XM_001515789

FIG. 1. Comparison of the deduced amino acid sequences of the ECL4 domains of the *Xpr1* genes of mice, other rodents, and selected other mammals. The five functionally characterized alleles in the genus *Mus* are listed by name. The box encloses the three *Mus* species that harbor viruses in the X/P-MLV family. Species with identical *Xpr1* sequences are listed together. Substitutions relative to the permissive *M. terricolor* (*dunni*) allele are highlighted in gray, and deletions are marked by dashes. Arrows and boldface indicate the three conserved residues.

(Fig. 2). The conservation of the *sxv*-like ECL4 sequence is consistent with the observation that most species of *Mus* as well as rat cells are permissive for X/P-MLV infection (6, 18, 23, 35). Surprisingly, this sequence was also found in all mice classed as *M. domesticus* (see below); these included six wild-trapped and three wild-derived lines of *M. domesticus* mice of Western Europe and North Africa, as well as three wild-caught California house mice and breeding lines of wild-derived mice from the Delmarva (Delaware-Maryland-Virginia) Peninsula and from South America.

(ii) $Xpr1^{p}$. The previously described variant $Xpr1^{p}$ restricts P-MLVs and Cz524, but not X-MLVs or CasE#1, and was originally described in *M. Coelomys pahari* (48). We also identified this variant in one additional Southeast Asian species, *M. Pyromys shortridgei*, although the other *Pyromys* species tested, *M. Pyromys saxicola*, carries $Xpr1^{sxv}$ (Fig. 2). Thus, *Mus* species of two subgenera carry a restrictive XPR1 receptor, while other Asian, palearctic, and African species, as well as other species of *Murinae*, carry the permissive *sxv* variant. Previous Southern blot analysis of genomic DNA from one of the two

ALLELE	TAXA	MLV SUSCEPTIBILITY			
		X-MLV	P-MLV	CasE#1	Cz524
Xpr1^{sxv}	*Mus domesticus* (ABUR, Birmingham, CalWM, Catalunya, Greece, Italy, LEWES, Mallorca, Papa Westray, PERA, *praetextus, poschiavinus,* WSB), *M. caroli, M. cervicolor, M. cookii, M. spicilegus, M. spretus, M. terricolor (dunni), M. sp. HL-2010, M. Nannomys bellus, M. Nannomys emesi, M. Nannomys gratus, M. Nannomys minutoides, M. Nannomys setulosus, M. Nannomys triton, M. Pyromys saxicola, Mastomys natalensis, M. fumatus, M. stella Grammomys dolichurus Rattus norvegicus*	+++	+++	+++	+++
Xpr1^{p}	*M. Pyromys shortridgei, M. Coelomys pahari*	+++	0	+++	0
Xpr1^{m}	*M. molossinus* (MOLD/RkJ, MOLF/EiJ, MOLG/DnJ), *M. musculus (Belgrade, BRNO, Czech-I, Czech-II,Skive, VEJ, Viborg)*	++	0	0	0
Xpr1^{c}	*M. castaneus* (CAST/EiJ, CAST/Ncr,CAST/Rp)	+++	0	0	+
Xpr1^{n}	-	0	+++	0	0
Xpr1^{t}	*M. tenellus*	NT	NT	NT	NT

FIG. 2. Distribution of ECL4 sequence variants of the six *Xpr1* alleles in *Murinae*. Subgenus designations are provided for species in the genus *Mus* that are not subgenus *Mus*. Former species designations, designations for wild-derived inbred strains, or trapping sites are given in parentheses. Three wild-trapped California mice and two wild-trapped *M. natalensis* mice carried identical *Xpr1* genes. Patterns of MLV susceptibility are given for mouse cells carrying prototype *Xpr1* alleles on the basis of $\log_{10}$ titers of infectivity with LacZ pseudotype: +++, >3; ++, 2 to 3; +, 0 to 1; 0, no blue cells. No wild mice were found to carry the laboratory mouse allele, *Xpr1*^{n}. The *Xpr1*^{t} sequence variant found in *M. tenellus* was not tested for receptor function.

species carrying the *Xpr1*^{p} variant, *M. pahari*, failed to identify X-MLV or P-MLV *env*-related endogenous copies (19). The absence of endogenous copies indicative of past virus exposure that might have selected for receptor mutations restricting entry suggests either that the evolutionary pressures resulting in the replacement mutations fixed in this species are unrelated to XPR1 receptor function, or these substitutions evolved in conjunction with an undiscovered retrovirus infection.

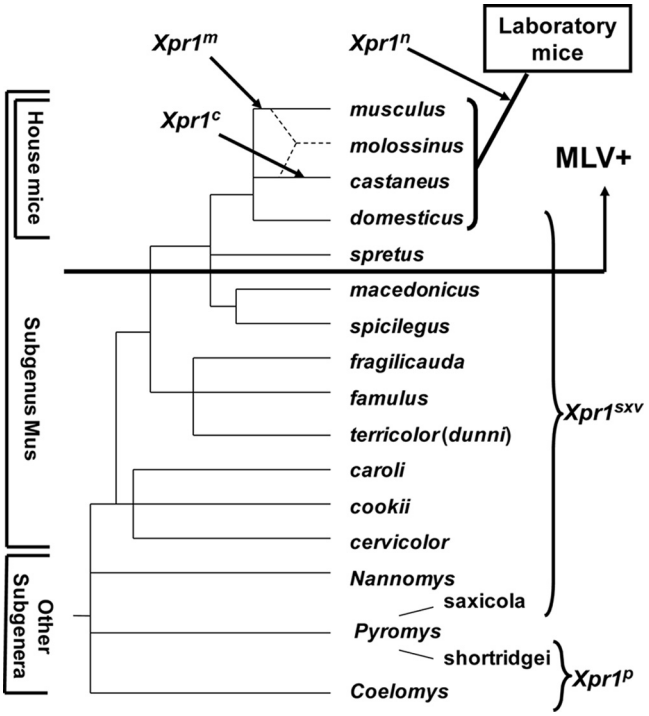

FIG. 3. Distribution of *Xpr1* variants and endogenous X/P-MLV *env* genes in the phylogenetic tree of the genus *Mus*. Brackets on the left indicate the four subgenera and the four house mouse species. Arrows and brackets mark the species that carry endogenous X/P-MLV *env* genes and the distribution of the five functionally defined *Xpr1* alleles among *Mus* species and strains. The tree is adapted from the synthetic trees developed by Guenet and others (15, 28, 45).

(iii) *Xpr1*^{c}. There are four species of house mice in three major lineages (Fig. 3). It is these house mouse species that uniquely carry endogenous X/PMV *env* sequences (3, 20, 42), and three different restrictive *Xpr1* variants are found in these mice (Fig. 2) (1, 18, 29, 41, 51). *Xpr1*^{c} restricts all viruses but X-MLV and was originally identified in *M. castaneus* (29) (Fig. 2). The sequence of this gene is identical to *Xpr1*^{sxv}, except for a five-codon deletion in ECL4 (31). *M. castaneus* mice are found in Southeast Asia and carry dozens of XMV *env* genes, some of which are capable of producing infectious virus (4, 20, 48), at least one of which is associated with interference-mediated resistance to infection (47). Three laboratory breeding lines of *M. castaneus*, CAST/EiJ, CAST/Rp, and CAST/Ncr, were found to carry *Xpr1*^{c} (Fig. 2). *Xpr1*^{c} was not, however, found in any of the three lines of the Japanese mouse *M. molossinus* (Fig. 2), which is a naturally occurring hybrid of *M. castaneus* and *M. musculus* (54) (Fig. 3).

(iv) *Xpr1*^{m}. A novel ECL4 sequence variant was identified in two house mouse species, the Japanese mouse *M. molossinus* and *M. musculus*, which has a range extending from eastern Europe to the Pacific. Both species carry dozens of copies of XMVs and few copies of PMV *env* genes (20, 42), and infectious virus has been isolated from both species (4, 19, 25, 49). The full-length *Xpr1*^{m} sequence is distinguishable from *Xpr1*^{sxv} at two sites, both of which are in ECL4: a deletion, I579Δ, and a substitution, T583K. To characterize receptor function of this novel variant, we infected cultured tail cells from three wild-derived strains carrying this gene: MOLG/DnJ, CZECHII/EiJ, and MOLF/EiJ. As shown for MOLG/DnJ (Fig. 4), these cells were susceptible to infection by amphotropic MLV but were not infected by P-MLVs or wild mouse X/P-MLVs and showed only low levels of susceptibility to CAST-X X-MLV. Such a restrictive phenotype can be due to the presence of interfering *env* genes as well as receptor polymorphism (47; reviewed in reference 39), so *Xpr1*^{m} was cloned and expressed in E36 Chinese hamster cells that show only trace levels of X-MLV susceptibility. Transfectants showed increased susceptibility to

Cells	Transfected *Xpr1*	ECL4	Log$_{10}$ LacZ Pseudotype Titer: CAST-X X-MLV	XMRV X-MLV	CasE#1	Cz524	P-MLV	A-MLV
MOLG/DnJ		S - T A T K F K P H V G D	0.4+/-0.4	<0	<0	<0	<0	3.4
E36	*Xpr1*^{m}	S - T A T K F K P H V G D	2.8+/-0.2	<0	<0	<0	<0	ND
E36	*Xpr1*^{sxv}-T583K	S I T A T K F K P H V G D	3.1+/-0.3	1.9+/-0.1	2.9+/-0.1	2.2+/-0.6	0.4+/-0.2	ND
E36	*Xpr1*^{sxv}-I579Δ	S - T A T T F K P H V G D	2.2+/-0.4	0.6+/-0.1	0.2+/-0.2	<0	<0	ND
E36			0.3+/-0.1	<0	<0	<0	<0	ND
E36	*Xpr1*^{sxv}	S I T A T T F K P H V G D	4.1+/-0.2	3.7+/-0.1	3.8+/-0.4	3.9+/-0.3	4.2+/-0.2	ND
M. dunni		S I T A T T F K P H V G D	5.2+/-0.3	3.8+/-0.4	5.1+/-0.1	4.7+/-0.1	4.4+/-0.7	4.1

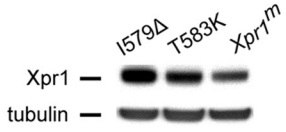

FIG. 4. Mutational analysis of *Mus* allele *Xpr1*^{m}. MOLG/DnJ cells and E36 cells expressing variant *Xpr1* genes were exposed to LacZ pseudotypes carrying the indicated Env protein; numbers represent titers ($\log_{10}$ in 50 μl) of LacZ-positive cells. ECL4 sequences are given for the exogenous *Xpr1* in transfected cells and the endogenous *Xpr1* in untransfected mouse cells. E36 cells show trace infectivity with CAST-X. The substitutions unique to *Xpr1*^{m} are highlighted in gray. On the right is a Western blot analysis of transfected E36 cells.

CAST-X compared to E36 cells and showed no susceptibility to other X/P-MLVs, indicating that this susceptibility profile is receptor mediated (Fig. 4).

The mutations found in *Xpr1*^{m}, I579Δ and T583K, were individually introduced into *Xpr1*^{sxv}. In transfected cells, the *Xpr1* clone with T583K showed a 100-fold reduction in susceptibility to X-MLVs and wild mouse X/P-MLVs relative to *M. dunni* cells but a 10,000-fold decrease in susceptibility to P-MLVs (Fig. 4). Cells expressing *Xpr1*-I579Δ showed poor but reproducible susceptibility to CAST-X and, at best, trace levels of infectivity with all other X/P-MLVs. This suggests that both mutations contribute to the virus resistance of this XPR1 variant, with the deletion having a greater effect on restriction. These results also confirm that the presence of K500 in ECL3 is not sufficient for efficient receptor function, consistent with the previous suggestion that the receptor interface requires residues in both ECLs (48).

(v) *Xpr1*^{n}. The *Xpr1*^{n} variant was originally cloned from laboratory mouse NIH 3T3 cells (1, 41, 51), which, like cells of various inbred laboratory strains, are resistant to infection by X-MLVs. The common strains of laboratory mice are a mosaic of three house mouse species, *M. musculus*, *M. domesticus*, and *M. castaneus*, with the largest genomic contribution coming from *M. domesticus* (50). The expectation that we would therefore find *Xpr1*^{n} in *M. domesticus* was also suggested by the fact that restriction of X-MLVs by *Xpr1*^{n} is consistent with the absence of XMV ERVs in *M. domesticus* (20). Surprisingly, however, all nine *M. domesticus* mice of Europe and North Africa carry *Xpr1*^{sxv} rather than *Xpr1*^{n}, as do wild-trapped or wild-derived mice from Peru and the East and West coasts of the United States (Fig. 2). Thus, if the *Xpr1*^{n} allele exists in wild mouse species, it has a limited distribution.

Previous mutational analysis of *Xpr1*^{n} determined that residues K500 and T582 function as independent receptor determinants for X-MLV and Cz524, but only T582 also produces a receptor for CasE#1. Because *Xpr1*^{n} also restricts XMRV, we tested *Xpr1*^{n} mutants carrying Δ582T or E500K for sensitivity to XMRV to determine if either or both of these residues are important for XMRV entry. Both transfectants were infected by XMRV, although XMRV infectivity was less efficient in the *Xpr1*^{n}-Δ582T mutant, with LacZ virus titers reduced 1 to 2 logs relative to the E500K mutant, although the cells were equivalently infectible by X-MLV (Fig. 5).

Sequence variation and positive selection of *Xpr1* in rodents. The substantial polymorphism seen in the *Xpr1* ECL4 of *Mus* is also characteristic of the 16 additional sequenced *Xpr1* genes of other rodent species (Fig. 1). ECL4 variability is due largely to replacement mutations; however, the three house mouse *Xpr1* alleles *Xpr1*^{c}, *Xpr1*^{m}, and *Xpr1*^{n}, and only these three alleles, carry deletions in this region of XPR1 (Fig. 1). These three alleles all encode restrictive XPR1 receptors that restrict two or more viruses in this family, and the appearance of these variants roughly coincides with the acquisition of MLV ERVs (Fig. 3). The hypervariability of this region in functional receptors suggests that XPR1 can tolerate multiple ECL4 replacement mutations, and the fact that three different deletions lie in this region in virus-positive mice suggests either that the six deleted residues are important for receptor function or, alternatively, that altering the size of this ECL may be an efficient way to disable receptor function in mice exposed to endemic retrovirus infection.

The additional sequence variants identified in other murid rodents were not tested for receptor function. Instead, we used a phylogenetic approach to try to identify residues showing

Cells	Transfected *Xpr1*	ECL3	ECL4	Log$_{10}$ LacZ Titer: CAST-X X-MLV	XMRV X-MLV	P-MLV
E36	*Xpr1*^{n}-Δ582T	E	S I T A T T F K P H V G N	4.4+/-0.9	2.1+/-0.1	4.4+/-0.3
	Xpr1^{n}-E500K	K	S I T A - T F K P H V G N	4.2+/-0.8	3.4+/-0.2	4.1+/-0.2
E36	*Xpr1*^{n}	E	S I T A - T F K P H V G N	<0	<0	3.9+/-0.2
E36	None			0.2+/-0.1	<0	<0
NIH 3T3	None	E	S I T A - T F K P H V G N	<0	<0	5.2+/-0.2
M.dunni	None	K	S I T A T T F K P H V G D	5.1+/-0.1	3.8+/-0.2	4.2+/-0.4

FIG. 5. Susceptibility of cells expressing *Xpr1*^{n} mutants to XMRV and X-MLV LacZ pseudotypes. ECL3 and ECL4 sequences are given for the exogenous *Xpr1* in transfected cells and the endogenous *Xpr1* in untransfected mouse cells. Expression of the *Xpr1* variants in these cells was previously described (49). Numbers represent titers ($\log_{10}$ per 50 μl) of LacZ-positive cells.

A

	M1-M2	M7-M8	Tree length	dN/dS (%)
F3x4	0.007794	0.000096	1.5697	7.75
F61	0.014047	0.000285	1.5599	8.58

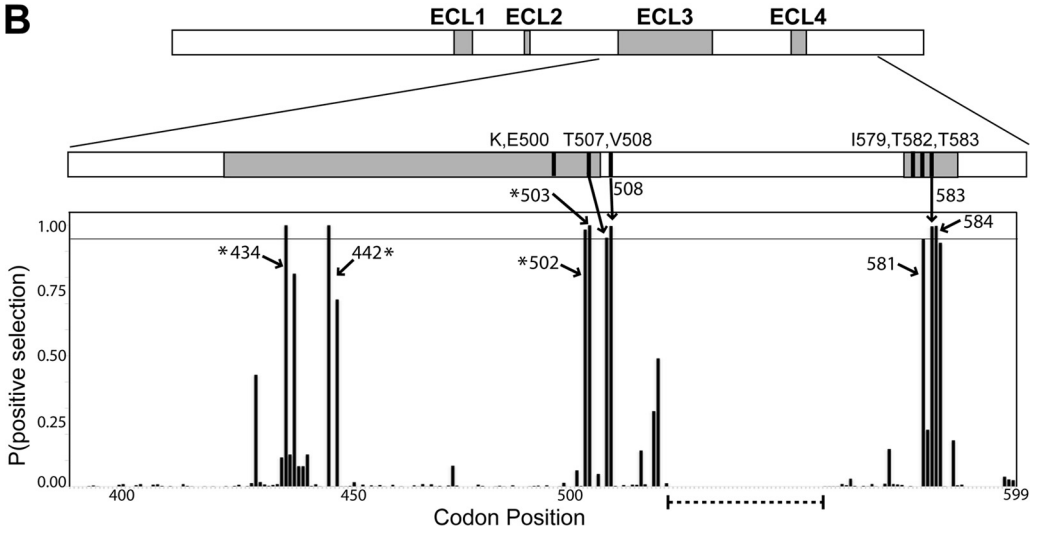

FIG. 6. Phylogenetic analysis of *Xpr1* in rodents. (A) Likelihood ratio tests were used to test for positive selection. Neutral models (M1 and M7) were compared with selection models (M2 and M8) using two different models of codon frequency (F3x4 or F61). Tree length is the average number of substitutions per codon along all branches. *dN*/*dS* ratios are given for the codons under selection, along with the percentage of codons in this category. Tree length and *dN*/*dS* (%) are given for the M8 analysis. (B) *Xpr1* sites under positive selection. At the top is a diagram of the *Xpr1* gene showing the locations of the four predicted extracellular loops and the locations of the six sites in ECL3 and ECL4 known to be involved in receptor function. At the bottom is a graph showing the posterior probability of positive selection at each codon based on an analysis of 19 *Xpr1* sequences using codon frequency model F3x4 and selection model 8 (see Table S2 in the supplemental material). A horizontal line marks $P = 0.95$, and arrows identify the nine codons under selection at $P > 0.95$ using the NEB calculation; asterisks mark codons also identified by the BEB calculation. The dotted line at the bottom marks the 35-codon segment that was not sequenced in all samples.

evidence of genetic conflicts. Host genes involved in antagonistic interactions with pathogens can be identified by sequence comparisons that reveal positive selection through an excess of nonsynonymous substitutions. To identify evidence of a possible antiviral role for *Xpr1* throughout *Mus* evolution, and to identify possible sites of virus interaction, we analyzed two sets of rodent *Xpr1* genes: a set of 28 sequences of the 138-bp exon that includes ECL4, and a 543-bp segment from 19 rodents that encodes ECL3 and ECL4 (see Fig. S1 and S2 in the supplemental material).

Taxonomy-based phylogenetic trees were generated for both sets of sequences. Values of *dN*/*dS* along each branch were calculated by using the free-ratio model of PAML. A *dN*/*dS* value of >1 suggests that positive selection has acted along that lineage. For the ECL3/4 DNA set, one branch shows clear positive selection, while none of the others had more than three inferred nonsynonymous changes when there were no synonymous changes (see Fig. S3 in the supplemental material); for the ECL4 set, only one branch had more than 3 nonsynonymous changes when there were zero synonymous changes (not shown).

Likelihood ratio tests of the ECL3/4 data set indicate that *Xpr1* has a significant probability of having experienced positive selection under the F3x4 model (Fig. 6A), whereas ratios for the ECL4 data set were below the level of significance (not shown). We calculated posterior probabilities of positive selection at individual codon sites in the ECL3/4 data set using two methods in the PAML suite of programs: the Bayes empirical Bayes (BEB) calculation and the naïve empirical Bayes calculation (NEB) (see Table S2 in the supplemental material). NEB identified nine codons under positive selection with posterior probability of >0.95; four of these sites were also identified by the more conservative BEB (Fig. 6B). Six of these selected codons are in ECL3. Two of these ECL3 codons, 434 and 442, lie near the 5′ end of this ECL, but the others (codons 502, 503, 507, 508) are clustered at the 3′ end in a region known to harbor entry determinants; positively selected residues T507 and V508 were previously identified as critical for entry of X/P-MLVs (49). The other ECL3 residues under positive selection have not been tested as receptor determinants, but 503 marks a potential glycosylation site in some XPR1 variants. The NEB analysis identified three codons in ECL4 under positive selection (Fig. 6B), although only one of these, 584, approached significance using BEB. Of these four codons, 583, as indicated in the $Xpr1^m$ analysis, affects P-MLV entry. Also, two of these selected ECL4 codons are deleted in $Xpr1^c$, and the deletion of one of the two adjacent Thr residues at 582 and 583 in $Xpr1^n$ eliminates an X-MLV receptor determinant (31). The fact that most of the amino acid sites identified as being under positive selection are at or near sites implicated in receptor function suggests that this gene may have experienced positive selection in *Mus* evolution because of an antiviral role.

***Xpr1* sequence variation in mammals.** Because of known examples of the transmission of mouse gammaretroviruses to other species, including humans (8, 26, 27, 33, 37), we extended our analysis of the functional and sequence variation of the

TABLE 1. Susceptibility of mammalian cells to LacZ pseudotypes of X/P-MLVs

Cells	Log_{10} LacZ pseudotype titer[a]							
	X-MLV		XMRV	X/P-MLV		P-MLV		A-MLV (4070A)
	CAST-X	AKR6		Cz524	CasE#1	HIX	FrMCF	
M. dunni	5.3 ± 0.2	5.1 ± 0.2	3.6 ± 0.2	4.6 ± 0.1	5.2 ± 0.1	4.2 ± 0.4	4.7 ± 0.9	4.7 ± 0.2
Human 293	4.9 ± 0.6	4.9 ± 0.1	4.1 ± 0.3	4.2 ± 0.2	5.1 ± 0.3	3.9	4.4 ± 0.1	4.4 ± 0.2
AGMK COS-1	5.2 ± 0.1	ND	3.9 ± 0.2	3.9 ± 0.3	4.3 ± 0.3	3.0 ± 0.5	2.7 ± 0.3	ND
Mink	5.8 ± 0.1	ND	4.3	5.8 ± 0.4	5.8 ± 0.7	4.6 ± 0.5	ND	4.4 ± 0.7
Ferret	5.2 ± 0.3	4.8 ± 0.3	3.7 ± 0.4	4.6 ± 0.3	5.1 ± 0.3	4.0 ± 0.6	4.7 ± 0.2	3.0
Rabbit SIRC	4.1 ± 0.4	3.8 ± 0.1	2.1 ± 0.6	3.1 ± 0.5	3.4 ± 0.4	2.0 ± 0.1	2.1 ± 0.3	ND
Cat CRFK	5.3 ± 0.6	5.0 ± 0.4	3.7 ± 0.7	4.7 ± 0.1	5.3 ± 0.5	3.8 ± 0.7	4.8 ± 0.7	ND
Armadillo DNI.Tr	3.9	3.4 ± 0.5	2.1	3.8	3.9 ± 0.1	1.8 ± 0.4	2.3 ± 0.2	1.0 ± 0.3
Rat-2	5.5	5.1 ± 0.1	3.1 ± 0.6	2.5 ± 0.3	5.1 ± 0.5	4.4 ± 0.5	0.3 ± 0.3	4.7 ± 0.6
Bat Tb-1-Lu	4.6 ± 0.2	3.9 ± 0.6	2.9 ± 0.3	3.8 ± 0.1	5.1 ± 0.1	0.6 ± 0.7	0.2 ± 0.2	ND
Guinea pig JH4	4.8 ± 0.2	4.0 ± 0.2	2.7 ± 0.1	2.4 ± 0.5	≤0	≤0	≤0	4.1 ± 0.2
Goat	4.9 ± 0.1	4.4 ± 0.1	2.9 ± 0.6	1.8 ± 0.4	≤0	≤0	0.1 ± 0.2	ND
Buffalo	4.4 ± 0.1	3.5 ± 0.1	1.8 ± 0.8	≤0	≤0	≤0	0.2 ± 0.2	ND
Dog MDCK	4.6 ± 0.1	4.0 ± 0.2	2.0 ± 0.7	≤0	≤0	≤0	≤0	ND
Gerbil GeLu	4.3 ± 0.3	3.5 ± 0.4	≤0	≤0	0.1 ± 0.2	≤0	≤0	2.1 ± 0.6
Chinese hamster E36	0.7 ± 0.2	≤0	≤0	≤0	≤0	≤0	≤0	ND

[a] Numbers represent titers (log_{10} in 50 μl) of β-galactosidase-positive cells. Where no SD is given, infectivity was only tested once. <0, no positive cells in cultures infected at least three times with undiluted pseudotype stock. Restriction is marked by underlining. ND, not done.

Xpr1 receptor to other mammals with special emphasis on species that serve as pets or farm animals. Cells from multiple species were tested for virus susceptibility (Table 1), and *Xpr1* sequences were determined or acquired from sequence databases (Fig. 1).

Alignment of these *Xpr1* genes confirms that there is striking hypervariability in ECL4, as seen in the murid genes (Fig. 1); however, this 13-residue ECL contains 3 nonvariant residues, S578, T580, and G589. Such nonvariant residues are often found near hypervariable regions in different virus receptors (40), and while mutational analysis has generally shown that such hypervariable residues are important for entry, it is also possible that these residues are not the actual receptor determinants. Instead, these variable residues may be negative control regions that govern accessibility to nearby conserved residues that serve as the actual viral attachment sites (40). Reliance on conserved rather than variable residues could provide the virus with a more reliable port of entry. To determine if these three conserved ECL4 residues are needed for receptor function, we mutated each of these sites in *Xpr1*^{sxv} (Fig. 7). When these *Xpr1* genes were expressed in hamster cells, however, all three were highly permissive receptors, suggesting these residues do not contribute significantly to the receptor interface.

Functional variation of XPR1 in mammals. Using the same seven X/P-MLV isolates that were used to test rodent cells for susceptibility, we identified distinctive patterns of virus susceptibility among the nonrodent mammals, including several patterns not seen in mice: for example, the restriction of P-MLVs and both wild mouse isolates by dog and buffalo cells and the restriction of P-MLVs by bat cells (Table 1). All of the cells were susceptible to X-MLV, as has been noted previously for some of them (6, 23, 35). The cells with restrictive receptors all restricted P-MLVs, and some of these cells also restricted CasE#1 or restricted both wild mouse MLVs. This indicates that X-MLVs are the most accommodating of receptor polymorphisms, and P-MLVs are the least accommodating. Comparison of the permissive and restrictive amino acid sequences identified 15 different replacements unique to restrictive genes in the regions involved in receptor function, namely, ECL4 and

Cells	Transfected Xpr1	ECL4	Log_{10} LacZ Pseudotype Titer				
			CAST-X X-MLV	XMRV X-MLV	CasE#1	Cz524	P-MLV
E36	*Xpr1*^{sxv}-S578A	A I T A T T F K P H V G D	4.7+/-0.2	3.4+/-0.1	4.5	3.9+/-0.1	3.7
E36	*Xpr1*^{sxv}-T580A	S I A A T T F K P H V G D	4.8+/-0.1	3.8+/-0.1	4.8+/-0.1	4.0+/-0.1	4.3+/-0.1
E36	*Xpr1*^{sxv}-G589A	S I T A T T F K P H V A D	4.7+/-0.3	3.5+/-0.2	4.8+/-0.2	4.1+/-0.1	3.8+/-0.1
E36	None		0.3+/-0.1	<0	<0	<0	<0
E36	*Xpr1*^{sxv}	S I T A T T F K P H V G D	4.1+/-0.2	3.7+/-0.1	3.8+/-0.4	3.9+/-0.3	4.2+/-0.2
M.dunni	None	S I T A T T F K P H V G D	5.2+/-0.3	3.8+/-0.4	5.1+/-0.1	4.7+/-0.1	4.4+/-0.7

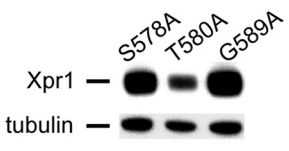

FIG. 7. Titers of LacZ pseudotypes on mouse and E36 cells expressing *Xpr1* genes with mutations in conserved residues. ECL4 sequence is given for endogenous *Xpr1* in untransfected *M. dunni* cells. Numbers represent titers (log_{10} per 50 μl) of LacZ-positive cells. Gray highlight indicates the three alanine substitutions. On the right is a Western blot analysis of transfected E36 cells.

Cells	*Transfected Xpr1*	ECL4	Log_{10} LacZ Titer CAST-X X-MLV	Log_{10} LacZ Titer XMRV X-MLV
GeLu gerbil		S I T V T A L Q P H V G D	4.3+/-0.3	<0
Lec8 hamster		S I T A T A F Q P H V G D	3.3+/-0.8	<0
E36 hamster		S I T A T A F Q P H V G D	0.3+/-0.1	<0
E36	*Xpr1-gerbil*	S I T V T A L Q P H V G D	2.9+/-0.2	<0
Gerbil	*Xpr1-monkey*	S I T S T T L L P H S G D	4.0+/-0.4	2.5+/-0.2
COS-1 monkey		S I T S T T L L P H S G D	5.2+/-0.1	3.9+/-0.2

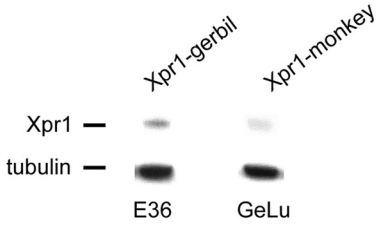

FIG. 8. Titers of LacZ pseudotypes on gerbil and hamster cells and cells expressing the indicated *Xpr1* variants. ECL4 sequences are given for endogenous *Xpr1* in untransfected cells. Numbers represent titers (log_{10} per 50 μl) of LacZ-positive cells. The Western blot analysis on the right shows lanes taken from the same exposure of the same film.

the 9-residue segment at the C-terminal end of ECL3 (data not shown). None of these replacements showed obvious correlations with specific patterns of restriction of P-MLVs, CasE#1 and Cz524.

Cells from two species were uniquely resistant to XMRV. As reported previously (34, 49), Chinese hamster cells are generally resistant to infection by X/P-MLVs, but interference with glycosylation by mutation or glycosylation inhibitors results in susceptibility to X-MLVs, but not to XMRV, AKR6 X-MLV, or P-MLVs. The current screening determined that gerbil GeLu cells are susceptible to all mouse X-MLVs, including AKR6, but are resistant to XMRV (Table 1 and Fig. 8). Expression of the gerbil XPR1 receptor in E36 hamster cells reproduced the gerbil susceptibility pattern (Fig. 8). To determine if there are additional factors in gerbil cells that restrict XMRV, we made a stable gerbil cell transfectant expressing exogenously introduced *Xpr1*. This transfectant expressed AGMK COS-1 *Xpr1*, which supports infection by both X-MLV and XMRV. These transfectants showed a small increase in susceptibility to X-MLV and were efficiently infected by XMRV, supporting the conclusion that gerbil cell resistance to XMRV is receptor mediated (Fig. 8).

We compared the sequences of the full-length XMRV-restricting *Xpr1* genes of Chinese hamsters and gerbils. Both genes carry K500 and T582, residues that function as independent X-MLV receptor determinants in the mouse and that also function as XMRV receptor determinants in the mouse (Fig. 5). The presence of both of these residues in hamster and gerbil cells suggests that infectivity by XMRV is restricted in these two cells by other residues. The XPR1 genes of these species differ from XMRV-permissive XPR1 genes at multiple sites, but these two restrictive genes have two shared substitutions not found in any XMRV-susceptible species (Fig. 1). Both of these residues, A583 and Q585, are in the ECL4 receptor-determining region, suggesting that one or both of these substitutions are responsible for this restriction. These results indicate that XMRV has unique receptor requirements that distinguish it from its X-MLV relatives.

DISCUSSION

Although no cellular function has been identified for mammalian XPR1, retrovirus receptors control the spread of elements responsible for genetic variation and disease and thus have important roles in the evolution of host species subject to infection (30). In this study, we describe XPR1 sequence and functional variation in the natural hosts of these gammaretroviruses and also in other mammalian species that could serve as virus reservoirs or targets for trans-species transmission. This analysis has allowed us to address two questions: (i) how receptor polymorphism helps natural populations adapt to exposure to infectious disease-inducing retroviruses and (ii) whether the potential for transspecies transmission in different mammalian species is restricted or facilitated by functional XPR1 polymorphisms. In the course of these studies, we characterized a novel restrictive *Xpr1* allele in *Mus* and identified a correlation between exposure to MLV infection and appearance of restrictive XPR1s caused by deletion. We used mutagenesis and phylogenetics to evaluate the functional contributions made by constrained, variable, and deleted residues in XPR1.

There are numerous examples of trans-species transmissions of retroviruses. The most well-known example is, of course, the derivation of HIV-1 from simian lentiviral precursors, but there are other examples, and retroviruses that cluster with mouse gammaretroviruses are found in multiple vertebrates. Martin and colleagues (33) found MLV-related ERVs in approximately one-fourth of vertebrate taxa and identified recent zoonotic transmissions from mammals to birds and from eutherians to metatherians. It would not be surprising to find more examples of interspecies transmissions involving MLVs, since MLV-infected house mouse species have a worldwide geographic distribution (32), and all mammalian species tested are permissive to infection by some or all X/P-MLVs.

The XMRV virus found in some human patients may have been acquired directly from mice or after transmission from mice to another species in contact with humans. We therefore examined the reported worldwide incidence of prostate cancer (7) relative to the geographic distribution of the various *Xpr1* alleles of house mice. The highest rates of this disease are found in the United States, and the lowest rates are found in Asian countries like Japan, India, and China. In Europe, rates are highest in Austria, France, and Scandinavia and lowest in Eastern Europe. This distribution roughly corresponds to the distribution of *Xpr1* receptor variants in mouse populations, with the most permissive allele, *sxv*, found in high-tumor-incidence areas like America and Western Europe, and with the

allele most restrictive of X-MLVs, *Xpr1*^{m}, found in low-tumor areas like Japan and Eastern Europe. This Eurasian group of mice also carries receptor-blocking genes (47) that further suggest that these mice might provide a poor reservoir for zoonotic transmission to humans. While the American and Western European mice with *Xpr1*^{sxv} carry only poorly expressed endogenous PMVs and few or no endogenous XMVs, wild mice are also known to carry viruses that have not become endogenized (12, 36, 46), and at least one virus in the X/P-MLV family, CasE#1, was isolated from a mouse trapped in California. Clearly, more studies are needed to examine wild mouse populations for infectious virus.

The isolation of XMRV from human patients raises the question of whether adaptation to this species has altered its tropism. XMRV shows 95% overall identity to X-MLVs and 93% identity in the 5′ *env* receptor binding domain. XMRV, as shown here, can utilize the receptor determinants for X-MLVs defined by residues K500 and T582 to enter mouse cells, although K500 is a more effective receptor for XMRV. Unlike the mouse X-MLVs that are able to infect all mammals, however, XMRV is restricted by hamster and gerbil cells and these XPR1 genes share two substitutions, T583A and L/K585Q. Residues A583 and Q585 are not found in other mammals, except for a few rodents that have not been tested for XMRV susceptibility. Other evidence implicates these two codons in virus entry: both residues are deleted in *Xpr1*^{c}, 583 is under positive selection, and the T583K substitution found in *M. molossinus* was associated with P-MLV restriction. This XMRV tropism difference demonstrates that the XMRV-receptor interaction differs from that of the mouse viruses and raises the possibility that this may represent an XMRV adaptation acquired through contact with humans or with an as-yet-undiscovered species before transmission to humans.

There are six XPR1 codons important for virus entry identified by mutagenesis of *Mus* alleles (500, 507, 508, 579, 582, and 583) (31, 48, 49). Expanding this receptor analysis to other mammals should be fruitful; analysis of the gerbil XPR1 has implicated A583 and Q585 in XMRV entry. That further studies on the human and other XPR1 genes could provide additional information on the receptor virus interface is suggested by several observations: the extensive sequence variation in permissive mammalian XPR1s in the regions involved in entry; the fact that various mammalian cells, like bat and dog cells, show resistance patterns not found in mice; the fact that studies on human/hamster *Xpr1* chimeras implicate as yet unidentified ECL4 residues in P-MLV entry (44); and the fact that X/P-MLV interference patterns differ in cells with different *Xpr1* genes (5, 6; C. A. Kozak, unpublished data).

Phylogenetic analysis of *Xpr1* indicates that it is under diversifying selection in rodents, suggesting that this gene has had a defensive role in rodent evolution. There is ample evidence that some wild mouse populations have been exposed to viruses that use the XPR1 receptor: Some species carry MLV ERVs acquired from past infections (20, 42), some of which can produce infectious virus (4, 19, 48, 49), and some wild mice also carry infectious MLVs that have not become endogenized (12, 46, 36). For such populations, survival is enhanced by host factors that restrict virus, and the XPR1 receptor is clearly one of those factors. Among the five *Mus* XPR1 variants, the one with the broadest susceptibility phenotype, *Xpr1*^{sxv}, is widely distributed among the Eurasian *Mus* species, and the *sxv* ECL4 sequence is also found in multiple species that predate *Mus*. The species with this ECL4 sequence either lack X/P-MLV ERVs or carry only PMVs, ERVs not known to produce infectious virus (Fig. 3) (17, 20).

The four restrictive *Mus* XPR1 polymorphisms appeared at two distinct points in *Mus* evolution. First, *Xpr1*^{p} appeared about 7.5 million years ago (MYA), shortly after the divergence of *Mus* from other *Murinae* (28, 45), but this allele is confined to two species of Southeast Asian mice. The presence of this restrictive receptor in mice that do not carry XMV or PMV ERVs (20) suggests that either the fixation of this variant is unrelated to receptor function, or these mice were exposed to an as-yet-undescribed retrovirus infection. The other three restrictive XPR1s arose later in *Mus* evolution, in the house mouse complex, which appeared about 0.5 MYA (13). House mice are distinguishable from other *Mus* species by two notable features. First, these mice are behaviorally different from other *Mus* species in their dependence on humans; these mice live in our houses, barns, warehouses, and ships and travel wherever we go. Second, these house mice have all been exposed to MLV infection and carry numerous endogenous copies of X/P-MLVs, some of which have remained active. Acquisition of these germ line viral sequences is roughly coincident with the appearance of the restrictive house mouse variants *Xpr1*^{m} and *Xpr1*^{c}, both of which, like the ERV sequences in these species, show an apparent species-wide distribution. Thus, the mutations in these two variants restricting *Xpr1* receptor function likely provided a survival advantage in the face of endemic infection by potentially mutagenic and pathogenic gammaretroviruses and may have contributed to the "arms race" between virus and host by providing the selective pressure that produced viral variants with altered receptor specificities.

We had expected *Xpr1*^{n} to be widespread in wild *M. domesticus* for several reasons. First, *Xpr1*^{n} was initially identified in laboratory mice, and although *M. musculus*, *M. castaneus*, and *M. domesticus* all contributed to the fancy mouse colonies used to generate the common inbred strains, *M. domesticus* is, by far, the most significant contributor to the laboratory mouse genome (50). Second, the complete absence of XMV *env* genes in *M. domesticus* is consistent with *Xpr1*^{n} restriction of X-MLV infection. The failure to identify *Xpr1*^{n} in any *M. domesticus* mouse trapped in geographically disparate locations in Western Europe and the Americas indicates, however, that this allele is not responsible for the absence of XMVs in these mice, an absence that also marks all other species with *Xpr1*^{sxv}. The failure to identify *Xpr1*^{n} in the wild may be a consequence of limited sampling, but the fact that *Xpr1*^{n} does not show the apparent species-wide distribution of *Xpr1*^{m} and *Xpr1*^{c} suggests that this allele evolved only recently in *Mus*, perhaps in the fancy mouse colonies that provided progenitors of the common laboratory strains (50). Testing of additional Eurasian wild mouse populations and laboratory breeding stocks may identify the origin of this restrictive allele.

The presence of different deletion mutations in the 13-residue ECL4 of the three restrictive alleles found in house mice is unusual in the natural history of this gene. No ECL4 deletion mutations were found in XPR1 genes of nonrodent mammals, and none appeared during the 7 million years of *Mus* evolution prior to the house mouse radiation. These independent dele-

tion mutations are thus unlikely to have resulted from inherent structural features in this region of the gene, because such deletions would not then be restricted to house mice. The appearance of these variants, each of which blocks two or more viruses in the XPR1 family, coincides with the acquisition of germ line MLV *env* genes, suggesting that, in the face of infection, replacement mutations may not be an effective way to disable or alter receptor function. After all, permissive receptors have sustained multiple ECL4 substitutions. That all three of the restrictive alleles carry deletion mutations suggests either that these six residues are critical for entry or that the size of the ECL4 loop may affect receptor function.

Finally, these studies do not provide any special insight into the origins or evolutionary relationships of the two major host range variants of MLVs that use XPR1, X-MLV and P-MLV. The fact that P-MLVs are found in *M. domesticus* that were thought to carry *Xpr1*^{n} prompted the reasonable suggestion that these viruses may have evolved in response to XPR1 mutations that enabled European mice to evade X-MLV infections (31, 49). Our present studies, however, indicate that the acquisition of P-MLV ERVs predates *Xpr1*^{n}. The observed distribution of ERVs and XPR1 variants also does not explain why P-MLVs integrated selectively into European mice carrying *Xpr1*^{sxv} but are not found in other Eurasian species carrying the same receptor variant.

ACKNOWLEDGMENT

This research was supported by the Intramural Research Program of the NIH, NIAID.

REFERENCES

1. **Battini, J. L., J. E. Rasko, and A. D. Miller.** 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.
2. **Bielawski, J. P., and Z. Yang.** 2005. Likelihood methods for detecting adaptive evolution, p 103–124. *In* R. Nielsen (ed.), Statistical methods in molecular evolution. Springer, New York, NY.
3. **Ch'ang, L.-Y., W. K. Yang, F. E. Myer, C. K. Koh, and L. R. Boone.** 1989. Specific sequence deletions in two classes of murine leukemia virus-related proviruses in the mouse genome. Virology **168:**245–255.
4. **Chattopadhyay, S. K., M. R. Lander, S. Gupta, E. Rands, and D. R. Lowy.** 1981. Origin of mink cytopathic focus-forming (MCF) viruses: comparison with ecotropic and xenotropic murine leukemia virus genomes. Virology **113:**465–483.
5. **Chesebro, B., and K. Wehrly.** 1985. Different murine cell lines manifest unique patterns of interference to superinfection by murine leukemia viruses. Virology **141:**119–129.
6. **Cloyd, M. W., M. H. Thompson, and J. W. Hartley.** 1985. Host range of mink cell focus-inducing viruses. Virology **140:**239–248.
7. **Curado, M. P., B. Edwards, H. R. Shin, H. Storm, J. Ferlay, M. Heanue, and P. Boyle. (ed.)** 2007. Cancer incidence in five continents, vol. IX. IARC scientific publication no. 160. IARC, Lyon, France. http://www.iarc.fr/publications/pdfs-online/epi/sp160/index.php.
8. **Dong, B., S. Kim, S. Hong, J. Das Gupta, K. Malathi, E. A. Klein, D. Ganem, J. L. DeRisi, S. A. Chow, and R. H. Silverman.** 2007. An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors. Proc. Natl. Acad. Sci. U. S. A. **104:**1655–1660.
9. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792–1797.
10. **Felsenstein, J.** 2004. PHYLIP (Phylogenetic Inference Package) version 3.6. Department of Genome Sciences, University of Washington, Seattle.
11. **Fischinger, P. J., S. Nomura, and D. P. Bolognesi.** 1975. A novel murine oncornavirus with dual eco- and xenotropic properties. Proc. Natl. Acad. Sci. U. S. A. **72:**5150–5155.
12. **Gardner, M.** 2008. Search for oncogenic retroviruses in wild mice and man: historical reflections. Cancer Ther. **6:**285–302.
13. **Geraldes, A., P. Basset, B. Gibson, K. L. Smith, B. Harr, H.-T. Yu, N. Bulatova, Y. Ziv, and M. W. Nachman.** 2008. Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. Mol. Ecol. **17:**5349–5363.
14. **Gillin, F. D., D. J. Roufa, A. L. Beaudet, and C. T. Caskey.** 1972. 8-Azaguanine resistance in mammalian cells. I. Hypoxanthine-guanine phosphoribosyltransferase. Genetics **72:**239–252.
15. **Guenet, J.-L., and F. Bonhomme.** 2003. Wild mice: an ever-increasing contribution to a popular mammalian model. Trends Genet. **19:**24–31.
16. **Hartley, J. W., N. K. Wolford, L. J. Old, and W. P. Rowe.** 1977. A new class of murine leukemia virus associated with development of spontaneous lymphomas. Proc. Natl. Acad. Sci. U. S. A. **74:**789–792.
17. **Jung, Y. T., T. Wu, and C. A. Kozak.** 2003. Characterization of recombinant nonecotropic murine leukemia viruses from the wild mouse species *Mus spretus*. J. Virol. **77:**12773–12781.
18. **Kozak, C. A.** 1985. Susceptibility of wild mouse cells to exogenous infection with xenotropic leukemia viruses: control by a single dominant locus on chromosome 1. J. Virol. **55:**690–695.
19. **Kozak, C. A., J. W. Hartley, and H. C. Morse III.** 1984. Laboratory and wild-derived mice with multiple loci for production of xenotropic murine leukemia virus. J. Virol. **51:**77–80.
20. **Kozak, C. A., and R. R. O'Neill.** 1987. Diverse wild mouse origins of xenotropic, mink cell focus-forming, and two types of ecotropic proviral genes. J. Virol. **61:**3082–3088.
21. **Lander, M. R., and S. K. Chattopadhyay.** 1984. A *Mus dunni* line that lacks sequences closely related to endogenous murine leukemia viruses and can be infected by ecotropic, amphotropic, xenotropic, and mink cell focus-forming viruses. J. Virol. **52:**695–698.
22. **Lander, M. R., B. Moll, and W. P. Rowe.** 1978. A procedure for culture of cells from mouse tail biopsies: brief communication. J. Natl. Cancer Inst. **60:**477–478.
23. **Levy, J. A.** 1975. Host range of murine xenotropic virus: replication in avian cells. Nature **253:**140–142.
24. **Levy, J. A., and T. Pincus.** 1970. Demonstration of biological activity of a murine leukemia virus of New Zealand black mice. Science **170:**326–327.
25. **Lieber, M., C. Sherr, M. Potter, and G. Todaro.** 1975. Isolation of type-C viruses from the Asian feral mouse *Mus musculus molossinus*. Int. J. Cancer **15:**211–220.
26. **Lo, S.-C., N. Pripuzova, B. Li, A. L. Komaroff, G.-C. Hung, R. Wang, and H. J. Alter.** 2010. Detection of MLV-related virus gene sequences in blood of patients with chronic fatigue syndrome and healthy blood donors. Proc. Natl. Acad. Sci. U. S. A. **107:**15874–15879.
27. **Lombardi, V. C., F. W. Ruscetti, J. Das Gupta, M. A. Pfost, K. S. Hagen, D. L. Peterson, S. K. Ruscetti, R. K. Bagni, C. Petrow-Sadowski, B. Gold, M. Dean, R. H. Silverman, and J. A. Mikovits.** 2009. Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. Science **326:**585–589.
28. **Lundrigan, B. L., S. A. Jansa, and P. K. Tucker.** 2002. Phylogenetic relationships in the genus *Mus*, based on paternally, maternally, and biparentally inherited characters. Syst. Biol. **51:**410–431.
29. **Lyu, M. S., and C. A. Kozak.** 1996. Genetic basis for resistance to polytropic murine leukemia viruses in the wild mouse species *Mus castaneus*. J. Virol. **70:**830–833.
30. **Maksakova, I. A., M. T. Romanish, L. Gagnier, C. A. Dunn, L. N. van de Lagemaat, and D. L. Mager.** 2006. Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet. **2:**e2.
31. **Marin, M., C. S. Tailor, A. Nouri, S. L. Kozak, and D. Kabat.** 1999. Polymorphisms of the cell surface receptor control mouse susceptibilities to xenotropic and polytropic leukemia viruses. J. Virol. **73:**9362–9368.
32. **Marshall, J. T.** 1981. Taxonomy, p 17–26. *In* H. L. Foster, J. D. Small, and J. G. Fox (ed.), The mouse in biomedical research, vol. 1. Academic Press, Inc., New York, NY.
33. **Martin, J., E. Herniou, J. Cook, R. W. O'Neill, and M. Tristem.** 1999. Interclass transmission and phyletic host tracking in murine leukemia virus-related retroviruses. J. Virol. **73:**2442–2449.
34. **Miller, D. G., and A. D. Miller.** 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.
35. **Oie, H. K., E. K. Russell, J. H. Dotson, J. M. Rhoads, and A. F. Gazdar.** 1976. Host-range properties of murine xenotropic and ecotropic type-C viruses. J. Natl. Cancer Inst. **56:**423–426.
36. **O'Neill, R., J. W. Hartley, R. Repaske, and C. A. Kozak.** 1987. Amphotropic proviral envelope sequences are absent from the *Mus* germ line. J. Virol. **61:**2225–2231.
37. **Schlaberg, R., D. J. Choe, K. R. Brown, H. M. Thaker, and I. R. Singh.** 2009. XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors. Proc. Natl. Acad. Sci. U. S. A. **106:**16351–16356.
38. **Steppan, S. J., R. M. Adkins, and J. Anderson.** 2004. Phylogeny and divergence-date estimates of rapid radiations in murid rodents based on multiple nuclear genes. Syst. Biol. **53:**533–553.
39. **Stocking, C., and C. A. Kozak.** 2008. Murine endogenous retroviruses. Cell. Mol. Life Sci. **65:**3383–3398.
40. **Tailor, C. S., D. Lavillette, M. Martin, and D. Kabat.** 2003. Cell surface receptors for gammaretroviruses. Curr. Top. Microbiol. Immunol. **281:**29–106.

41. **Tailor, C. S., A. Nouri, C. G. Lee, C. Kozak, and D. Kabat.** 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.
42. **Tomonaga, K., and J. M. Coffin.** 1998. Structure and distribution of endogenous nonecotropic murine leukemia viruses in wild mice. J. Virol. **72:**8289–8300.
43. **Urisman, A., R. J. Molinaro, N. Fischer, S. J. Plummer, G. Casey, E. A. Klein, K. Malathi, C. Magi-Galluzzi, R. R. Tubbs, D. Ganem, R. H. Silverman, and J. L. DeRisi.** 2006. Identification of a novel gammaretrovirus in prostate tumors of patients homozygous for R462Q *RNASEL* variant. PloS Pathog. **2:**e25.
44. **Van Hoeven, N. S., and A. D. Miller.** 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.
45. **Veyrunes, F., G. Dobigny, F. Yang, P. C. M. O'Brien, J. Catalan, T. J. Robinson, and J. Britton-Davidian.** 2006. Phylogenomics of the genus *Mus* (Rodentia; Muridae): extensive genome repatterning is not restricted to the house mouse. Proc. R. Soc. Lond. B **273:**2925–2934.
46. **Voytek, P., and C. A. Kozak.** 1989. Nucleotide sequence and mode of transmission of the wild mouse ecotropic virus, HoMuLV. Virology **173:**58–67.
47. **Wu, T., Y. Yan, and C. A. Kozak.** 2005. *Rmcf2*, a xenotropic provirus in the Asian mouse species *Mus castaneus*, blocks infection by polytropic mouse gammaretroviruses. J. Virol. **79:**9677–9684.
48. **Yan, Y., R. C. Knoper, and C. A. Kozak.** 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.
49. **Yan, Y., Q. Liu, and C. A. Kozak.** 2009. Six host range variants of the xenotropic/polytropic gammaretroviruses define determinants for entry in the XPR1 cell surface receptor. Retrovirology **6:**87.
50. **Yang, H., T. A. Bell, G. A. Churchill, and F. P.-M. de Villena.** 2007. On the subspecific origin of the laboratory mouse. Nat. Genet. **39:**1100–1107.
51. **Yang, Y.-L., L. Guo, S. Xu, C. A. Holland, T. Kitamura, K. Hunter, and J. M. Cunningham.** 1999. Receptors for polytropic and xenotropic mouse leukaemia viruses encoded by a single gene at *Rmc1*. Nat. Genet. **21:**216–219.
52. **Yang, Z.** 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13:**555–556.
53. **Yang, Z., W. S. W. Wong, and R. Nielsen.** 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. **22:**1107–1118.
54. **Yonekawa, H., O. Gotoh, Y. Tagashira, Y. Matsushima, L.-I. Shi, W. S. Cho, N. Miyashita, and K. Moriwaki.** 1986. A hybrid origin of Japanese mice "*Mus musculus molossinus*." Curr. Top. Microbiol. Immunol. **127:**62–67.