# Characterization of a Molecularly Cloned Retroviral Sequence Associated with *Fv-4* Resistance

HIDETOSHI IKEDA,\* FREDERIC LAIGRET, MALCOLM A. MARTIN, AND ROY REPASKE

Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20205

Received 4 March 1985/Accepted 9 May 1985

The murine leukemia virus (MuLV) sequence associated with the resistance allele of the Fv-4 gene (Fv-4') was molecularly cloned from genomic DNA of uninfected mice carrying this allele. The 5.2-kilobase cloned EcoRIDNA fragment (pFv4) was shown by nucleotide sequencing to contain 3.4 kilobases of a colinear MuLV-related proviral sequence which began in the C-terminal end of the *pol* region and extended through the *env* region and the 3' long terminal repeat. Cellular sequences flanked the 3' as well as the 5' ends of the truncated MuLV sequence. Alignment of the N-terminal half of the pFv4 *env* sequence with ecotropic, mink cell focus-forming, and xenotropic MuLV *env* sequences established the relatedness of pFv4 and ecotropic MuLV *env* sequences. A subcloned 700-base pair segment (pFv4<sub>env</sub>) from the 5' *env* region of pFv4 was used as an Fv-4-specific probe; it hybridized specifically to the Fv-4'-associated proviral sequence but not to endogenous ecotropic MuLV proviral DNA under high stringency. All Fv-4-resistant mice contained the same retroviral segment associated with the same flanking cellular DNA. Expression of Fv-4'-specific mRNA was demonstrated in the spleens of Fv-4' mice but not  $Fv-4^s$  mice, supporting the previously proposed resistance model based on interference.

Retroviral infections and retrovirus-induced diseases are modulated by several host genes. Some of these genes (for example, Fv-1 and Fv-4) control the susceptibility of target cells to viral replication, others (such as SI and W) control the size of the target cell population, and others (H-2-linked genes such as Rfv-1 and Rfv-2) control immunological reactions of host to viruses or virus-infected cells (43). The Fv-4gene, which controls susceptibility to infection by ecotropic murine leukemia viruses (MuLVs) (18, 41, 47), was first described in a G stain of laboratory mouse (now called FRG). Subsequently, Fv-4-like resistance has been identified in wild Asian mice (Fv-4w') (31, 32) and in wild California mice ( $Akvr-I^R$ ) (10). Genetic studies have shown that resistance in these different mice is due to the same genetic locus on chromosome 12 (17, 30, 32, 42).

A viral interference model has been proposed as the mechanism of Fv-4 restriction (15). A unique cell surface antigen immunologically related to major MuLV envelope glycoprotein gp70 has been identified in normal  $Fv-4^r$  mice (15, 48). This glycoprotein in serological assays exhibits characteristics of ecotropic MuLV gp70s (16). More recently, an MuLV env sequence present in mouse chromosomal DNA was linked to  $Fv-4^r$  resistance; this sequence was initially identified by reactivity with an ecotropic MuLV env probe (22). These results are consistent with the observation that Fv-4 resistance affects ecotropic MuLVs but not other host range types of MuLVs (47; C. A. Kozak, unpublished data). The interference model suggests that the Fv-4gene product (gp70) competes with ecotropic MuLV for cell surface receptors. To further describe this proviral locus and its regulation, we have molecularly cloned and characterized the endogenous MuLV env-related sequence associated with Fv-4 resistance.

# MATERIALS AND METHODS

Mice. BALB/c and AKR/J mice were obtained from T. Odaka, Institute of Medical Science, University of Tokyo,

Tokyo Japan, who also provided frozen livers from three partially Fv-4-congenic mice and FRG mice. Two of the congenic mice, BALB/c-Fv-4w'(Fu) and BALB/c-Fv-4w'(Hz), carry the resistance allele from wild mice (*Mus musculus molossinus*) originally trapped in different areas of Japan. Both mouse strains were inbred after six backcrosses to BALB/c (32; T. Odaka, unpublished data). A third congenic mouse strain, AKR-Fv-4'', carries the Fv-4'' gene of the FRG strain (41) on an AKR/J background and was also inbred after backcross generation 6 (T. Odaka, unpublished data).

**Preparation and analysis of DNA.** DNA was extracted from liver or thymus as previously described (4). Restricted DNA was fractionated by electrophoresis on 0.6% agarose gels, and cleavage products were transferred to nitrocellulose membranes for Southern blotting (40). Membranes were hybridized with <sup>32</sup>P-labeled probes in the presence of 50% formamide-20% dextran-sulfate, as previously described by Thomas (44). Hybridization was carried out for 12 to 24 h at 42°C (low stringency) or 55°C (high stringency). The membranes were washed four times, for 5 min each time, at room temperature in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-0.1% sodium dodecyl sulfate (SDS); 2 times, for 15 min each time, at 55°C in 0.1× SSC-0.1% SDS; and 4 times, for 5 to 20 min per time, at 55°C in 2× SSC.

Molecular cloning of the  $Fv-4^r$ -associated MuLV sequence. EcoRI-digested BALB/c- $Fv-4w^r$ (Fu) liver genomic DNA was ligated to EcoRI-cleaved  $\lambda$ gtWES arms (23), packaged in vitro, and propagated in Escherichia coli LE392. Cloned recombinant DNA was screened as previously described (3) with a <sup>32</sup>P-labeled ecotropic env-specific probe, pEc-B4 (4). A recombinant lambda phage ( $\lambda$ Fv4), containing a 5.2kilobase (kb) pEc-B4-reactive EcoRI fragment, was subcloned into pBR322 and was designated pFv4.

**DNA probes.** Ecotropic MuLV gp70-related sequences were identified with the ecotropic *env*-specific pEc-B4 probe (4) derived from the DNA clone of  $\lambda$ AKR623 (26). This probe specifically hybridized to endogenous ecotropic MuLV proviral DNAs at high stringency or to  $Fv-4^r$ -related

<sup>\*</sup> Corresponding author.

sequences under relaxed conditions. The pFv4<sub>env</sub> probe, which specifically hybridized to Fv-4-associated MuLV sequences at high stringency, was a 0.7-kb BamHI fragment, subcloned from the env region of clone pFv4 (see Fig. 1) and subsequently inserted into the BamHI site of pBR322. Other MuLV probes used included a 1.9-kb BamHI fragment derived from the gag region of AKV (28), a 1.4-kb XhoI-HpaI fragment isolated from C-terminal half of the AKV pol region (13), a 0.9-kb BamHI-XbaI fragment from the AKV gp70-p15(E) region (28), and a 0.6-kb KpnI fragment containing long-terminal-repeat (LTR) sequences from cloned Harvey sarcoma proviral DNA (11, 28). Probes were labeled ([ $\alpha$ -<sup>32</sup>P]ATP) by the nick-translation procedure of Maniatis et al. (27).

Restriction enzymes, gel electrophoresis, nucleotide sequencing, and data analysis by computer. Restriction enzymes were purchased from New England Biolabs, Beverly, Mass., or Boehringer Mannheim Biochemicals, Indianapolis, Ind., and used as suggested by the manufacturers. Restriction fragments were separated in horizontal agarose gels (0.8 to 2.0%) and purified by the glass bead procedure (45).

T4 polynucleotide kinase and calf intestinal alkaline phosphatase were purchased from P.L. Biochemicals, Milwaukee, Wis.  $[\gamma^{-32}P]ATP$  (3,000 Ci/mmol) from Amersham Corp., Arlington Heights, Ill., was used for the kinase reaction. DNA was sequenced by the partial-degradation method of Maxam and Gilbert (29). The computer programs of Queen and Korn (34), Wilbur and Lipman (46), and Hopp and Woods (14) were used for translation to amino acids, identification of restriction sites, determination of sequence homology, and generation of the hydrophilicity profiles of the deduced amino acid sequences.

Cellular RNA isolation and Northern blots. Cellular RNA was isolated from cells by the guanidine thiocyanate method and was purified by ultracentrifugation onto a CsCl cushion (8, 38). RNA preparations were fractionated on 1% agarose gels in the presence of 2.2 M formaldehyde (24), transferred to nitrocellulose filters (44), and hybridized to labeled DNA probes in 50% formamide at 45°C as previously described (35). The filters were washed twice (10 min) in  $2 \times SSC-0.1\%$  SDS at 50°C, twice (15 min) in  $0.1 \times SSC-0.1\%$  SDS at 50°C, and twice (5 min) in  $2 \times SSC$  at 50°C.

## RESULTS

Cloning of pFv4. BALB/c and AKR/J mice contain 1 and 3 endogenous ecotropic MuLV loci (4, 6), respectively, that react with the ecotropic env-specific probe pEc-B4 (4). The single endogenous ecotropic provirus present in the BALB/c mouse genome is located within a 23-kb EcoRI fragment. EcoRI digestion of genomic DNA from BALB/c mice congenic for Fv-4 generates the same 23-kb reactive fragment as well as an additional 5.2-kb fragment originally derived from the Fv-4<sup>r</sup> parent (22). This 5.2-kb EcoRI fragment was cloned from the EcoRI-digested liver DNA of a BALB/c mouse congenic for Fv-4 into  $\lambda$ gtWES phages. Of the  $1.3 \times 10^6$  recombinant phages screened, 2 contained a 5.2-kb insert that hybridized to the pEc-B4 probe. The insert present in one of these phage ( $\lambda$ Fv4A) was subsequently subcloned into the EcoRI site of pBR322 and designated pFv4.

**Restriction mapping and alignment of pFv4 sequences.** The restriction map of pFv4 is presented in Fig. 1 and compared with that of AKV623, a cloned ecotropic MuLV provirus (26). The positions of the *PstI*, *SmaI*, and *KpnI* sites in the





FIG. 1. Alignment of the cloned  $Fv-4^r$ -associated retroviral sequence pFv4 with the *pol*, *env*, and 3' LTR regions of AKV proviral DNA. The 0.7-kb *Bam*HI fragment (heavy-lined areas from map positions 2.1 to 2.8) is the *env*-specific probe sequence (pFv4<sub>env</sub>) subcloned from pFv4. Broken lines represent cellular DNA sequences. The nucleotide sequencing strategy is shown at the bottom. Abbreviations: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; K, *Kpn*I; P, *Pst*I; Pv, *Pvu*I; S, *Sma*I; Xb, *Xba*I; SA, consensus splice acceptor sequence site.

LTR of AKV623 and of the single XbaI and BamHI sites in the env gene at 3.7 and 2.8 map units, respectively, are identical to those present in pFv4 (Fig. 1). Identification of retroviral sequences present in pFv4 was determined by hybridizing a series of labeled MuLV probes (see above) to restricted preparations of pFv4 DNA. These experiments indicated that MuLV-related pol, env, and LTR sequences (located between map units 1.3 and 4.7 [Fig. 1]) were present in the pFv4 clone (data not shown), a result confirmed by nucleotide sequencing (see below).

The pFv4 env probe and its specificity. A number of reports have shown that different classes of MuLVs contain unique sequences in the N-terminal half of the gp70 coding region, whereas the C-terminal portion of gp70 as well as p15(E) is highly conserved (13, 19, 21, 25, 33, 36, 39). Aligned sequences show distinguishing differences between ecotropic and mink cell focus-forming (MCF) or xenotropic MuLVs in the 5' portion of gp70 (32, 35), which account for the differential specificities of the ecotropic pEc-B4 (4) and the xenotropic and MCF MuLV-reactive pXenv (3) probes. Differences between MCF and xenotropic gp70 MuLV sequences in the 5' half of gp70 coding regions are not extensive; however, within this segment several short runs of heterology exist (33, 36). These differences may be analogous to the short variable domains in gp85 of avian retroviruses which have recently been shown to be responsible for subgroup specificity (9).

Comparison of the *env* sequences of ecotropic MuLVs with those of pFv4 indicated that significant variability occurred within a 0.7-kb pFv4 segment defined by *Bam*HI restriction sites (positions 70 to 820 [Fig. 2]; Fig. 1). This *Bam*HI fragment was subcloned into pBR322 and was designated pFv4<sub>env</sub>. Genomic DNA from BALB/c-*Fv*-4*w*<sup>r</sup>(Fu) mice, containing the endogenous ecotropic BALB/c provirus as well as the *Fv*-4<sup>r</sup> determinant, was digested with *Pst*I and hybridized to either the pFv4<sub>env</sub> or the pEc-B4 ecotropic *env* probes. Under low-stringency conditions, the pEc-B4 probe hybridized more efficiently with the endogenous ecotropic proviral DNA of BALB/c mice (the 8.3-kb fragment) than with the *Fv*-4<sup>r</sup>-associated (4.2-kb fragment) retroviral sequence (Fig. 3A, lane 1). The reverse was true when the

Fv-4		A T G ME T	GAG Glu	AGT SER	CCA Pro	GCG Ala	TTC Phe	TCA SER	AAA Lys	CCC Pro	CTT Leu	AAA LYS	GAT ASP	AAG LYS	ACT THR	ATC ILE	AAA LYS	50 AAA LYS	GCT Ala	CTC LEV	CTA Leu	GGG GL Y	GTG Val	TTG LEU	Bai GGG GLY	ATC	CTA LEU	CTC LEU		
Akv Eco		••••	•••	••••	A Thr	A Thr	C LEU	•••	•••	•••	Т РНЕ	••••	A ASN	C GLN	GT. Val	. A . Asn	CCG Pro	TGG TRP	. GC GL Y	.C. Pro	·	ATT ILE	c	с.т *	CT. LEU	т *	c	GGA GL Y	GGG GL Y	GTC
NZB Xeno		••••	• • •	6 Gl y	T SER	•••	•••	••••	••••	••••	•••	•••	••••	•••	. T. Ile	. A . Asn	CCG PRO	TGG TRP	. GC GL Y	.C. Pro	•	ATA ILE	т •	А МЕТ	·	· · ·	Т.G	G.G	AGG	GCA
Eu-6			<b>6</b> 76		00																150									
A.L			VAL NH2	THR 9070	GLY	GLY	LEU	ALA	HIS	LYS	ASP	SER	PRO	HIS	LEU	ILE	TYR	ASN	LEU	ACC	TGG TRP	GAA Glu	GTA VAL	ACA Thr	AAT ASN	GGA GLY	GAA Glu	CAA GLN	GAA Glu	ACT Thr
×	ASN	PRO	`*'	*	LEU						ASN		*	*	GLN Bgi	VAL II	PHE	*	с *			•	• • •	т *	*	*	C ASP	.G. Arg	•	G #
Aeno	GLY	ALA	SER	VAL	GLN	ARG					•	•	*		.A. GLN	с *	.T. PHE	*	G.T VAL	T *	•	AG. Arg	·T *	с *	с *			.T. Leu	ATG MET	• • •
Fv-4	GTG	TGG	6CA	GTA	ACC	eec	AAC	CAC	ccc	TTG	TEE	ACT	TGG	TGG				ccc	GAC	стс	ACA	CCA	GAC	стс	TGT	ATG	25 CTG	6CC	CTA	CAT
Akv				A			азн т	н1S 	РКО Т	C	1RP		TRP	TRP				PR0	ASP	LEU	THR 	PRO	ASP 1 1 T	LEU (	CYS	ME T	LEU T	AL A	LEU C	м15 с
Xeno	* . ga	*	* 	.ст	. <u>.</u>	* .c.	* .c.]	ж тс.	.т.	* c	* 6	* e	# АТ.	*	GAC	ACC	ттс	* т	* A.A	*	* TAT	* ттт	•	* 	*	*	* T.A	*	*	٠
	61.1	GUN	146	ALA	ASM	<u>ALA</u>	THK	SER	LEU	•	6L Y	•	MET	THR	ASP	THR	PHE	•	LYS	*	TYR	PHE	*	*	•	ASP	*			
Fv-4	eec eec	CCA Pro	ACT Thr	CAT HIS	TGG TRP	GEC GEC Y	CTA Leu	GAC ASP	AAC Asn	CGA Arg	CCT Pro	CCG Pro	TAC Tyr	300 TCC SER	TC T SE R	SI CCC PRO	CCG PRO	GGG GL Y	CCC Pro	CCT Pro	TGT CYS	TGC Cys	TCG SER	GGA GLY	GAT ASP	AAG Lys	GGG GL Y	GCC AL A	GTG VAL	TCG SER
Akv	G *	e	T.C SER	Т ТУR	••••	••••	•••	 GL U	T.T TYR	G #	G Ala	·;T	.TT PHE	т #	C PRO	S • • • • *	•• I ••• *	••••	••••	• • • • •	c	т #	A	· *	AGC Ser	. GC SER	. AC ASP	T SER	AC. THR	C.A PRO
Xeno	.TA Val	GG. GLY	GAC ASP	T.C TYR	•••												GAT ASP	. AC ASP	· A	GAA Glu	CCC PRO	GAT ASP	ATT ILE	G	ASP	GG T GL Y	T.C CYS	CG. Arg	S ACT THR	mal C.C PRO
F 4	350																41	00												
FV-4	GLY	CYS	ALA	ARG	ASP	CYS	ASP	GLU	PRO	LEU	THR	SER	TAC	SER	CCT PRO	CGG ARG	TGC CYS	AAT	ACC THR	GCC ALA	TGG TRP	AAT	AGA Arg	A T G ME T	AAA Lys	CTG LEU	GCC ALA	CAG Gln	GTG Val	ACA Thr
ARV	•		SER	*		•	6LU	G *	•	¢.6 *	т *	•	•	A.T THR	с *	*	•	•	G #	*	•	с *	•	C.T LEU	G *	т. А *	T.T SER	A.A LYS	*	•
Xeno	*	G.A GLY	AGA Arg	*	AGG ARG	ACA Thr	AGA Arg	LEU	TAT TYR	GAC ASP																				
Fv-4	CAT	GCA	ccc	450	GAG	GGA	TTC	тат	STO	TEC	сст	666	тса	CAC			466	TCC	607	Pvu	I	5	05	667						101
Aku	HIS	ALA	PRO	LYS	GLU	GLY	PHE	TYR	VAL	CYS	PRO	GLY	SER	HIS	ARG	PRO	ARG	TRP	ALA	ARG	SER	CYS	GLY	GLY	PRO	GLU	ALA	TYR	TYR	CYS
Xeno	*	*	HIS	ASN	GL Y	•	*	*	т		c	т	PRO	*	*	*	* . TA	*		*	*	*				*	SER	PHE		•
							*	*			*		HIS	THR	VAL	*	ILE	GL Y				*	*	*	*	GLY	GLU	GLY	*	*
Fv-4	GC T	TCC SER	TGG	GGA GL Y	TGT	GAA	55 ACT THR	50 ACA THR	GEC	CGA	GCA	GCC	TGG	AAA		ACT	TCA	TCA	TGG	GAC	TAC	ATC	ACA	600 GTA	AGC	AAT	AAC	TTG	TCA	TCC
Akv	c	т		c	c		c					T				T.C	G	.:c								c	T	C.A	A.C	A
																	-	-	-	-	-	-	-	-	-		*	*	1.0.8	
Xeno	. GC GL Y		•••	· · ·	•••	G	c	· . T		. AG	••••	TAC	• • •	G	• • •	T.A	••••	• • •	• • •	• • •	CTA	· . T	T.C	С.Т	. AG	CGA	GGA	AAC	A. T	C.T
Xeno	. GC GL Y	AAA Lys	•	•	•	G *	c *	T *	• • •	. AG Gln	••••	TAC Tyr	•	G *	•	T.A SER	•	*	•	•	CTA Leu	• • • •	T.C SER	C.T LEU	.AG Lys	CGA Arg	GGA GLY	AAC ASN	A.T THR	C.T PRO
Xeno Fv-4	.GC GLY	AAA LYS CCG PRO	* CAG	•	•	G *	c *	T *	· A	. AG GLN	 * 650	TAC Tyr		G *		T.A SER	•	GCC	CCC PRO	 # AAAA LYS	CTA LEU GCC ALA	T * TGC CYS	T.C SER AAA LYS	C.T LEU AAT ASN	. AG LYS AAT ASN	CGA ARG GGC GLY	GGA GLY 70 TGG TRP	AAC ASN DO TGC CYS	A.T THR AAT ASN	C.T PRO CCC PRO
Xeno Fu-4 Aku	.GC Gly	AAA LYS CCG PRO GAC ASP	CAG GLN	•	•	G *	c *	<b>T</b>	<b>A</b>	. AG Gln	•••• *	TAC TYR	*	6 *	*	T.A Ser	•	GCC ALA A	CCC PRO A THR	AAAA LYS CC. PRO	CTA LEU GCC ALA .TA VAL	T * TGC CYS 	T.C SER AAA LYS 	C.T LEU AAT ASN GG. GLY	. AG LYS AAT ASN 	CGA ARG GGC GLY .AG GLU	GGA GLY TGG TRP 	AAC ASN TGC CYS	A.T THR AAT ASN C	C.T PRO CCC PRO T SER
Xeno Fv-4 Akv Xeno	.GC GLY AAG Lys	AAA LYS CCG PRO GAC ASP GAT ASP	CAG GLN  *	 # GGC GLY	 * CCC PRO	G * TGT CYS	C * TAT TYR	GAT	A * TCC SER	. AG GLN TCG SER	 * 650 GTC VAL	TAC TYR TCC SER	AGT	6 * 66C 6LY	GTC	T.A SER CAG GLN	GGT	GCC ALA A *	CCC PRO A THR A.A THR	AAAA LYS CC. PRO CCG PRO	GCC ALA .TA VAL	T # TGC CYS 	T.C SER LYS 	C.T LEU AAT ASN GG. GLY	. AG LYS AAT ASN  # GGG GLY	GGC GGC GLY .AG GLU T	GGA GLY TGG TRP  * C.A ARG	AAC ASN TGC CYS  *	A.T THR AAT ASN C *	C.T PRO CCC PRO T SER 
Xeno Fv-4 Akv Xeno Fv-4	.GC GLY AAG LYS CTA	AAA LYS CCG PRO GAC ASP GAT ASP	CAG GLN 	GGC GLY CGC	CCC PRO	TGT CYS	C * TAT TYR GGT	GAT ASP	TCC SER	.AG GLN TCG SER	 650 GTC VAL AGG	TAC TYR TCC SER GCT	AGT SER	6 * 66C 6LY 750	GTC VAL	T.A SER CAG GLN ACT	GGT GLY	GCC ALA A # 	CCC PRO A THR A.A THR	AAAA LYS CC. PRO CCG PRO	CTA LEU GCC ALA .TA VAL	T * TGC CYS 	T.C SER AAA LYS 	C.T LEU AAT ASN GG. GLY	.AG LYS AAT ASN  # GGG GLY CTA	CGA ARG GGC GLY .AG GLU T *	GGA GLY TGG TRP  * C.A ARG	AAC ASN TGC CYS  *	A.T THR AAT ASN C *	C.T PRO CCC PRO T SER 
Xeno Fv-4 Akv Xeno Fv-4 Akv	.GC GLY AAG LYS CTA LEU T	AAAA LYS CCG PRO GAC ASP GAT ASP GTC VAL	CAG GLN  * GTA VAL	GGC GLY CGC ARG G	CCC PRO TTT PHE C	TGT CYS ACG THR	TAT TYR GGT GLY A.C	GAT ASP CCA PRO TTT	TCC SER GGA GLY	AG GLN TCG SER AAG LYS	GTC VAL AGG ARG CA.	TAC TYR TCC SER GCT ALA C	AGT SER ACC THR	6 * 66C 6LY 750 TCC SER 	GTC VAL TGG TRP	T.A SER CAG GLN ACT THR GTC	GGT GLY ACA THR	GCC ALA A *  GGT GLY C	CCCC PRO A THR A.A THR CAT HIS	AAAA LYS CC. PRO CCG PRO GAA GLU TGG	CTA LEU GCC ALA .TA VAL TGG TRP	GGA GLY	T.C SER AAAA LYS  # CTG LEU T	C.T LEU AAT ASN GG. GLY CGC ARG	.AG LYS AAT ASN  GGG GLY CTA LEU 	CGA ARG GGC GLY .AG GLU T * TAC TYR	GGA GLY TGG TRP  * C.A ARG ATC ILE G	AAC ASN TGC CYS  *  *	A.T THR AAT ASN C *	C.T PRO CCC PRO T SER  # GGG GLY
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno	.GC GLY AAG LYS CTA LEU T *	AAAA LYS CCGG PRO GAC ASP GAC ASP GTC VAL ACT THR	CAG GLN * * GTA * *	GGCC GLY CGCC ARG ECC GAA	* CCCC PRO TTTT PHE C * RI	G * CYS ACG THR 	C * TAT TYR GGT GLY A.C SER .AC	T * GAT ASP CCCA PRO TTT PHE G.G	A # TCC SER GGA GLY 	.AG GLN TCG SER AAG LYS A *	GTC VAL AGG ARG GLN .A.	TAC TYR TCC SER GCT ALA C	AGT SER ACC THR	G * GGC GLY 750 TCC SER  *	GTC VAL TGG TRP	T.A SER CAG GLN ACT THR GTC VAL GA.	GGT GLY ACA THR 	GCCC ALA A * GGTY C *	* CCCC PRO A THR A.A THR CAT HIS  *	AAAA LYS CCC. PRO CCG PRO GAAA GLU TGGGU	CTA LEU GCC ALA .TA VAL TGG TRP 	T * TGC CYS  *	T.C SER AAAA LYS  * CTG LEU T *	C.T LEU AAAT ASN GG. GLY CGC ARG 	.AG LYS AAT ASN  # GGG GLY CTA LEU  *	CGA ARG GGC GLY .AG GLU T * TAC TYR 	GGA GLY TGG TRP  * C.A ARG ATC ILE G VAL	AACC ASN TGC CYS  *  * TCT SER  *	A.T THR AAT ASN C * C	C.T PRO CCC PRO T SER  # GGGG GLY A *
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno	.GC GLY AAG LYS CTA LEU T * G	AAAA LYS CCCG PRO GAC ASP GAC ASP GAC VAL ACT THR 	* GLN * * * * * * *	GGC GLY CGC ARG Ec GAA GLU	CCCC PRO TTTT PHE C *	G * TGT CYS ACG THR  *	C * TAT TYR GGT SER .AC ASP	T * GAT ASP CCA PRO TTT PHE G.G ALA	A * TCC SER GGA GLY  *	. AG GLN TCG SER AAG LYS A # .GA	GTC VAL AGG ARG CA. LYS	TAC TYR TCC SER GCT ALA C *	AGT SER ACC THR *	G * GGC GLY 750 TCC SER  *	•••• •• •• •• •• •• ••	T.A SER CAG GLN ACT THR GTC VAL GA. ASP	GGT GLY ACA THR G.C ALA	GCC ALA A * GGTY C PRO	CCCC PRO A THR A.AA THR CATT HIS  # A.AA LYS	AAAA LYS CC. PRO CCG PRO GAAA GLU TGG TRP .TT VAL	CTALLEU GCC ALA .TAVAL TGG TRP 	T TGC CYS  # GGA GLY 	T.C SER AAAA LYS  # CTG LEU T #	C.T LEU AAT ASN GG. GLY CGC ARG  *	.AG LYS AAT ASN  # GGG GLY CTA LEU  * C	CGA ARG GGC GLY .AG GLU T * TAC TYR *	GGA GLY 7 TGG TRP  * C.A ARG ATC ILE G VAL CGA	AACC ASN 00 TGC CYS  * * TCT SER  *  *	A.T THR AAT ASN C * C *	C.T PRO CCCC PRO T SER  # GGGG GLY  *
Xeno Fu-4 Aku Xeno Fu-4 Aku Xeno Fu-4	.GC GLY AAG LYS CTA LEU T # G CAT HIS	AAAA LYS CCG PRO GAC ASP GAT ASP GAC THR 	CAGG GLN * GLN * * * * * * * * * * * *	GGC GGC GGC CGC ARG GAA GGU GGA GGC GGC GGC GGC	CCCC PRO TTTT PHE C * CTCLEU	G * * * * * * * * *	C * TAT TYR GGT GLY A.C SER .AC SER TTT	T * GAT ASP CCA PRO TTT PHE G.G ALA Ba GGG GLY	A * TCC SER GGA GLY  * * *	.AG GLN TCG SER AAG LYS A * .GA ARG	GTC VAL AGG CA. GLN .A. LYS CTA LEU	TAC TYR TCC SER GCT ALA C # AGA	AGT AGT ACC THR *	G * GGC GLY 750 TCC SER  * AGA ACA THR	GTC VAL TGG TRP  *  *	T.A SER CAG GLN ACT THR GTC VAL GA. ASP I CTG LEU	GGT GGT GLY ACA THR G.C ALA SA GGLY	GCC ALA A * GGT GLY C PRO	CCCC PRO A THR A.A. THR A.A. LYS AAGG	AAAA LYS CC. PRO CCGG PRO GAAA GLU TGGGLU TGGGLU TGGC VAL	CTALEU GCC ALA .TA VAL TGG TRP  *	T * TGC CYS  * GGA GLY  *	T.C SER AAAA LYS  * CTG LEU T * GGGGGLY	C.T LEU AAT ASN GG. GLY CGC ARG  * A.A *	.AG LYS AAT ASN  GGG GLY CTA LEU  * *	CGA ARG GLV .AG GLU T * TAC TYR  * *	GGA GLY TGG TRP  * C.A ARG G VAL CGA ARG GTC VAL	AACC ASN DD TGC CYS  * *  * *  * * 	A.T THR AAT ASN C * C * ACA THR TCA SER	C.T PRO CCCC PRO T SER  # GGGG GLY A #  #
Xeno Fu-4 Aku Xeno Fu-4 Aku Xeno Fu-4 Aku	.GC GLY AAG LYS CTA LEU T * * 800 CAT HIS 	AAAALYS CCCG PRO GAC ASP GAC ASP GAC VAL ACHT THR  *	CAGN GLN  * GTA VAL A.C ILE T LEU CCA PRO  *	GGC GGLY CGCC ARG GAA GGLU GGA GGLY SGGG GLY 	CCCC PRO TTTT PHE C * CTCC LEU *	G TGT CYS ACG THR  # T THR 	C * TAT TYR GGTY A.C SER .AC SER .AC *	GAT * GAT ASP CCA ASP PRO TTT PHE G.G ALA Ba GGLY Ba 	A * TCC SER GGA GLY  * MHI ATC ILE mHI  *	AGG GLN TCG SER AAG LYS A # A # CGG ARG A #	 # 650 GTC VAL AGG ARG GLN .A. LYS CTA LEU T	TAC TYR TCC SER GCT ALA C * C * AGA ARG .LYS	AGT AGT SER ACC THR  *	G # GGC GLY 750 SER  # AG. # BA ACA THR 	GTC VAL TGG TRP  #  GAT ASP C	T.A SER CAG GLN ACT THR GTAL GA. ASP I CTG LEU TC. SER	ACA GGUY ACA THR G.C ALA SGA GGAY G	GCCC ALA A *  GGTY CC PRO SO CCTT PRO SO C *	CCCC PRO A THR A.A. THR HIS  # A.A. LYS AAGG ARG B 3 I C *	AAAA LYS CC. PRO CCG PRO GAAA GLU TGG TRP .TT VAL GTC VAL	CTALEU GCCCALATA VAL TGGG TRP * CCGG PRO	T * TGC CYS  * GGA GLY  * *	T.C SER AAAA LYS * CTG LEU T * * GGGG GLY *	C.T.LEU AAT ASN GG.GLY CGCC ARG  # A.A # CCCA PRO  #	.AG LYS AAT ASN  GGG GLY CTA LEU  * C * AAT 	CGAARG GGC GLY .AG GLU T * TAC TYR  * * CCT PRO C	GGAA GLY TGGGT TRP  * C.A AARG C.A AARG G YAL CGA ARG GTC YAL  *	AACC ASN TGC CYS  #  * TCT SER  *  * TTG LEU  *	A.T THR AAT ASN C * C * ACA THR TCA SER 	C.T PRO CCCC PRO T SER  # GGGG GLY A #  # GAT ASP C #
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno	.GC GLY AAG LYS CTA LEU T * * 800 CAT HIS  * GCC ALA	AAAALYS CCCG PRO GACP GACA ASP GACT VAL ACT THR  # GACG ASP VAL  *	CCAA GLN  # GLN  # GTA VAL T LEU CCAA PRO  # G	GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC	CCCC PRO TTTT PHE .C CTCU CTCU  AC.R	G * TGT CYS ACG THR *  * ACC THR *  *	C * TAT TYR GGLY A.CC SER A.SP TTT PHE  *	T * GATP CCA PRO TTT PHE G.G ALA Ba GGG GLY Ba  *	A * TCCC SER GGAY  * * * * * * * * * * * *	AGG GLN TCG SER AAGG LYS A * A * A * A GLN A * A GLN A GLN A GLN A GLN A GLN 	GTC GTC VAL AGG ARG GLN LVS CTA LEU T * * .GC ARG	TAC TYR TCC SER GCT ALA C # C # AGA ARG ARG LYS CAG GLN	AGT AGT AGT AGT AGT AGT AGT A ILE G.C VAL	G # GGC GGLY 7500 TCC SER # AGA THR # B ACA THR  #	GTC VAL TGG TRP  *  *  A A.SN	T.A SER CAG GLN ACT THR GTC VAL GA. ASP I CTG LEU TC. SER G.A VAL	GGT GLY ACAA THR GGCA ALA GGLY GLY 	GCC GALA A GGTY GGLY C CCCC PRO SC CCT PRO SC CCC CC CC CC CC CC CC CC CC CC CC CC	C.CC PRO A THR A.A. A.A. A.A. A.A. A.A. A.A. C.T. C.C. X	AAAA LYS CC. PRO GAAA GLU TGG TRP .TT YAL GTCC YAL 	CTALEU GCC ALA .TA YAL TGG TRP  * CCG PRO A	T TGCC CYS  # GGA GLY  # ATA ILE  #  *	T.C SER AAAA LVS  * CTG LEU T * GGGG GLY  *	C.TLEU AATN ASN GG.GLY CGCC ARG A.A * * * * * * *	AAT ASN GGGGCLY CTA LEU  * * AAT ASN C *	CGAAGG GGCY .AGGLU T * TACCTTYR T * CCTTPRO C *	GGAA GLY TGGG TRP C.A ARG ARG G VAL CGAA ARG GTC VAL  *	AACC ASN DO TGC CVS * TCT SER * TTG LEU *	A.T THR AAAT AAAT AAAT ACA THR ACA TCA SER A.T THR	C.T PRO CCCC PRO T SER  # GGGG GLY  # GAT ASP C #
Xeno Fu-4 Aku Xeno Fu-4 Aku Xeno Fu-4 Aku Xeno	.GC GLY AAGG LYS CTA LEU T # G CAT  # G CAT  # G CAL *  CAL *  `  C `  `  `  `  ` 	AAA LYS CCCG PRO GACP GACP GACP GACP VAL ACHR  # GACP 	CAGG GLN  * GTA VAL A.C ILE T LEU CCA PRO  * G	GGC GGC GGC CGC ARG GC GAA GLU GGG GLY  *  *	CCCC PRO TTTT PHE C * CTCC LEU  *	G # TGT CYS ACG THR  #  # ACCC THR  ECGG ARG	C * TAT TYR GGLY A.CC SER .ACP PHE  *	T * GAT ASP CCAA PRO TTT PHE G.GA ALA Ba GGG GLY Ba  * TCT SER	A * TCC SER GGAY  * * MHI ATC ILE LEU	AGG GLN TCG SER AAGG LYS A * * A * * A * * A * CGG ARG A * CGG ARG A * CGG THR	GTC VAL AGG ARG GLN .A. LYS CTA LEU T * .GC ARG	TAC TYR TCC SER GCT ALA C # AGA ARG ARG LYS GLN	AGT AGT SER ACC THR  # ATA ILE  G.C VAL	G # GGC GLY 750 TCC SER  # ACA THR  # CTC LEU	GTC VAL TGG TRP * * * * *	T.A SER CAG GLN ACT THR GTC GA. ASP I CTG LEU. SER G.A VAL	GGT GLY ACAA THR G.C ALA BGCA GCA GCA SCA SCA SCA SCA SCA SCA SCA SCA SCA S	GCCC ALA  GGTY GLY C PRO SO CCTO PRO SO C *  *	CCCC PRO A THR A.A. THR A.A. THR A.A. A.A. A.A. CATS  B.A. C.C. C.C. C.C.	AAAA LYS CC. PRO CCG PRO GAAA GLU TGGP TRP 	CTA LEU GCC ALA .TA YAL TGG TRP  * CCG PRO  * *	T * TGC CYS  * GGA GLY  * ATA ILE  * 	T.C SER AAAA LYS  * CTG LEU T * GGGG GLY  *	C.T LEU AAAT ASN GG. GLY CGC ARG A.A W CCAA PRO  X T X	AAAT AAAT GGGGGLY * AAAT AASN C * AAAT ASN *	CGAARG GGC GLY TAC TYR  * CCTTYR  * CCTT PRO C * 	GGA GLY 710 TRP * C.A ARG ATC ILE G VAL CGA ARG GTC VAL * *	AACC ASN DO TGC CYS  * * TCT SER  * * TTG LEU  * A.CC ILE	A.T THR AAT ASN C * ACA THR TCA SER  A.T THR	C.T PRO CCCC PRO T SER GGLY A * GAT ASP C *
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Fv-4	.GC GLY AAGG LYS CTA LEU T * * * * * * * * * * * * * * * * *	AAAALYS CCGGPRO GACCASP GATASP GATASP GACTATHR 	CCCG CCG CCG CCCG CCCG CCCG CCCG CCCG	GGC GLY CGCC ARG GLY GGG GLY  * .T. VAL	CCCC PRO TTTT PHE C RI C CTCU LEU  AC THR TCCC	G # TGT CYS ACG THR  # ACC THR  # ACC THR  # CGG ARG	C * TAT TYR GGTY A.CC SER ASP TTT PHE  * CCT PRO	T * GAT PRO FRO FRO GLA Ba GGGY Ba  * TCT SER GTA	A * TCC SER GGA  * MHI  * T.G LEU	AGG GLN TCG SER AAGG LYS A * GA ARG ARG ARG THR	GTC GTC VAL AGG ARG GLN .A. LUYS CTA LEU T * .GC ARG	TAC TYR GCT ALA C # C # AGA ARG GLN CAG GLN	AGT SER ACC THR  # ATA ILE T # G.C VAL	G # GGC GLY 750 TCC SER  # AG. # BA ACA # CTC LEU	GTC VAL TGG TRP  *  SGAT ASP C * A ASN CCT PRO	T.A SER CAG GLN ACT THR GTC VAL GA. SER G.A VAL GCCC ALA	GGT GLY ACA THR G.C ALA GGA GGA GGA SGA ARG	GCCC ALA A GGTY C CCCC PRO SO CCCT PRO SO CCCT PRO	CCCC PRO A THR A.A.A THR A.A.A THR A.A.A THR A.A.A CATS  * C.CC * C.CC PRO	AAAA LYS CC. PRO CCG PRO GAAA GLU TGG PRO GAAA SLU STC VAL  #  #	CTA LEU GCC ALA .TA VAL TGG TRP # PRO #		T.C. SER AAAA LYS  * CTG LEU T * GGGGGLY  * TCA	C.T LEU AAT AASN GG. GLY CGCC ARG GLY CGCC ARG  * * * * * *	AG LYS AAT ASN  # GGG GLY  #  * * * * * * * * * * * * * * *	CGAAGGCLY GGLY .AGGLU .T TYR TYR  * * * * * * *	GGAA GLY TGGC TRP * C.A ARG C.A ARG G.A. VAL C.A ARG G.A. VAL C.A * * * * * * * * * * * * * * * * * * *	AACC ASN DO TGC CYS  *  * TTCT SER *  * TTG LEU  * A.CC THR	A.T THR AAT ASN C * C * ACA THR TCA SER  * A.T THR GGLY	C.T PRO CCCC PRO T GGGG GLY  GGGG SLY  GGGG M SL C C C C C C C C C C C C C C C C C C
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv	.GC GLY AAG LYS CTA LEU T * * SOO CATS  * * GCC ALA CAG GLN  *	AAAALYS CCCG PRO GACP GACP GACP GACP GACP VAL ACTR H GACP  # CGAC ASP  # CCGG ASP CVAL ************************************	CCGG CCAG GLN  * * * * * * * * * * * * * * * * *	GGCY GGCY CGC GAAG GGCY GGGG CCCC CCCC C	CCCC PRO TTTT PHE * CTCC LEU * AC. THR TCCC SER *	GGG ACCG THR  ILE CCGG ARG CCGG ARG	C TATT TYR GGTY A.C SER .ACP PHE  * CCT PRD  *	T # GATP PRO TTT PHE GGG GLY Ba  SER GTA VAL	A # TCCC SER GGAY  # MI ATCC ILE LEU	.AG GLN TCG SER AAG LYS .A ARG CGGG ARG .A ARG THR	GTC VAL AGG CA. GLN LYS CTAU T " " GC ARG	TAC TYR TCC SER ALA C # C # AGAG ARG CAG CLN	AGT SER ACC THR  # ATA ILE T # G.C VAL	G # GGCY 7500 TCCC SER  # AGA ACA THR  # CTC LEU	GTC VAL TGG VAL TGG TRP  *  *  *  *  *  *  *  *  *  *  *  *	T.A SER CAG GLN ACTT THR GAC GA. SER GA. SER VAL GCCC ALA A THR	GGT GGT GGLY ACA THR G.C ALA SGCA GGLY  # AGA ARG  *	GCCC ALA A GGTY C CCCC PRO SO CCCC PRO SO CCCC CCC SO CCCC CCC SO CCCC CCC SO CCCC CCCC CCCC CCCC CCCC CCCC CCCC CCCC	CCCC PRO A THR A.AA THR A.AA LYS AGGG AGGG C.CC CPRO 	AAAAALLYS CC PRO GGLU TGG GLU TGG GLU TGG CLU TGG CLU CCCG CCC PRO CCCG CCC CCCG CCC CCCG CCCC CCCC CCCC	CTALEU GCCCALA .TA YAL TGG TRP # CCGG PRO #	T TGC CYS  # GGA GLY  # ATA ILE  #	T.C SER AAAA LYS  # CTGU T # CTGU T # GGGG GLY  # TCA SER  #	C.TLEU AATAASN GG.GLY CGCCAGG A.A.W CCCAO  X AACCAASN 	.AG LYS AAT ASN  GGGG GLY CTA LEU CTA LEU  * * * * * * * * * * * * * * * * *	CGAA ARG GLU T * TAC TYR  * CCTT PRO C * *	GGAA GLY TGGGT TRP C.A ARG ARG ARG GTC VAL CGAA ARG GTC VAL  * * * * * * * * * * * * * * * * *	AACC ASN DD TGC CYS  * * * * * * * * * * * * * * * * *	A.T THR AAT ASN C # C # ACA THR TCA ATHR GGGG GGLY .A. GLU	C.T PRO CCCC PRO T GGGG GLY  # CGGG GLY  # C CCC * * * * * * * * * * * * * * * *
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno	.GC GLY AAG LYS CTA LEU T * * * * * * * * * * * * * * * * *	AAAA LYS CCGG PRO GAC ASP GAC ASP GAC ACT THR 	CCCG CCG CCG CCG CCG CCG CCG CCG CCG CC	GGCC GGCY GGCY GGCY GGCY GGCY GGCY GGCY	CCCC PRO TTTT PHE C * CTCC EUU  * AC SER  *	G TGT TGT TGT CVS ACG THR  # ACC CVS ACG ACG ACG ACG ACG ACG ACG ACG	C * TAT TYR GGTY A.C SER .AC SER  *  *  * CCT *	T # GAT PRO FRO FRO GLA Bas GGLY Bas * * TCT SER GTA VAL G	A * TCCC SER GGAY  * * T.G LEU CAG GLN	AGG GLN TCGG SER AAGG ARG CGGG ARG ARG THR ACC THR	GTC VAL AGG VAL AGG ARG GLA. LYS CTA LEU T * .GC ARG ARG	TAC TYR TCC SER GCTT ALA C # C # AGA ARG GLN CLEU	AGT AGT SER ACC THR  # ATA ILE  # G.C VAL	G # GGC GGLY 750 SER  # AG. # AG. # CTC LEU AGA AGA AGG ARG	GTC VAL TGG TRP  #  #  ASN CCTT PRO  #	T.A SER CAG GLN ACT THR GA. ASP I CTGU GA. SER VAL GCCC A.A THR C.T PRO	GGT GLY ACA THR G.CA SGA GGA CAT HIS	GCCC ALA A GGTY GGLY CCCC CCCC CCCC CCCC CCCC CCCC CCC	CCCC PRO A.A. THR A.A. THR A.A. THR A.A. A.A. C.T. C.C. C.C. C.C. C.C. C.C.	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	CTALEU GCCCALA .TAYAL TGGGTRP  # CCGGPRO  # C # SSO CCT T.A SER	  GGGA GLY  # ATA ILE  # GGCC GLY	T.C. SER AAAA LYS  # CTGGLEU T #  # GGGGY  # TCAA SER  # TCAA	C.TLEU AATAASN GG.GLY CGCCARG ARG A.A. * * CCAASN CCAASN * * * * * *	AGG LYS AAT ASN  GGG GLY CTA LEU  *  *  *  * TCT SER  *	CGAARG GGC GLY GGLY TYR TACC TYR  CPRO C *  *	GGAA GLY TGGGC TRP C.A ARG ARG GVAL CGAA ARG GTC CVAL  * * CCCC PRO STA VAL	AACC AASN DO TGCC CYS. * * * * * * * * * * * * * * * * * * *	A.T THR AAT ASN C * C * C * C * C * C * C * C * C * C	C.T PRO CCCC PRO SER  SER  GGGG GGGY A *  GGGG GGGY A *  B GGT ASP C * C * C * C
Xeno       Fv-4       Akv       Xeno	.GC GLY AAGG LYS CTA LEU T # 8000 CAT HIS  GCC ALA CAG GLN  ARG CAC	AAAA LYS CCGG PRO GACP ASP GACP ASP GACP CVAL ACT THR  * CGAA ASP CVAL CCG CO CASP CVAL CCG CASP CVAL CCG CASP CO CASP CO CASP CO CO CASP CO CO CO CASP CO CO CASP CO CO CO CASP CO CO CO CO CO CO CO CO CO CO CO CASP CO CO CO CO CO CO CO CO CO CO CO CO CO	CAGG GLN         	GGCC GGC GGC GGC GGC GGC GGC GGC GGC GCC CCC CCC GCCC GCCC	CCCC PRO TTTT PHE C CTCU  AC THR TCCC SER 	G TGTT CYS ACCG THR  THR  THR  CGGG ARG GLN CCA	C * TATT GGTY A.C SER .ACP TTTT * C * CCTT PHE  *  C *    	T * GAT PRO FOR GLGA GGGA SER CCCA ASP CCCA ASP CCCA ASP CCCA ASP CCCA ASP CCCA ASP CCCA ASP CCCA ASP CCCA SEC SEC CCA SEC CCA SEC CCA SEC CCA SEC CCA SEC SEC CCA SEC SEC SEC SEC SEC SEC SEC SEC SEC SEC	A * TCCC SER GGA GLY  * MHI ATCC LEU CAG GLN CCCC	AGG GLN TCG SER AAGG SER AAG ARG ARG ARG ARG ARG THR ATC ILC	GTC VAL AGGG CA. LYS CTA LEU  ARG ARG ARG CTA LCTA CTA CTA CTA CTA CTA CTA CTA CTA CTA	TAC TYR TCC SER GCTAALAC # C # AGAAARG LYS CAGGLN CTCC LEU	AGT SER ACC THR  * ATA ATA G.C VAL CCCC PRO	G # GGCY 750 TCCC SER  AG. * BACA THR  CTC LEU AGA AGG AGG CAA	GTC VAL TGG TRP GAT ASP A ASN CCT PRO C *	T.A SER CAG GLN ACTT THR GTC VAL GA.A SP I CTG LEU TC. SER G.A VAL GCC ALA A THR C.TC CTG C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.	GGT GLY ACAA THR G.C ALA BGGAY  GGAY  AGA ARG CAT HIS ACA	GCCCA ALAA GGCCA ALAA GGLY C GGLY C GGCC GGCC GGCC GGCCA GGCCA GGCCA CCCCO CCCCCO CCCCO	CCCC PRO A.A. THR A.A.A THR HIS A.G.G.C.C. PRO 	AAAA LVS CCCG PRO GAAAG GLU TGGGPRO GAAA VAL  FRO  GAC	CTALEU GCCAALA .TAGGCCAALA .TAGTRP   CCCGCA   T.A. SER CCGA		T.C. SER AAAA LYS  CTGUU T  GGGG GLY  TCA SER  A.G GAR A.G	C.TLLEU AATAASN GG.Y CGCCAARGG GLY CGCCAARG A.A.A. A.A. A.A. CCCAA PRO  GT. UD50	AGG LYS AAT ASN * GGG GLY CTA LEU  *  * *  * * *  * * * * *	CGAAG GGU GLU T TAC TYR  CCTT PRO C * ACC THR   GGU AGU AGU AGU AGU AGU AGU AGU AGU AGU	GGAA GLY TGGGATRP C.A.GARG ATCCAARG GT.CAARG GT.CAARG CCCCAARG CCCCCCAARG CCCCCAARG CCCCCAARG CCCCCCAARG CCCCCCAARG CCCCCCAARG CCCCCCCAARG CCCCCCCAARG CCCCCCCCAARG CCCCCCCCAARG CCCCCCCCCC	AACC ASSN DO TGCC CYSS  * TCTT SER  * C * C * TTGG SEC  *  * C *  *  *  *  *  *	A.T THR AAAT AASN C * C * ACAA THR ACAA THR GGLY  CAG	C.T PRO CCCC CPRO T SER  GGGG GLY    GGGG GLY       
Xeno       Fv-4       Akv       Xeno       Fv-4	GCC GLY AAGG LYS CTA LEU T # BOO CAT CAT ARG ALA ARG GLN  CCC CPRO	AAAA LYS CCCGPRO GACPRO GACPRO GACPRO GACPRO CASP GACT XACT THR  # CCGA AARG  # CCGA CCGA CCCC CCCC CASP CONT CONT CONT CONT CONT CONT CONT CONT	CAGG GLN. * * GTA VAL CAG VAL T EU VAL CCAG PRO * * * * * * * * * * * * * * * * * * *	GGCC GGLY CGCC GARG GGLY GGAA GGLU GGGG GLY VAL S000 CCCC PRO CCCC PRO CCCC CCCC CCCC CCCC C	CCCCC PRO TTTT PHE C RO CTCU LCU CCCC RO CTCU THR CCCCC PRO	G * * * * * * * * * * * * * * * * * *	C TATT TYR GGLY A.CC SER  ASP TTT PHE  CCT PRO       	T * GATP GATP GATP GALA Ba GGGY Ba  CCCA Ba GGGY Ba  CCCA Ba GGGY Ba  CCCA CCCCA CCCA CCCA CCCA CCCA CCCCA CCCCA CCCCA CCCCA CCCA CC	A * TCCC SER GGAA GLY * * * * TLG LEU CAG GLN CCCC CPRD	AGG GLN TCG SER AAGS AAG AAG AAG AAG AAG AAG THR ATC ILE CCT TO RO	GTC VAL AGGG CGLN .A. LVS CTA LUS .T ARG ARG ARG ARG ARG ALS 	TACCTYR TCCCSER GCTALA C * C * AGAGAAAGGLN CTCCLEU TCCCSER G.G	AGT SER ACCC THR ATA ILE T # G.CC VAL CCCC PRO CACC	G GGCY GGLY 7500 SER  BA ACA BA ACA AGA AGG AGG AGG AGG AGG AG	GTCC VAL TGG TRP  #  GATP C # C # C # C #  ASN CCTT TRP  #  ACCC THR	T.A SER CAGGLN ACTT THR GTC TGU GAA.P ICTGU GCCC ALA.A.THR C.T PRO TCCC SER	GGTY GLY ACA THR GCC ALA AGA GGAY  GCC ALA AARG AACA THR ACA THR	GCLA A. A * * GGLY CCTO PROSC * * GGLY CCTO PROSC * * GGLY CPRO SC * * GGLY		AAAAS L CC. PRO GAAU TGG PRO GAAU TGG TTR TT TT T CCCG PRO 	CTALLU GCCA A.LA TGG PRO  TRP PRO  TRP PRO  TSC CCCG CCCT PRO  T.A SCCCA AGA.GA.	T TGC CYS  GGA GLY  # ATA ILE  # GGC GLY GATP A.C	T.C. SER AAAA LYS  GGGQLY  TCAA THR AAGGC.A	C.TLEU AATN GGLY CGCC ARG GLY CGCC ARG CCA CCA CCA CCA CCA CCA CCA CCA CCA CC	.AG LYS AAT ASN  GGGGGLY     TCT SER   	GGA GGLUT * TACC GLUT * TACC T *T *	GGA Y TGGG C.A.G. ARG ATCC C.A.G. AILE C.A.G. VAL C.G.A.G. VAL C.G.A.G. C.C.C. PRO GTA TLEU C.C.A.G. C.C.C. C.C.C. C.A.G. C.C.C. C.A.G. C.G.G. C.A.G. C.G.G. C.A.G. C.G.G. C.A.G. C.G.C. C.G.G. C.G.C.C. C.G.C. C.G.C. C.G.C. C.G.C. C.G.C. C.G.C.C. C.G.C.C. C.G.C.C. C.G.C.C. C.G.C.C. C.G.C.C. C.G.C.C.C. C.G.C.C.C.C	AACC ASSN DO TGCS  * TCTT SER  * TTG LEU  A.CC TTG ILE ACG C.T.T PRO GTCL 	A.T TR AAAT AAAT AAAT AAAT AAAT AAAT AAAT	C.T PRO CCCC CPRO T SER  GGGG GLY A # C # ACCC # A GGT ACCC # A C # C # C
Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno Fv-4 Akv Xeno	GCC GLY CTAAG LYS CTA LEU CTA LEU CAT HIS CAT ALA ARG GLN  CAG GLN  CCC CAT CAC GLY  CCC CAT CCC CAT CCCC CAT CCCC CCCCCCCCC	AAAA LYS CCCGPRO GACPRO GACPRO GACPAL ACTTHR 	CAG GLN  * * * * * * * * * * * * * * * * *	GGCC GGLY CGCC ARG GLY GGA GGLY  YAL 9000 CCCC PRD T A  CTC LEU  TCT	# CCCC PRO TTTT PHE * CTCU * CTCU * CTCU * CTCU * CTCU * CTCU * CTCU * CTCU * CTCU * CTCU * CTCU * * * * * * * * * * * * * *	G TGT CYS ACG THR  THR  ACC THR  ACC THR  ACG GLN CCGG GLN CCGA CCA PRD	C TATT TYR GGLY A.CC SER C * C * C * C	T * GATP GATP GATP GALA BGGG GGLY CCCAO C		AGG GLN TCG SER AAGG ARG ARG ARG ARG ARG ARG THR ATCC THR ATCC THR ATCC THR ATCC THR ATCC SCT TCG SCT SCT SCT SCT SCT SCT SCT SCT SCT SCT	GTC VAL AGG CA.N LYS CTA LUYS CTA LUYS T * GC ARG ARG ARG CA.N PRO	TACC TYR GGLTAC # AGA GLN CLCU CCC SER G.G ALA	AGT SER ACC THR  # ATA ILE T # G.CC VAL CCCC PRO	G GGCY GGCY AGA AGA AGA AGA AGA AGG AGG AG	GTC VAL TGG TRP JIIT ASP  ASN  ASN  CCT THR ACC THR	T.A. SER CAG GLN ACTT THR GA.A ASP I CTG LEU TC.R G.A.A THR TPRO TCC SER	GG GG L	GCCA ALA A.A. A.A. A.A.A.A.A.A.A.A.A.A.A.A.	CCCC PRO A THR A.A.A THR A.A.A THR A.A.A CATS HIS C.C C C C C C C C C C C C C C C C C C	AAAAS CC PRO GAAAUYS CC.G PRO GAAU TGGPU GAAU TGGPU GAAU CCGO CPRO CCG GAAU TGGPU CCGO CPRO CCGO CPRO CCGO CPRO CCGO CPRO CCGO CPRO CCGO CPRO CCGO CPRO CCGO CPRO CCGO CPRO CCGO CCGO CPRO CCGO CCGO CCGO CCGO CCGO CCGO CCGO CC	CTALLEU GCCA .TA TGG PRO CCCCO  T.A. SER GALA .TA SER CCCCO CCCO CCCO CCCO CCCO CCCO CCCO C	T TGCS CCVS  GGLY      GGCC GLY   	T.C. SER AAAS L.V.S. CTGUUT GGLY T.A. GGLY T.A. GGGQLY T.A. A.GGGCY T.A. A.GGCCA A.G. A.G. A.G. A.G. A.G. A.G	C.TULEU AATN GG.Y GGLY CGCGARG 	.AG LYS AAT ASN  GGGG GLY  * *  * * * * * * * * * * * * *	CGAA GGC GLU GLU  TAC CT PRO  C C C C C C C C C C C C C C C C C	GGA A GLY TGGP C.A A A RG A RG A RG A RG A RG A RG A RG A	AACC AASH DDTGCS  * TTCT SER  * TTGU SER  * TTGU SER  * TTGU SER  *	A.T THR AAATN AAATN	C.T PRO CCCC PRO T SER  GGGG GGGY A *  GATP C * * C * * C * * C * *  GGGA GGA GGLY *
Хело Fv-4 Akv Xело Fv-4 Akv Xело Fv-4 Akv Xело Fv-4 Akv Xело Fv-4 Akv Xело Fv-4 Xело	GCC GLY CTA LEU CTA LEU CAT HIS CAT HIS CAT ALA CAG GLN  CAG CAT CALA CAG CAT CALA CAG CAT CALA CAG CAT CAT CALA CALA CALA CALA CALA CALA C	AAAA LYS CCCG PRO GACP GACP GACP GAC ASP GAC ASP 	CAGGUN CAGUN CGUN CCAGUN CCAGUN COLO CCAGUN COLO CCAGUN COLO CCAGUN COLO CCAGUN	GGCY GGCY GGCY GGCY GGCY GGCY GGCY GGCY	CCCC PRO TTTTE 	G TGTS CYS ACGG THR  * ACG THR  * ACG CARG GLN CCAA GLN	C TATT GGLY GGLY A.CC SC CCGG CGGG ARG C PRO	T * GATP GATP GATP GASP GGG GGG GGG GGA GGA GGA GGA GG	A H TCCC SER GGLY  HIC ATCC SER ATCC CCC GLN CCCC PRO CCCC  H	AGGLN TCGG SER AAGG ARG ARG ARG ARG ARG ARG A	GTCL GTCL GTCL AGGGGGAA AAGGGAA AAGGAAA AGGGAA AAGGAAA AGGAAAA AGGAAAA AGGAAAA AGGAAAA AGGAAA AGGAAA AGGAAA AGGAAAA AGGAAAA AGGAAAA AGGAAAA AGGAAAA AGGAAAA AGGAAAAA AGGAAAA AGGAAAAA AGGAAAAA AGGAAAAA AGGAAAAA AGGAAAAAA	TACC TYR TCC SER GALA C *  AGA ARG GLN CTC LEU TCC CAG GLN CTC CLEU	AGT SER ACCC THR  # ATA ATA ILE  G.CC PRO CACC HIS	GGCY GGCY TTCC SER ACCA ACCA ACCA ACCA ACCA ACCA ACCA AC	GTCL VAL TGG CAT GATP  GATP  ASN CCT PRO  ACCC THR	T.A. SER CAG GLN ACT THR GTC VAL GCC ASP VAL GCC ALA. THR C.T PRO TCCC SER	GGT GLY GLY GLY GLY GLY GLY GLY GLY GLY CAT HIS ACA THR ACA THR	 GALA A  GGLY GGLY  GGLY  GGLY  CPR  CPR  CPR  CPR  CPR  CPR  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCCC  CCCCC  CCCCC  CCCCCCCC		AAAA LYS CC. PRO GALU TGC PRO GAAU TTRP TTRP  SGACP CCCO CCCO CCCO CCCO CCCO CCCO CCCO C	CTALLEU GCCA .TAGP TRPP  * CCGO PRO  * SCCTO PRO  T.AR SCCA GGLU GGLY	T TGCS CYS  # GGLY GGLY GGLY GGT AC SALC ASN C	T.C.C.A.A.A.S. CLEUUT	C.TU LEU AASN GGLY CARG GGLY CARG CARG CARG CARG CARG CARG CARG CARG	.AG LYS AAT A GGGG GLY   GGGG CTA LEU   AASN  TCT SER  CTT LEU CTTA  CTT SER  CTTA 	GGARG GGLU GGLY GGLY TYR TTYR TTYR TTYR TTYR TTYR CCTD CC THR  ACC THR  TG MET AASN 	GGA TGGP TGGP TGGP TGGP TGGP TGGP TGGP T	AACC AASH DO TGCCSS  * TTT SER  * TTTG ILE ACG TTTG ILE ACG TTTG CSS  *	A.T THR AAATN AAATN AAATNA AAATN AAATNA AAAATNA AAATNA AAATNAAATNA AAATNAA	C.T PRO CCCC PRO T SER GGGG GGG GGG C C C C C C C C C C C C C
Xeno   Fv-4   Akv   Xeno   Fv-4	GCC GLY AAG LYS CTA LEU T * * * * * * * * * * * * * * * * *	AAAA LYS CCCG PRO GACP ASP GACP ASP GACP CGAC ASP CVAL CCCG ASP CVAL CCCG CCCG CVAL CCCG CCCG CCCG CCCG CCCG CCCG CCCG CC	CCGG CCGC CCCC	GGCC GGCC GGCC GGCC GGCC GGCC FRO  SCTCC CCCC FRO  SCTCC CCCC FRO  SCTCC CCCC FRO  SCTCC CCCC CCCC CCCC CCCC CCCC CCCC	CCCCO TTHE C PRO TTHE C C RO C C C C C C C C C C C C C C C C C	G TGT CYS ACGG THR  # ACCCT THR  # ACCCT THR  # ACCCGG ARG GLN  AAAT ASN	C TATT GGLY A.SER CPRO  CCC CGGG ATG ATG ATG	T * GAATP CCCA PROT FHE GALA BBGG GLLA G GGAA G GGAA G LU ACTTHR	A TCCC SER GGAY T HI ILEU CAGG GLN CCCC PRO  AAACC	AGG GLN TCCG SER AAGS ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG	GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC	TAC TYR TCC SER GCTA A.C. * C * AGA ARG GLN CTCC LEU TCCC SER G.G ALA ALAA	AGT SER ACC THR ATA ATA ILE  # G.CC VAL CCCC PRO CACCHIS	G GGCY 750 SER AG. BA THR  CTCLEU AGAA GGLN CGAN	GTC VAL TGG VAL TGG TRP  JI II GAT ASP  ASN CCT TPRO  GGLU	T.A SER CAG GLN ATHR GAL GAL ASP ICTEU CTEU CTEU CTEU CTEU CTEU CTEU CTEU	GGT GLY ACA THR GGA GGA GGA CA CA THR CA THR TGG TRP	GCCC GALA A.A. GGLY GGLY CCCC CCCC CCCC CCCC CCCC CCCC	CCCC PRO A.A. THR CATS CATS CATS CATS CATS CATS CATS CATS	AAAA LVS CCCO GGLU GGLU GGLU GGLU GGLU GGLU GGLU CCG GGLU CCG GGLU CCG GACA CCG CCG CCG CCG CCG CCG CCG CC	CTALLEU GCC ALA. TGG TTRP CCCGO PRO  TLA SER CCGA GGLY 		T.C.C.A.AAAA ALYS # CLEU T # GGLY # TCAA TCAA TCAA AGG GCY # TCAA C #	C.T.LEU AAATN GGLY CGCGCARGG A A.A CCAAGG CAR	AGG CTA ASN GGG CTA ASN GGG GLY GGG GLY GGG GLY TTCT SER CTT LEU T.A TTCT T.A A	GGA GGU T TACR GGL Y TACR T TACR T TACR	GGA GLY TGG TRP CARG GAA GTC CCC CCC CCC CCC CCC CCC CCC CCC CCC	AACC ASN DO TGCC CYS  TTCT TSCR  TTCT TTCT TCC TCC TCC TCC TCC TCC T	A.T THR AAAT AAAT AAAT AAAT AAAT AAAT AAAT AAA	C.T PRO CCCC PRO T S * GGLY A * GAT ASP C * * ACC * * GAT HR A * GGA * * * GGLY * * *
Xeno   Fv-4   Akv	GCC GLY CTA LEU CTA LEU CAT HIS CAT HIS CAT ALA ARG GLN  CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT	AAAA LYS CCCGPRO GACPRO GACPRO GACPRO GACPRO GACPRO GACPRO GACPRO GACPRO GACPRO CCCGGA GACPRO GACPRO GACPRO CCCGGA GACPRO GACO GACPRO GACPRO GACPRO GACPRO GACPRO GACPRO GACPRO G	CCCGC GLN GTAA GTAA GTAA CCCA PRO  GTAA CCCA PRO  CCCA CCA CCA CCA CCA CCA CCA CCA CC	GGCC GGC GGC GGC GGC GGC GGC GGC GGC GG	CCCC CPRO TTHE CCR CCCC CPRO TTHE CCR CCCC CCCC CCCC CCCC CCCC CCCC CCCC	G TGT CYS ACGG THR  T ACCR THR  ACCR ACGG ARG GLN  AAA CR CARG ARG GLN  AAAN CCAA CCAA CCAA CCAA CCAA CON CON CON CON CON CON CON CON CON CON	C TATT TYR GGLY A.G GLY A.G GLY A.G CT PRO ATG ATG MET C.LU	T GATP GASP CCCA C	A TCCC GGL GGL HI ATCC HI ATCC HI LEU CAGN CCCC PRO  MACCN  AACCN 	AGG GLN TCGG SER AAGS ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG	GTC	TACC TYR TCCR GALA  AARA CAG GLN CLEU TCCR GALA AARA CLEU TCCR GALA AARA	AGT AGT SER ACCCTHR ATA ATA ATA ATA ATA CCCC PRO CACC HIS ACCCTHR 	G GGLY GGLY TTCER 	GTCL GTCL GTCL GTCL GTCL GASP GLU GAGU GLU GAGU	T.A. SER CAG GLN ACT THR GTAL GA. SER G.A. SER VAL GCCC ALA A TCC SER TCCC SER TCCC SER TCCC SER TCCC SER	GGTY GGTY GGTY ACACA GGAY GGAY ACACA CATHR GGAY ACACA CATHR TRP TRP TRP TRP TRP TRP TRP TRP TRP TR	GCLA A. A. A. GGLY CPR OF CCCCO SOCCO SOCCO SC CPR OF CPR	CCCCO PRA.THRA.AR THRA.THRA.THRA.THRA.THRA.THRA.THRA.THRA.	AAAAS LCCCO GGLU GGLU GGLU GGLU GGLU GGLU GGLU GGL	CTALLEU GALAA.TAL TGGP	GGATP GGGATP GGGCY GGLY GGLY GGLY GGLY GGLY GGLY GGCY GGLY GGL	T.CCR ALYS ALYS CLEU CLEU CLEU CLEU CLEU CLEU CLEU CLEU	C.TU LEU AAATN GGLY CGCGCGARG ARG ARG CARG CARG CARG CARG C	AGG GLY AAATN GGG GLY CTT LEU CTT LEU T.A CTT LEU T.A CTT LEU CTT LEU CTT LEU CTT LEU CTT LEU CTT LEU CTT LEU CTT LEU	CGA AGG GGLY T TACK T TACK T TACK T CCTO C C T ACCT C T T ACCT C T T ACCT C T T ACCT C T T ACCT C T T T T	GGAA GLY TGGP TGGP C.A A TGG VGA CARG GTA CARG GTAL CARG GTAL TTG PRO C C C C C C C C C C C C C C C C C C C	AACC AASN DO TGCC CCVS  TTEER  TTEGU  A.LE ACCG TTEC C.T CCVAL  A.LC TTEC C.T CCVAL  A.LC CCVAL 	A.THR AAATNA AAATNA AAAA	C.T PRO CCCC PRO T * GGGG GLY  * GGAT ASP C * C * C *  *



FIG. 3. Reactivity of MuLV *env* probes to  $Fv-4^r$  and  $Fv-4^s$  mouse DNAs. (A) BALB/c-Fv-4w'(Fu) DNA was digested with *PstI*, electrophoresed through 0.6% agarose, transferred to a nitrocellulose membrane, and hybridized to the pEc-B4 probe at 42°C (lane 1) or to the pFv4<sub>env</sub> probe at either 42 (lane 2) or 55°C (lane 3). The 8.3-kb fragment represents the BALB/c endogenous ecotropic provirus; the 4.2-kb fragment is the  $Fv-4^r$ -associated retroviral sequence. (B) EcoRI-digested genomic DNA from BALB/c (lane 1), BALB/c-Fv-4w'(Fu) (lane 2), BALB/c-Fv-4w'(Fu) (lane 3), AKR/J (lane 4), AKR- $Fv-4^r$  (lane 5), FRG ( $Fv-4^r$ ) (lane 6), C57BL/6 (lane 7), C3H/He (lane 8), and NFS (lane 9) strains was immobilized on filters as described above and hybridized to the  $^{32}P$ -labeled pFv4<sub>env</sub> probe at 55°C. (C) The same filter as that for panel B was stripped (washed for 2 h at 68°C in 50% formamide-3× SSC to remove the first probe and rinsed for 30 min at room temperature in 2× SSC) and then rehybridized to the  $^{32}P$ -labeled *PstI-EcoRI* 3' cellular flanking DNA probe (Fig. 1) at 55°C.

pFv4<sub>env</sub> probe was used (Fig. 3A, lane 2). At high stringency, the pFv4<sub>env</sub> probe reacted only with the 4.2-kb  $Fv-4^r$ -specific sequence (Fig. 3A, lane 3). Having determined that the pFv4<sub>env</sub> probe differentiated Fv-4-related env sequences from ecotropic MuLV env sequences, we then used this probe to identify  $Fv-4^r$  sequences in resistant mice. Labeled pFv4<sub>env</sub> DNA hybridized to a 5.2-kb fragment of *Eco*RIdigested genomic DNA from every  $Fv-4^r$  mouse tested; no reactivity with  $Fv-4^s$  (sensitive) mouse DNA was observed (Fig. 3B).

A second probe derived from pFv4 was a 700-base-pair (bp) PstI-EcoRI fragment located in the 3'-flanking cellular DNA (between map units 5.5 and 6.2 [Fig. 1]). This probe was used to ascertain whether the Fv-4-related retroviral sequence was associated with the same 3'-flanking cellular DNA in all  $Fv-4^r$  mice. The 5.2-kb EcoRI segments in the DNA of  $Fv-4^r$  mice that hybridized to the pFv4<sub>env</sub> probe (Fig. 3B) also reacted with the 700-bp PstI-EcoRI 3'-flanking, cellular-sequence probe (Fig. 3C). Both probes also hybridized to KpnI, HindIII, PvuII, or PstI cleavage products of  $Fv-4^r$  genomic mouse DNAs as would be predicted from the map of pFv4 (Fig. 1; data not shown). Genomic DNA of  $Fv-4^s$  mice, which did not hybridize with the pFv4<sub>env</sub> probe (Fig. 3B), did hybridize with the 3'-flanking, cellularsequence probe, as seen by the reactivity of a 1.5-kb EcoRI fragment in all five  $Fv-4^s$  mice tested (Fig. 3C). These data suggest that in all  $Fv-4^r$  mice examine, the  $Fv-4^r$ -associated retroviral sequences were inserted 5' to the same flanking cellular sequence.

Nucleotide sequence of pFv4. Critical regions of pFv4 were

sequenced to establish (i) the limits of the retrovirallike sequence and (ii) similarities or differences with known MuLV sequences. The strategy used to sequence a total of 3,900 nucleotides is shown in Fig. 1. The entire Fv-4'associated retroviral sequence with the exception of the midportion of *env* is shown in Fig. 2, 4, and 5. Analysis of the nucleotide sequence indicated that 3.4 kb of the 5.2-kb pFv4 insert represented MuLV-related sequences. The retroviral sequences included C-terminal *pol* nucleotides, the entire *env* gene, and 3' LTR sequences all of which could be colinearly aligned with AKV ecotropic MuLV DNA as illustrated in Fig. 1. The nucleotides 5' to the partial *pol* sequence and those 3' to the LTR were not related to known MuLV proviruses and were presumed to be of cellular origin (Fig. 4 and 5).

**C-terminal pol region.** The first identifiable retroviral sequences occurred 250 bp from the 5' EcoRI site in pFv4 and represented the last 850 bp of a MuLV *pol* gene (Fig. 4). This C-terminal *pol* sequence had 86% nucleotide and 92% deduced amino acid homology with AKV623 (13). The highly conserved MuLV *env* splice acceptor signal (13, 20, 33, 39), located in the *pol* region, was present in pFv4 275 nucleotides upstream from the beginning of *env* (positions 5508 through 5516 [Fig. 4]), the same position as in other MuLVs.

General features of the pFv4 env sequence. The nucleotide sequence of the N-terminal 1,200 bp of the env region (Fig. 2), as well as 255 bp of the C-terminal portion of env encoding a p15(E) analog (Fig. 5), was sequenced. Both segments of the pFv4 env region has a single open reading frame specifying amino acids which were homologous to

FIG. 2. Sequence of the amino-terminal portion of the pFv4 *env* region and its alignment with corresponding ecotropic and xenotropic MuLV sequences. The 5' *env* sequence of pFv4 is compared with analogous sequences of ecotropic AKV623 (13, 25) and xenotropic NZB IU6 (33) MuLV sequences. Dots or asterisks, respectively, indicate nucleotide or deduced amino acid identities with pFv4. Nucleotide substitutions are shown. Gaps in sequences result from best-fit alignment and do not represent interruptions in the continuous sequence. Boxed amino acids indicate potential glycosylation sites. Recognition sites for several restriction enzymes are included for reference. Numbers shown above each group of the three aligned sequences refer only to the nucleotide position in this alignment.

		5' c	ellu	lar	flan	king	seq	venc	es																					
FV-4	TAA END	TTG LEU	TGA END	TAT TYR	GTA VAL	TAT TYR	GTA VAL	GTC VAL	TTT PHE	AGT SER	AAC ASN	TGG	GAA GLU	ATT	GAT ASP	TTT Phe	ACT Thr	GAA Glu	GTT Val	AAG Lys	CCA Pro	GGG GL Y	CTG LEU	TAT Tyr	GGA GLY	TAC Tyr	AAA LYS	TAT TYR	CTC LEU	CTG LEU
AKV	GTG Val	AGA Arg	GTA Val	CGA Arg	.G. Gly	С НIS	CGG Arg	CCA Pro	GGC GLY	тсс *	C.T HIS	•••• *	G *	c *	 *	•••• *	A *	*	c *	*	•••• *	*	•••• *	••• *	G *	*	G *	с *	••• *	*
FV-4	GTG Val	TTT	GTG Val	GAC ASP	GCG Ala	TTC PHE	TCT SER	GGC GL Y	TGG TRP	GTA Val	Hir GAA GLU	GCT ALA	TTC PHE	CCA Pro	ACC Thr	AAA Lys	CGT Arg	GAA GLU	ACT THR	GCC Ala	AGA Arg	GTT Val	GTA Val	ACC Thr	AAG Lys	AAG Lys	CTG Leu	CTA Leu	GAA Glu	GAA Glu
AKV	••••	c *	•••• *	•••• *	A.C Thr	*	••• *	т *	*	G *	•••• *	с *	т *	••• *	•••• *	G *	A . A *	*	• • A *	• • A *	*	с *	G *	T SER	*	A *	т *	G *	*	••• *
FV-4	ĂTA Ile	TTT Phe	CCA Pro	AGA Arg	TTC Phe	GGA GL Y	ATG Met	CCC Pro	CAG Gln	GTA Val	TTG Leu	GGA Gly	ACT Thr	GAT ASP	AAT ASN	GGG GL Y	CCT Pro	GCC Ala	TTC Phe	ATC Ile	TCC SER	CAG GLN	GTA Val	AGT SER	CAG GLN	TCG SER	GTG Val	GCC Ala	GGT GLY	TTA Leu
AKV	••• *	••• *	G *	···· *	•••• •••	••• *	••• *	• • • •	•••• *	••• *	••• *	••• *	T Ser	••• *	с *	•••• *	••• *	•••• *	••• *	.C. Thr	•••• •	··· *	••• *	••• *	••• *	•••• *	•••• *	•••• *	.A. Asp	••• *
FV-4	CTG Leu	GGG Gly	ATT ILE	GAT ASP	TGG TRP	AAA Lys	CTA Leu	CAT HIS	TGT Cys	GCT Ala	TAC Tyr	AGA Arg	CCC Pro	CAG Gln	AGT SER	TCA SER	GGT GL Y	CAG GLN	GTA Val	GAA Glu	AGA ARG	ATG MET	AAT ASN	AGA Arg	ACA	ATT ILE	AAG LYS	GAG Glu	ACT Thr	TTG Leu
AKV	 *	 *	с *	 *	•••• *	•••• *	т *	52! • • • *	50 •••• *	•••• *	т *	 *	•••• *	•••• *	••• *	••• *	•••• *	 *	••• *	••• *	•••• *	 *	•••• *	···· *	5300 C *	c *	••• *	••• *	•••• *	C.A *
FV-4	ACC	<b>A A A</b>	TŤA	ACG	стт	P GCA	/U I) GCT	I GGC	ACT	AGA	GAC	TGG	GTA	стс	CTA	стс	ccc	TTG	GCC	стт	ТАС	CGA	GCC	CGG	ΑΑΤ	ACT	CCG	GGC	cce	CAT
AKV	THR	L Y S	LEU	THR	LEU	ALA P	ÁLA VU II	GLY I 	тн <b>r</b>	ARG 	ASP 5	TRP 350	VAL	LEU 	LEU	LEU	PR0	LEU A	ALA	LEU 	т у R 	ARG	AL A	ARG 	ASN	тн <b>r</b>	PR0	GLY 541	PR0	HIS
	×	×	*	×	¥	¥	×	¥	*	*	×	¥	¥	*	*	*	¥	×	*	*	*	¥	*	×	×	*	*	*	*	*
FV-4	GGA GLY	CTT Leu	ACT Thr	CCG Pro	TAT Tyr	GAA Glu	ATT ILE	CTA LEU	TAT Tyr	GGG GL Y	GCG Ala	CCC Pro	CCG Pro	CCC PRO	CTT Leu	GTT Val	AAT ASN	TTT Phe	CAT HIS	GAT ASP	CCT Pro	GAA Glu	ATG Met	TCA SER	AAG LYS	CTG Leu	ACT Thr	AAC ASN	AGT Ser	CCT Pro
XENO																							-			-T.A - *	•••• *	••• *	••• *	c *
<b>ÅK V</b>	••• *	G *	•••• *	 *	 *	••• *	c *	т.с *	c *	 *	••• *	•••• *	 *	•••• *	5450 ••• *	••• *	c *	c *	••• *	c *	c *	C Asp	 *	 *	G.A Glu	т.а *	••• *	т *	•••• *	A *
FV-4	TCT SER	CTC LEU	CAA GLN	GCT Ala	acc <u>CAC</u> HIS	epto TTA LEU	r CAG GLN	GCC Ala	CTC LEU	CAA GLN	GCA Ala	GTA VÁL	CAA Gln	CGA Arg	GAG GLU	GTC VAL	TGG TRP	AAG Lys	CCG Pro	CTG Leu	GCC Ala	GCT Ala	GCC Ala	TAT Tyr	CAG GLN	GAC ASP	CAG GLN	CTA LEU	GAT ASP	CAG GLN
XENO	•••	•••	•••	•••	acc •••	epto	r • •	••••	••••	• • •	••••	••••	•••	• • •	•••	•••	·	•••	A	•••	••••	•••	т *	•••	•••	•••	 *	G	 *	•••
Ξάκν	* 	500 	*	*	* acc	epto	r	*			* A.G	G	G	*	A	х. А.Т		A	50 <sup>°</sup> A			. A A	•••	c	.G.				c	A
	*	*	*	*	¥	×	×	¥	¥	*	THR	×	×	*	*	ILE	*	*	×	¥	*	GLU	×	*	ARG	×	*	*	¥	*
FV-4	CCA Pro	GTG Val	ATA Ile	CCA Pro	CAC HIS	CCC Pro	TTT Phe	CGT ARG	GTC Val	GGT GL Y	GAC Asp	GCC Ala	GTG Val	TGG TRP	GTA Val	CGC Arg	CGG Arg	CAC HIS	CAG GLN	ACT Thr	AAG LYS	AAC ASN	TTA Leu	GAA Glu	CCT Pro	CGC ARG	TGG TRP	AAA Lys	GGA Gly	CCC Pro
XENO	••• *	••• *	••• *	••• *	••• *	•••• *	с *	••• *	••• *	••• *	••• *	••• *	••• *	••• *	••• *	•••• *	•••• *	••• *	••• *	••• *	···· *_	*	••• *	••• *	c *	••• *	••• *	••• *	••• *	••• *
AKR	••• *	•••• *	•••• *	••• *	5600  *	•••• *	c *	••• *	A.T Leu	•••	 *	T SER	•••• *	••• *	G *	 *	•••• *	•••• *	••• *	c *	5 A *	650 ••• *	•••• *	••• *	•••• *	••• *	•••• *	G *	•••• *	•••• *
FV-4	TAC	ACC	GTC	CTG	стб	ACC	ACC	ccc	ACC	GCT	стс	<b>A A A</b>	Ac Gta	C I GAC	GEC	ATC	тст	ece	TGG	GTA	CAC	ecc	GCT	CAC	GTA	AAG	GCG	ece	ACA	ACT
XENO		· · · ·	VAL			тнк •••	тнк 	PRU	тнк •••	ALA 			Ac	c I	GL T		5EK	ALA 		¥AL		ALA 	ALA 		• • • •		ALA	ALA 		
AKV	*	*	*	*	*	*	*	* 57	• • •	*	*	*	* Ac	с I	*	*	*	*	*	ILE	*	*	*	*	* 5750 C	*	* A	*	* c	* G
	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	ILE	*	×	*	*	*	×	*	*	*	*
FV-4	CCT Pro	CCG Pro	GCC Al A	GGA GLY	ACA Thr	S ACA THR	tart TGG TRP	of Aga Arg	env GTC VAL	CAG Gln	CGT Arg	TCT SER	CAA Gln	AAC ASN	CCC PRO	TTA Leu	AAG Lys	ATA Ile	AGA Arg	CTA Leu	TCA SER	AAA Lys	A A G L Y S	CTC LEU	TCC SER	TAG END				
XENO	••• *	A *	•••• *	•••• *	••• *	G Ala	 *	.AG LYS	env T *	•••• *	•••• *	•••• *	•••• *	•••• *	••• *	•••• *	••• *	••• *	•••• *	т *	A.C Thr	CGT Arg	GG. Gly	GC. Ala	C Pro	A *				
AKV	c *	•••• *	ATA Ile	AAA Lys	C Pro	s T SER	tart  *	of  *	env A *	A *	5 C *	800 •••	•••• *	· · · · *	т *	•••• *	A *	c *	• • • G *	т *	A.C Thr	CGT Arg	GG. Gly	GC. Ala	C Pro	A *				

FIG. 4. Alignment of pFv4 sequences with the *pol* regions of ecotropic and xenotropic MuLV proviruses. Nucleotide and deduced amino acid retroviral sequences of pFv4 are aligned with corresponding sequences of AKV623 ecotropic (13, 25) and reported NZB IU6 xenotropic (33) MuLV sequences. Nucleotide and deduced amino acid identities are indicated by dots (.) and asterisks (\*), respectively. Several restriction sites are identified for reference. The *env* splice acceptor sequence and the beginning of the overlapping *env* reading frame are indicated. AKV623 sequence numbering is by the method of Herr (13).

Α																														
							4										50													
FV-4	CAG Gln	AAA Lys	LEU	PHE	GAA Glu	TCA SER	GGA GL Y	CAA GLN	GEG GL Y	TGG TRP	PHE	GAG Glu	GGA Gly	CTG LEU	PHE	AAC ASN	AGG Arg	SER	CCA Pro	TGG	PHE	ACG Thr	THR	LEU	ILE	SER	THR	ILE	MET	GEC GL Y
XENO	A *	··· *	G *	•c *	••• *	•••• *	•••• *	••• *	•••• *	•••• *	••• *	•••• *	•••• *	••• *	••• *	••• *	•••• *	•••• *	••• *	*	*	*	••• *	••• *	•••• *	•••• *	•••• *	•••• *	••• *	••••
AKV	A *	G *	с.с *	*	••• *	c *	CA. Gln	••• *	···· *	•••• *	••• *	• • • •	G *	••• *	••• *	т *	.A. Lys	•••• *	т *	••• *	*	с *	· · · · *	••• *	•••• *	•••• *	•••• *	c *	•••• *	т *
			1	00																150										
FV-4	CCT Pro	CTG Leu	ATA	GTA VAL	CTT Leu	TTG Leu	TTA Leu	ATC Ile	CTA Leu	CTT Leu	TTC Phe	GGA Gly	CCC Pro	TGC Cys	ATT Ile	CTC Leu	AAC Asn	CGG Arg	TTG Leu	GTC Val	CAG Gln	TTT Phe	GTA Val	AAA Lys	GAC Asp	AGA Arg	ATT Ile	TCA Ser	GTA Val	GTA Val
XENO	 *	•••• *	•••• *	•••• *	••• *	A *	•••• *	•••• *	 *	с *	C LEV	•••• *	•••• *	•••• *	•••• *	••••	•••• *	с *	•••• *	••• *	•••• *	 *	•••• *	••• *	•••• *	•••• *	••• *	•••• *	•••• *	••• *
AKV	c *	••• *	•••• *	ATC Ile	•c *	••• *	•••• *	т *	т *	с *	т *	G *	т *	т *	•••• *	•••• *	т *	с *	с *	•••• *	 *	•••• *	A.C Ile	•••• *	•••• *	G *	••• *	G *	••• *	G *
							200																2	50						
FV-4	CAG Gln	GCC Ala	CTG Leu	ILE	CTG Leu	ACC Thr	CAA GLN	CAG Gln	TAT Tyr	CAC HIS	CAA Gln	CTC LEU	AAA Lys	TCA Ser	ATA Ile	GAT Asp	CCA Pro	GAA Glu	GĄA Glu	GTG Val	GAA Glu	TCG SER	CGT Arg	GAA Glu	TAA End	AAG	ATTT	TATT	CAGT.	TTCC
XENO	••• *	••• *	••• *	••• *	••• *	 *	••• *	••• *	•••• *	 *	•••	••• *	 *	•••• *	 *	A Glu	 *	••• *	••• *	A *	 *	 *	•••• *	•••• *	 *	.т.	••••	••••	••••	••••
AKV	 *	 *	•••• *	G Val	•••• *	т *	 *	• • • • •	•••• *	т *	•••• *	т *	G *	A Thr	•••• *	A Glu	GAT ASP	TGT Cys	A Lys			A *	 *	*	••• *	•••	••••	••••	••••	
FV-4	AGA	AAGA	6666	GGAA																										
XENO	•••	• • • •	• • • •	••••																										
AKV																														

B FV-4 NZB Xeno AKV Eco	1 [U3] Pst1 50 TGAAAGACCCCACCATAAGGCTTAGCAAGCTAGCTGCAGTAACGCCCATTITGCAAGGCAT GAAAAAGTACCAGAGCTGAG TTCT T. TTTT	100 CCGAGGCTACAA .AA.A.T AAAAAC.A.A	
FV-4 Xeno AKV	150 GGAAGTTTAGTTGAAAAATAAGGCTG AACAATÁCTGGGACAGGGG AAGAGTTT	200 GGCCAAGAACAGAT	
FV-4 Xeno AKV	250 CCAAACAGGATATCTGTGGTCAGGCACCTGGGCCC GGCTCAGGGCC GGTCCCCAGAAACAGAGAGGGCTGGAAAGTACCGGGACTAGGG	SOO AAGAACAGATGGTAC C.	
FV-4 Xeno AKV	350 400 TCAGATAAAGCGAAAACCAGCAACAGTTTCTGGAAA GTCCCA CCTCAGTTTCAAG TTCCCCCAAAAGACCGGG AAATACCCCCAGCC 	CAT bo TTGTTCAAACTAACC 	0×'
FV-4 Xeno AKV	450 'TATA box' [R] AATCAGCTCGCTTCTGGCTCGGGCCTTTTGCTCCCCAGCCCTAGCCCTATAAAAAGGGTAAAAACTTCACACTCGGCGCGC 	510 Cagtcatccgatagá act ct	
FV-4 Xeno AKV	Sma1 Kpn1 'polyA signal' 565 [U5] 600 CTGAGTCGCCCGGGTACCCGTGTTCCCAATAAAGCCTTTTGCTGTT GCATCCGAATCGTGGTCTCGCTGATCCTTGGGAGGGGTCTCCT AT.	CAGAGTGATTGACTA	
FV-4 Xeno AKV	640 3' cellular flanking sequences 685 CCCAGCTCGGGGGTCTTTCACTTCCAGTTACTTTTTTTTT		

FIG. 5. Carboxy-terminal sequence of pFv4 env region and associated 3' LTR and flanking cellular sequences. pFv4 retroviral sequences are aligned with corresponding AKV623 ecotropic (13, 25) and NZB IU6 xenotropic (33) MuLV sequences. Symbols for nucleotide or deduced amino acid identities are as indicated in the legend to Fig. 4. Gaps indicate alignment differences resulting from additional nucleotides in one sequence relative to the others. (A) Alignment of the carboxy-terminal p15(E) env region of pFv4 with AKV623 and NZB IU6 MuLV sequences. A reference nucleotide number for AKV623, calculated by the method of Herr (13), is given. (B) Alignment of pFv4-associated 3' LTR sequence with those of AKV623 ecotropic and NZB IU6 xenotropic MuLV LTRs. The limits of U3, R, and U5 regions are indicated by vertical bars. The inverted terminal repeats are enclosed in boxes. The absence of directly repeated sequences in the U3 region of pFv4 and xenotropic LTRs is represented by the gap between positions 150 and 251. Smaller gaps indicate deletions in one aligned sequence relative to another. The numbering system used refers only to alignment positions. Characteristic LTR signal sequences and restriction sites are noted.



FIG. 6. Expression of  $Fv-4^r$ -associated retroviral RNA in mouse tissues. Total cell RNA (5 µg) was prepared from BALB/c spleen (lane 1), BALB/c- $Fv-4w^r$ (Fu) spleen (lane 2), or an AKR thymic lymphoma (lane 3), electrophoresed in a 1% agarose gel containing formaldehyde, and hybridized, as described in the text, to <sup>32</sup>P-labeled pFv4<sub>env</sub> (lanes 1 and 2) or pEc-B4 (lane 3) DNA probe. The size markers indicate the positions of full-length and processed *env* RNAs.

ecotropic MuLV sequences. As is the case for other MuLVs, the pFv4 *env* sequence overlapped the 3' terminus of *pol*. Alignment of the sequenced pFv4 gp70 and p15(E) regions with comparable segments of AKV623 proviral DNA indicated that the unsequenced portion of the pFv4 clone was identical in size (about 740 bp) to the corresponding AKV sequence. This suggested that a complete MuLV *env* gene was present in the pFv4 clone.

Characteristics of the N-terminal env region of pFv4. The N-terminal half of the MuLV env gene is a variable region in which characteristic differences distinguish ecotropic from MCF or xenotropic MuLVs. For example, a best-fit alignment of ecotropic and xenotropic env sequences generates a series of characteristic reciprocal gaps previously described by Repaske et al. (36). Xenotropic and MCF MuLVs, on the other hand, have similar patterns of insertions and deletions that clearly distinguish them from ecotropic MuLVs (33, 36). Alignment of the env sequence of pFv4 within the variable region fits the pattern of ecotropic and not xenotropic or MCV MuLV sequences (Fig. 2). In this region of pFv4, 73% of the nucleotides or deduced amino acids were shared with AKV. This value was comparable to the 70 to 75% conservation observed when the env regions of Moloney, Friend, and AKV ecotropic MuLVs were compared (13, 21, 25, 39). A similar comparison of xenotropic or MCF with ecotropic MuLV env sequences indicated only 30 to 40% amino acid identity. Similarly, 38% deduced amino acid homology existed between the pFv4 and MCF or xenotropic env sequences (19, 33, 36). Thus, by several criteria, the gp70specific segment of pFv4 was clearly related to the ecotropic class of MuLVs. Although the N-terminal half of the gp70 coding region of the ecotropic MuLVs and the corresponding pFv4 sequences shared several structural features, the 0.7-kb BamHI env-specific segment of the pFv4 could be used to differentiate the Fv-4 resistance determinant from ecotropic proviral DNAs (Fig. 3).

Putative glycosylation sites are located in the same positions in both pFv4 and AKV623 with a single exception of one additional site in pFv4 at positions 1093 to 1101 (Fig. 2). Glycosylation of this additional site could account for the apparent higher molecular weight of the gp70 identified on membranes of Fv-4-resistant mouse cells (16).

**Characteristics of the 3' portion of p15(E).** In the Cterminal portion of p15(E), pFv4 had 86% deduced amino acid sequence homology to ecotropic and MCF MuLVs and 98% homology to xenotropic MuLV (Fig. 5A). There were 50 nucleotide differences between pFv4 and AKV or MCF, whereas only 8 nucleotide differences existed between pFv4 and NZB xenotropic p15(E) sequences. At the 3' end of the p15(E) coding region, both AKV623 and MCF247 have several amino acids deleted relative to NZB xenotropic MuLV (33) or to pFv4 sequences (Fig. 5A, positions 238 to 243).

pFv4 LTR sequence. A complete LTR containing U3, R, and U5 regions was also sequenced (Fig. 5B). Noncoding, nonretroviral sequences, presumably cellular flanking sequences, were found 3' to the LTR (Fig. 5B). A notable feature of the flanking sequence was a 15-bp polythymidylate tract interrupted by a single C. The polyadenylate complement on the opposite strand terminated a short (132-bp) open reading frame (data not shown). With regard to the LTR, all of the characteristic landmarks were present including a 11-bp inverted repeat, a TATA box, a CCAAT box, and a polyadenylation signal. A single copy of a direct-repeat element (positions 251 to 357 [Fig. 5B]) similar to that found in NZB xenotropic (33) and MCF247 MuLVs (19) was present in the pFv4 LTR. Alignment of pFv4 LTR sequences with those of NZB xenotropic (33), MCF247 (19), and AKV ecotropic MuLVs (13, 25) indicated that the pFv4 sequence similarity with the xenotropic p15(E) region extended into the LTR. The pFv4 LTR differed from the NZB xenotropic and AKV623 ecotropic LTRs by 28 and 88 nucleotides, respectively. Furthermore, a five-nucleotide insertion characteristic of xenotropic MuLVs (33) was also identified in the pFv4 LTR (Fig. 5B, positions 457 to 461). Individual nucleotide insertions or deletions found in the xenotropic MuLV LTR sequences were conserved in the pFv4 LTR.

Analysis of mRNA expression. To ascertain whether unique species of RNA were present in Fv-4-resistant mice, total RNA was purified from spleens of BALB/c-Fv-4w'(Fu) and BALB/c ( $Fv-4^s$ ) mice and analyzed by Northern blot hybridization with the pFv4<sub>env</sub> probe. Labeled pFv4<sub>env</sub> DNA reacted strongly with a 3.0-kb RNA species from spleen of BALB/c-Fv-4w'(Fu) mice (Fig. 6, lane 2), but not with RNA from a normal BALB/c ( $Fv-4^s$ ) mouse (Fig. 6, lane 1). The 3.0-kb env message expressed in BALB/c-Fv-4w' spleen reacted with an LTR probe but not with an MuLV pol probe (data not shown). RNA isolated from thymic tumors of AKR/J mice was analyzed in the same gel and hybridized with the pEc-B4 probe to show the sizes of full-length (8.2-kb) and env (3.0-kb) mRNAs transcribed from a typical ecotropic provirus (Fig. 6, lane 3).

### DISCUSSION

We have molecularly cloned the segment of mouse genomic DNA that is associated with Fv-4 resistance. The cloned DNA contains a truncated proviral sequence containing a portion of the *pol* gene, the entire MuLV *env* region, and the 3' LTR; it is flanked at its termini by cellular sequences (Fig. 2, 4, and 5). Classification of the  $Fv-4^r$  *env* sequence as ecotropic was based on (i) the nucleotide and deduced amino acid homologies of 73 and 73%, respectively, with the analogous segment of ecotropic MuLV *env* genes, (ii) the pattern (ecotropic and xenotropic) of MuLV *env* gene nucleotide insertions and deletions which correspond to ecotropic rather than xenotropic gp70 sequences (36; Fig. 2), and (iii) the conservation of the number and position of potential glycosylation sites present in ecotropic MuLV *env* regions (13, 21, 25, 39).

While these data indicate that the  $Fv4^r$  retroviral segment is closely related to ecotropic MuLVs derived from inbred mice, this segment also contained restriction enzyme cleavage sites previously recognized only in ecotropic MuLVs isolated from wild mice. For example, a *Hin*dIII site located at 1.3 maps units (Fig. 1) in pFv4 was conserved in almost all wild-mouse ecotropic MuLVs and was never found in ecotropic MuLVs derived from inbred mice (5, 7). The *PvuI* site of pFv4 at 2.5 map units has been observed only at this position in the Castaneus-E wild-mouse ecotropic MuLV (7). Conservation of these specific restriction sites suggests that pFv4 is evolutionarily related to wild-mouse ecotropic MuLVs, which is consistent with the identification of Fv4like resistance determinants in wild Asiatic (31, 32) and California (10) mice.

Nucleotide sequencing revealed the existence of two unique and potentially significant differences between the gp70 coding sequences of pFv4 and ecotropic MuLVs. A region of amino acid heterology including a six-amino-acid insertion was present from positions 994 to 1044 (Fig. 2); in addition, a novel potential glycosylation site occurred at positions 1093 to 1101 (Fig. 2). The six-amino-acid insertion results in a peak of hydrophilicity (14) unique to Fv-4. This difference in an otherwise ecotropiclike gp70 sequence could be responsible for the unique antigenicity of the  $Fv-4^r$  gp70 (16). The second major differences, an additional glycosylation site, could explain the slightly higher molecular weight of the  $Fv-4^r$  gp70 (15, 16) if this site is glycosylated.

As noted above, sequence analysis of the pFv4 retroviral segment indicated that it had the structure of a truncated proviral since it was lacking approximately 5.4 kb of MuLVrelated sequences. The presence of 250 bp of nonretroviral sequences that abutted the shortened MuLV pol region of the pFv4 insert raised the possibility that a stretch of cellular DNA had been inserted within a full-length Fv-4 provirus, thereby displacing viral DNA sequences in the 5' direction. In search of the missing retroviral DNA, we carried out two experiments. In one,  $Fv-4^r$  cellular DNA was cleaved with KpnI, an enzyme known to cleave within the U5 portion of MuLV LTRs including the proviral DNAs of wild-mouse MuLVs (7). In this regard, the sequencing of the pFv4 segment revealed the presence of a KpnI site at its usual location in the 3' LTR. After restriction of  $Fv-4^r$  cellular DNA with KpnI and transfer to a nitrocellulose membrane, a single 15-kb pFv4<sub>env</sub>-reactive fragment was detected by Southern blot hybridization. In a second experiment, a 17-kb partial MboI cleavage product of Fv-4r cellular DNA was molecularly cloned in a lambda phage vector by using the pFv4<sub>env</sub> probe to screen recombinant plaques (data not shown). The cloned DNA segment contained approximately 12 kb of cellular DNA located upstream from the  $Fv-4^r$ associated retroviral sequences. This cellular DNA segment failed to hybridize to an MuLV LTR probe. Taken together, the results of both experiments are consistent with the absence of an LTR element within 12 kb of the 5' terminus of the Fv-4-associated retroviral sequences.

The detection in  $Fv-4^r$  animals of a 3.0-kb RNA species which hybridized to the pFv4<sub>env</sub> probe and comigrated with

the 3.0-kb *env* mRNA present in an AKR thymoma (Fig. 6) is good evidence that the  $Fv-4^r$ -associated retroviral DNA is located near a strong promoter that functions in conjunction with the pFv4 splice acceptor (Fig. 4) to generate a normallength *env* RNA. It has previously been shown that MuLV *env* RNAs hybridize to U3- and U5-specific probes as well as with labeled MuLV *env* DNA (2). The U<sub>5</sub>-reactive segment is transcribed from the 5' LTR and represents retroviral RNA sequences that are spliced to *env* sequences during the processing of *env* mRNA. The 3.0-kb  $Fv-4^r$  mRNA, while hybridizing to *env*- and U3-specific probes, fails to react with an MuLV U5 LTR probe (F. Laigret and H. Ikeda, manuscript in preparation), indicating, again, that this  $Fv-4^r$  mRNA is not initiated from within a 5' LTR.

Fv-4 may function in a way similar to that of the chicken loci ev3, ev6, and ev9, all of which express high levels of avian retroviral env glycoprotein (37) and restrict susceptibility to exogenous viruses of the same subgroup. These loci represent endogenous defective proviruses; ev3 has a large deletion in the junction of gag and pol, ev6 lacks a 5' LTR and gag region, and ev9 has an undetermined defect (1, 12). Since both ev6 and Fv-4 lack a 5' LTR, it is quite likely that the expression of env gene products is under the control of a cellular promoter.

A viral interference model has been proposed for restriction of ecotropic MuLV infection in Fv-4<sup>r</sup> mice. This concept has been supported by the detection of an  $Fv-4^r$ -associated gp70 having antigenic characteristics different from those of typical ecotropic MuLV gp70s from inbred laboratory mice. Data presented in this paper show that the sequence associated with resistance in  $Fv-4^r$  mice contained a specific env sequence related to an ecotropic MuLV env sequence. The  $Fv-4^r$ -associated env gene had overall characteristics which permitted its gp70 to be classified as an ecotropic type, but sufficient sequence differences existed in one region to account for a potential antigenic differences and an additional glycosylation site which could account for the larger apparent molecular weight of gp70 (see above). RNA studies showed that the  $Fv-4^r$ -associated env gene sequence is expressed. All of the data were compatible with Fv-4 resistance based on the expression of an ecotropic MuLV-related gp70 which competes with the env glycoproteins of ecotropic MuLVs for a cell-specific receptor(s), thereby inhibiting exogenous viral infection. Additional studies on the promoter region of Fv-4 as well as adjacent cellular sequences are in progress.

#### ACKNOWLEDGMENTS

We thank T. Okada for Fv-4 congenic mice, T. Theodore for lambda phage packaging extracts, and R. O'Neill and A. Rabson for DNA sequencing and RNA blotting procedures. We thank C. Kozak and C. Buckler for reviewing the manuscript and S. Rosenfeld and J. Barnhart for typing the manuscript.

#### LITERATURE CITED

- 1. Baker, B., H. Robinson, H. E. Varmus, and J. M. Bishop. 1981. Analysis of endogenous avian retrovirus DNA and RNA: viral and cellular determinants of retrovirus gene expression. Virology 144:8-22.
- 2. Boccara, M., M. Souyri, C. Magarian, E. Stavnezer, and E. Fleissner. 1983. Evidence for a new form of retroviral *env* transcript in leukemic and normal mouse lymphoid cells. J. Virol. **48**:102–109.
- 3. Buckler, C. E., M. D. Hoggan, H. W. Chan, J. F. Sears, A. S. Khan, J. L. Moore, J. W. Hartley, W. P. Rowe, and M. A.

Martin. 1982. Cloning and characterization of an envelopespecific probe from xenotropic murine leukemia proviral DNA. J. Virol. 41:228–236.

- Chan, H. W., T. Bryan, J. L. Moore, S. P. Staal, W. P. Rowe, and M. A. Martin. 1980. Identification of ecotropic proviral sequences in inbred mouse strains with a cloned subgenomic DNA fragment. Proc. Natl. Acad. Sci. U.S.A. 77:5779–5783.
- Chattopadhyay, S. K., M. M. Cloyd, D. L. Linemeyer, M. R. Lander, E. Rands, and D. R. Lowy. 1982. Cellular origin and role of mink cell focus-forming viruses in murine thymic lymphomas. Nature (London) 295:25-31.
- Chattopadhyay, S. K., M. R. Lander, E. Rands, and D. R. Lowy. 1980. Structure of endogenous murine leukemia virus DNA in mouse genomes. Proc. Natl. Acad. Sci. U.S.A. 77:5774–5778.
- 7. Chattopadhyay, S. K., A. I. Oliff, D. L. Linemeyer, M. R. Lander, and D. R. Lowy. 1981. Genomes of murine leukemia viruses isolated from wild mice. J. Virol. 39:777–791.
- Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299.
- 9. Dorner, A. J., J. P. Stoye, and J. M. Coffin. 1985. Molecular basis of host range variation in avian retroviruses. J. Virol. 53:32-39.
- Gardner, M. B., S. Rasheed, B. K. Pal, J. D. Estes, and S. J. O'Brien. 1980. Akvr-1, a dominant murine leukemia virus restriction gene, is polymorphic in leukemia-prone wild mice. Proc. Natl. Acad. Sci. U.S.A. 77:531-535.
- Hager, G. L., E. H. Chang, H. W. Chan, C. F. Garon, M. A. Israel, M. A. Martin, E. M. Scolnick, and D. R. Lowy. 1979. Molecular cloning of the Harvey sarcoma virus closed circular DNA intermediates: initial structural and biological characterization. J. Virol. 31:795–809.
- Hayward, W. S., S. B. Braverman, and S. M. Astrin. 1980. Transcriptional products and DNA structure of endogenous avian proviruses. Cold Spring Harbor Symp. Quant. Biol. 44:1111-1222.
- Herr, W. 1984. Nucleotide sequence of AKV murine leukemia virus. J. Virol. 49:471–478.
- 14. Hopp, T. P., and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. U.S.A. 78:3824–3828.
- Ikeda, H., and T. Odaka. 1983. Cellular expression of murine leukemia virus gp70-related antigen on thymocytes of uninfected mice correlated with Fv-4 gene-controlled resistance to Friend leukemia virus infection. Virology 128:127-139.
- 16. Ikeda, H., and T. Odaka. 1984. A cell membrane "gp70" associated with Fv-4 gene: immunological characterization, and tissue and strain distribution. Virology 133:65-76.
- Ikeda, H., H. Sato, and T. Odaka. 1981. Mapping of the Fv-4 mouse gene controlling resistance to murine leukemia viruses. Intl. J. Cancer. 28:237-240.
- Kai, K., H. Ikeda, Y. Yuasa, S. Suzuki, and T. Odaka. 1976. Mouse strain resistant to N-, B-, and NB-tropic murine leukemia viruses. J. Virol. 20:436–440.
- Kelly, M., C. A. Holland, M. L. Long, S. K. Chattopadhyay, D. R. Lowy, and N. H. Hopkins. 1983. Nucleotide sequence of the 3' end of MCF247 murine leukemia virus. J. Virol. 41:435-448.
- Khan, A. S. 1984. Nucleotide sequence analysis establishes the role of endogenous murine leukemia virus DNA segments in formation of recombinant mink cell focus-forming murine leukemia viruses. J. Virol. 50:864–871.
- Koch, W., G. Hunsmann, and R. Friedrich. 1983. Nucleotide sequence of the envelope gene of Friend murine leukemia virus. J. Virol. 45:1–9.
- 22. Kozak, C. A., N. J. Gromet, H. Ikeda, and C. E. Buckler. 1984. A unique sequence related to the ecotropic murine leukemia virus is associated with the *Fv-4* resistance gene. Proc. Natl. Acad. Sci. U.S.A. 81:834–837.
- Leder, P., D. Tiemeier, and L. Enquist. 1977. EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the λgtWES system. Science 196:175–177.

- Lehrach, H., D. Diamond, J. M. Wazney, and H. Boedtker. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry 16:4743-4751.
- Lenz, J., R. Crowther, A. Straceski, and W. Haseltine. 1982. Nucleotide sequence of the AKV env gene. J. Virol. 42:519-529.
- Lowy, D. R., E. Rands, S. K. Chattopadhyay, C. F. Garon, and G. L. Hager. 1980. Molecular cloning of infectious integrated murine leukemia virus DNA from infected mouse cells. Proc. Natl. Acad. Sci. U.S.A. 77:614-618.
- 27. Maniatis, T., A. Jeffrey, and D. G. Kleid. 1975. Nucleotide sequence of the rightward operator of phage  $\lambda$ . Proc. Natl. Acad. Sci. U.S.A. 72:1184–1188.
- Martin, M. A., T. Bryan, T. F. McCutchan, and H. W. Chan. 1981. Detection and cloning of murine leukemia virus-related sequences from African green monkey liver DNA. J. Virol. 39:835-844.
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499-560.
- O'Brien, S. J., E. J. Berman, J. D. Estes, and M. B. Gardner. 1983. Murine retroviral restriction genes Fv-4 and Akvr-1 are alleles of a single locus. J. Virol. 47:649-651.
- Odaka, T., H. Ikeda, K. Moriwaki, A. Matsuzawa, M. Mizuno, and K. Kondo. 1978. Genetic resistance in Japanese wild mice (*Mus musculus molossinus*) to an NB-tropic Friend murine leukemia virus. J. Natl. Cancer Inst. 61:1301-1306.
- 32. Odaka, T., H. Ikeda, H. Yoshikura, K. Moriwaki, and S. Suzuki. 1981. Fv-4: gene controlling resistance to NB-tropic Friend murine leukemia virus. Distribution in wild mice, introduction into genetic background of BALB/c mice, and mapping of chromosomes. J. Natl. Cancer Inst. 67:1123-1127.
- 33. O'Neill, R. R., C. E. Buckler, T. S. Theodore, M. A. Martin, and R. Repaske. 1985. Envelope and long terminal repeat sequences of a cloned infectious NZB xenotropic murine leukemia virus. J. Virol. 53:100–106.
- Queen, C. L., and L. J. Korn. 1980. Computer analysis of nucleic acids of proteins. Methods Enzymol. 65:595-609.
- Rabson, A. B., P. E. Steele, C. F. Garon, and M. A. Martin. 1983. mRNA transcripts related to full-length endogenous retroviral DNA in human cells. Nature (London) 306:604–607.
- Repaske, R., R. R. O'Neill, A. S. Khan, and M. A. Martin. 1983. Nucleotide sequence of the *env*-specific segment of NFS-Th-1 xenotropic murine leukemia virus. J. Virol. 46:204–211.
- Robinson, H. L., S. M. Astrin, A. M. Senior, and F. H. Salazar. 1981. Host susceptibility to endogenous viruses: defective, glycoprotein-expressing proviruses interfere with infections. J. Virol. 40:745-751.
- Schweizer, J., and K. Goerttler. 1980. Synthesis in vitro of keratin polypeptides directed by mRNA isolated from newborn and adult mouse epidermis. Eur. J. Biochem. 112:243-249.
- Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe. 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) 293:543-548.
- Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-518.
- Suzuki, S. 1975. Fv-4: a new gene affecting the splenomegaly induction by Friend leukemia virus. Jpn. J. Exp. Med. 45:473-478.
- Suzuki, S., K. Tsuji, and K. Moriwaki. 1981. Friend murine leukemia virus resistance in Japanese wild mice: possible allelism with Fv-4 in FRG mice. J. Natl. Cancer Inst. 66:729-731.
- 43. Teich, N., J. Wyke, T. Mak, A. Bernstein, and W. Hardy. 1982. Pathogenesis of retrovirus-induced disease, p. 785–997. In R. Weiss, N. Teich, H. Vermus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.
- 44. Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. 77:5201–5205.
- 45. Vogelstein, B., and D. Gillespie. 1979. Preparation and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci.

U.S.A. 76:615-619.

- Wilbur, W. J., and D. J. Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. U.S.A. 80:726-730.
- 47. Yoshikura, H., Y. Naito, and K. Moriwaki. 1979. Unstable

resistance to G mouse fibroblasts to ecotropic murine leukemia virus infection. J. Virol. 29:1078-1086.

 Yoshikura, H., and T. Odaka. 1982. Surface antigen expressed in hematopoietic cells derived from Fv-4<sup>r</sup> mouse strains. J. Natl. Cancer Inst. 68:1005-1009.