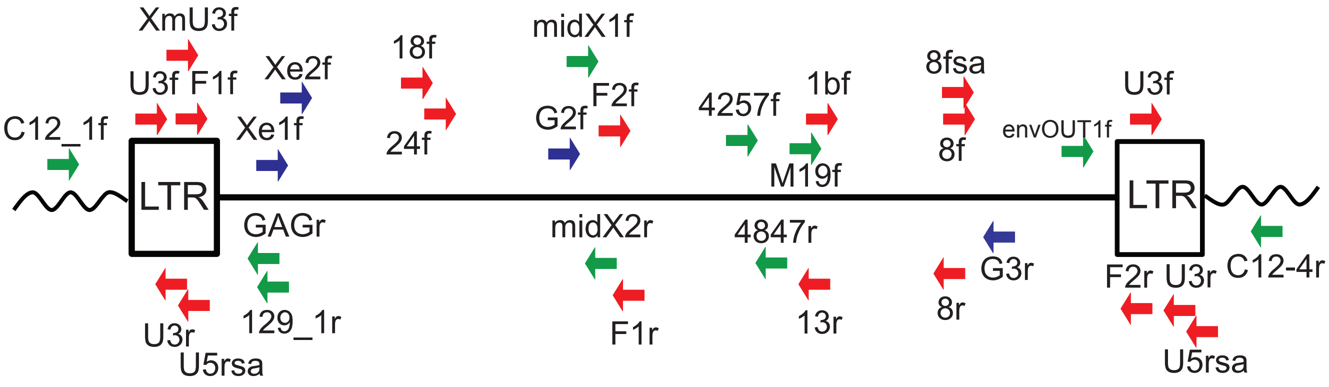
Supplemental Fig. S1:
Short Tandem Repeat (STR) Analysis

	Xenografts: 736, 777, 9216R, 9218R, 8R, and 8L		**22R*v1***		**CWR-R1**	
locus	**alleles**	**frequency, 1 in ___**	**alleles**	**frequency, 1 in ___**	**alleles**	**frequency, 1 in ___**
D5S818*	**11, 12**	3.91	**11, 13**	9.50	**11, 12**	3.91
D7S820†	**9,10,11**	11.60	**9,10,11**	11.60	**9,10,12**	11.60
D13S317	**8, 12**	15.94	**9, 12**	26.39	**8, 12**	15.94
D16S539	**12, 12**	9.77	**12, 12**	9.77	**12, 12**	9.77
TH01	**6, 9.3**	7.39	**6, 9.3**	7.39	**6, 9.3**	7.39
TPOX	**8, 8**	4.11	**8, 8**	4.11	**8, 8**	4.11
CSF1PO	**10, 11**	5.86	**10, 11**	5.86	**10, 11**	5.86
Random Match Probability, 1 in ___ :		1,259,877		5,062,112		1,259,877
Random chance that the xenografts and 22R*v1* or R1 would have the same allele pattern:				1.6×10^{-13}		6.3×10^{-13}

*STR analysis at locus D5S818 was not determined for sample 736.

†Frequency for D7S820 was determined using only the 9 and 10 allele frequencies.

Supplemental Fig. S2: Primers used for PCR

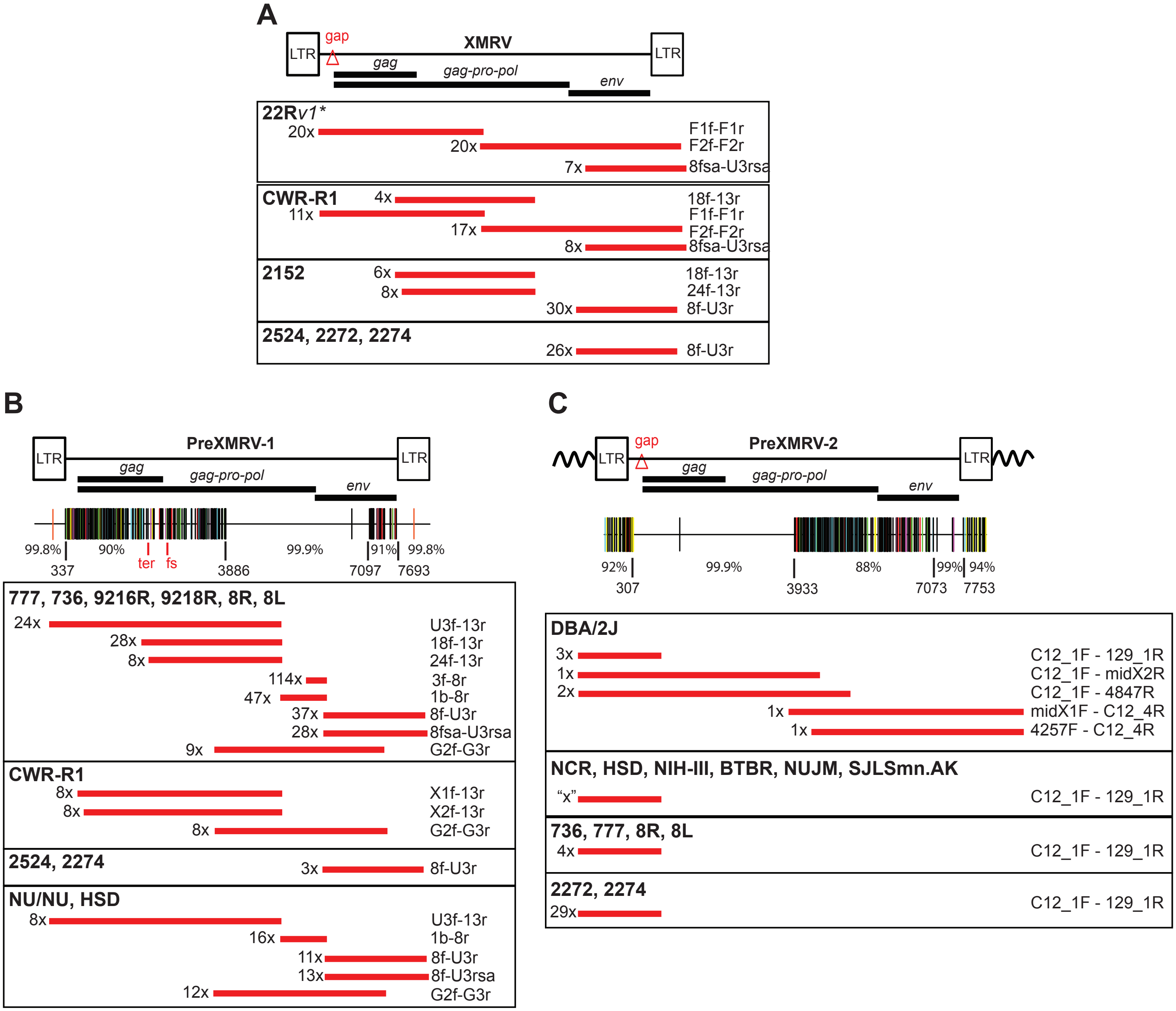


Primer:	Sequence	5' nucleotide position#	1	2	X
U3f	5'-GTTTAATTAAAGAATAAGGCTGAATAAC-3'	7826	+	-	+
13r	5'-ATGTCTTCTAACAGCTTTTTGGACACG-3'	5108	+	-	+
18f	5'-GGCAGAGGATGAGCAGAGAGAGAG-3'	1974	-!	+	+
24f	5'-AAGAAAAGGGACACTGGGCTAAGG-3'	2129	+	+	+
1bf	5'-AGGCATTCCCGACCAAGCG-3'	5048	+	-	+
8r	5'-CTGGATGCTACCGGAGCCC-3'	6284	+	-	+
8f	5'-GGGCTCCGGTAGCATCCAG-3'	6266	+	-	+
U3r	5'-CCCCTTTTTTATAGGGCTAGGAC-3'	8096	-!	-	+
8fsa	5'-CCTGTTTTGATTCCTCAGTGGG-3'	6247	+	-	+
U5rsa	5'-TCTGAGGAGACCCTCCCAAGG-3'	112	+	+	+
F1f	5'-GCGCCAGTCATCCGATAGACTGAGTCGCCCGGGTACCCGTGTTCCCAATAAAGCC-3'	1	-	+	+
F1r	5'-GCCGACGCCAAGGTCCCAGTTTTTGCGTTAGGACGCCTTTGGCGTAGCCCTGCTTCTCGTCGACAAAGAGTTC-3'	3754	-	+	+
F2f	5'-GAACTCTTTGTCGACGAGAAGCAGGGCTACGCCAAAGGCGTCCTAACGCAAAAACTGGGACCTTGGCGTCGGC-3'	3682	-	+	+
F2r	5'-TTGCAAACAGCAAAAGGCTTTATTGGGAACACGGGTACCCGGGCGACTCAGTCTATCGGATGACTGGC-3'	8185	+	-	+
XmU3f	5'-GTCCTAGCCCTATAAAAAAGGGG-3'	8074	-	-	+
G2f	5'-CCCTTATACCCGCTCACCAAGAC-3'	3550	+	-	-
G3r	5'-TGGAGCTGCTCAAATTGTTGGG-3'	7204	+	-	-
Xe1f	5'-GTGGCCCAATCAGTAAGTCCGAG-3'	410	+	-	-
Xe2f	5'-CACTCCCTTGAGTCTGACCCTTG-3'	630	+	-	-
GAGr	5'-TCCCCCAACAAAGCCACTCCA-3'	473	-	+	+
midX1f	5'-TTGTCGACGAGAAGCAGGGC-3'	3689	-	+	+
midX2r	5'-TGCGTTAGGACGCCTTTGGC-3'	3731	-	+	+
envOUT1f	5'-CTGACCCAACAGTATCACCAACTC-3'	7629	+	+	+
C12_1f	5'-TGCTGGACAGAATCTCTGGTCTCT-3'	Ch12	-	-	-
C12_4r	5'-GATACTCAAGTGGTTCCCACCC-3'	Ch12	-	-	-
129_1r	5'-GCGGTTTCGGCGTAAAACCGAAAGCA-3'	537	-	+	+
4847r	5'-CTTTGCTGGCATTTACTTGGGCA-3'	4909	+	-!	+
4257f	5'-GATGGCAGAAGGTAAGAAGCTAAATGTTTA-3'	4296	+	-!	+
M19f	5'-TGGCCTTACTGAAAGCTCTCTTCC-3'	4436	-	+	-

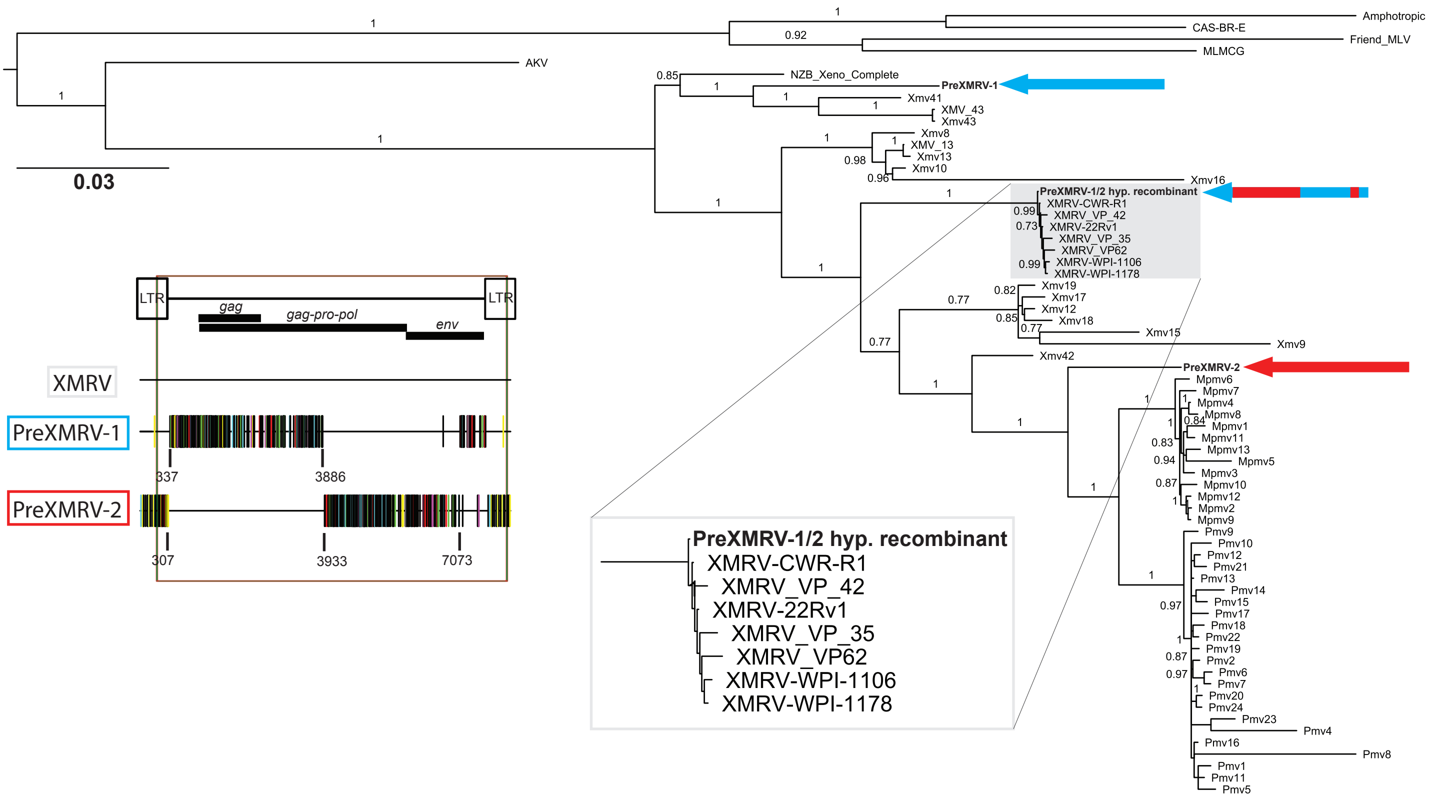
#: 5'nucleotide position relative to XMRV-22Rv*1* (Acc. FN692043).

-!: primer has mismatches with the proviral genome, but was able to amplify the proviral sequence.

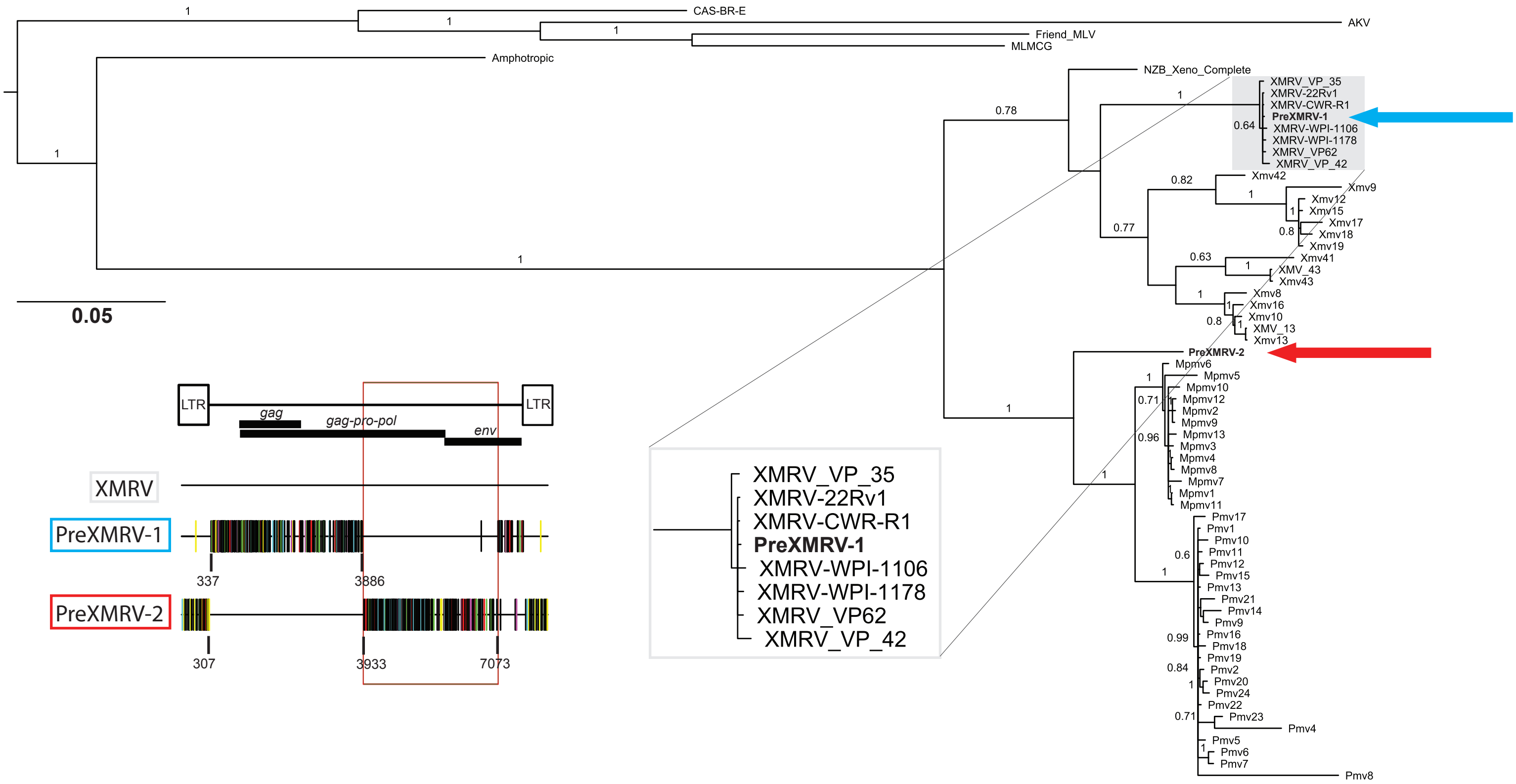
SUPPLEMENTAL FIG. S3: Location and Numbers of Cloned PCR Products Used for Sequencing Analysis



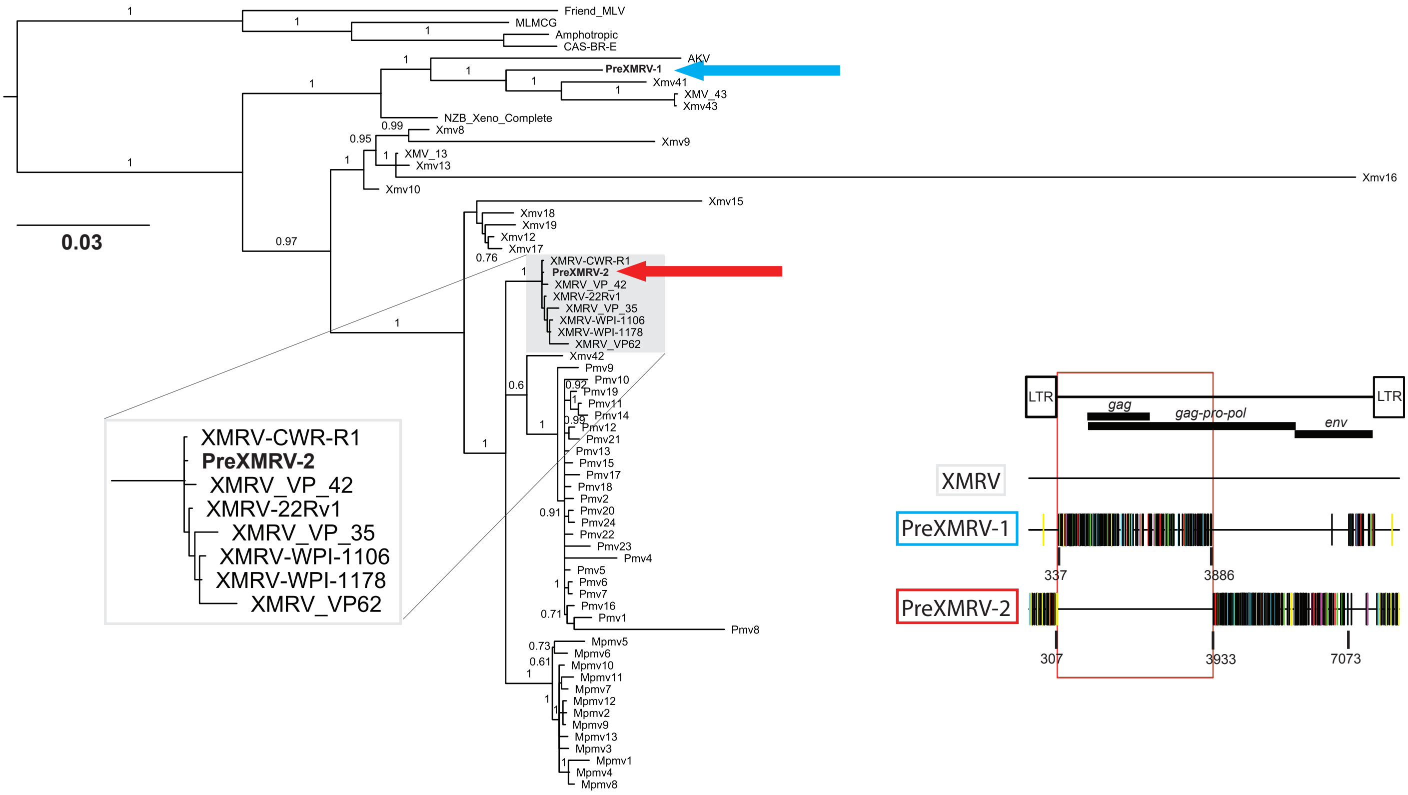
Supplemental Figure S4A: Phylogenetic analysis of full-length genomes



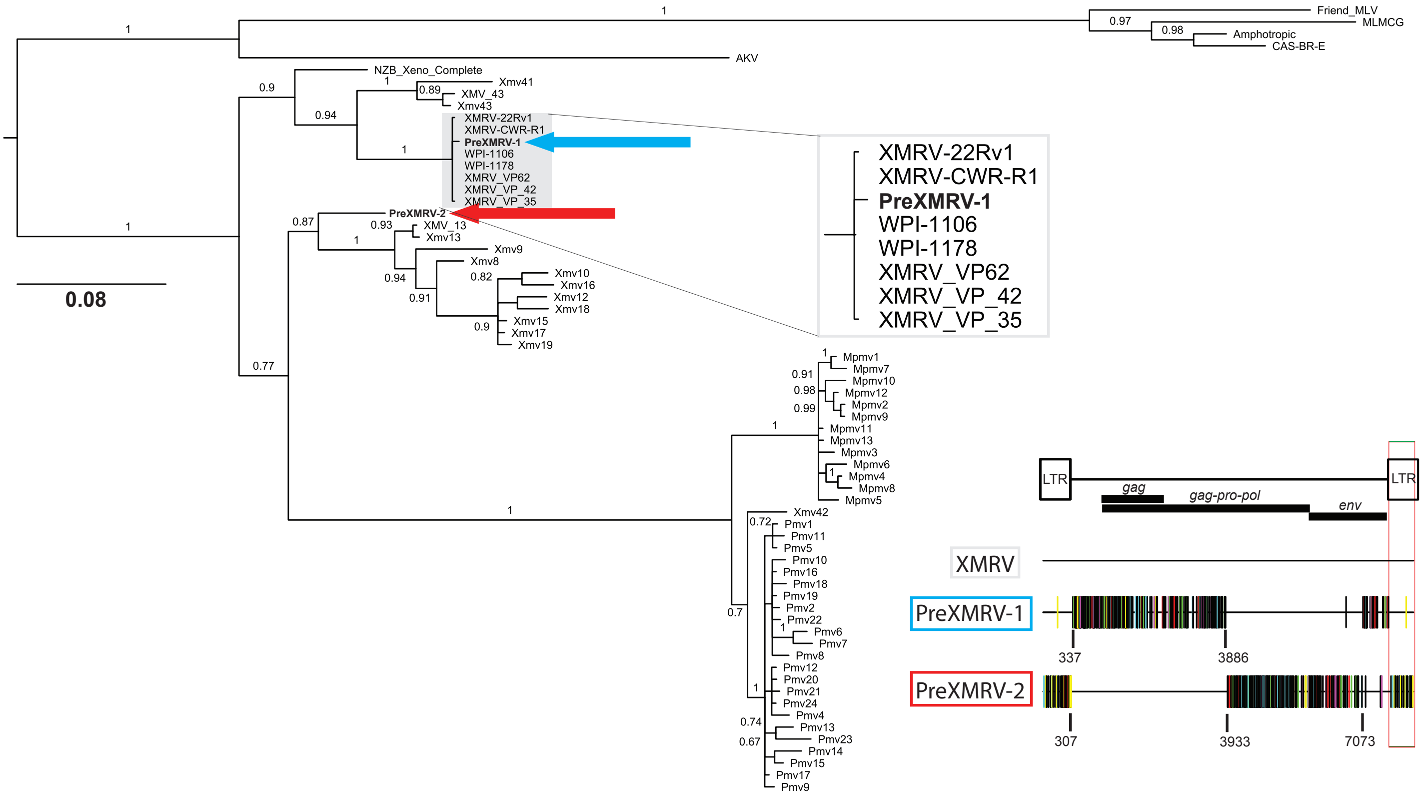
Supplemental Figure S4B: Phylogenetic analysis of XMRV and PreXMRV-1 homology



Supplemental Figure 4C: Phylogenetic analysis of XMRV and PreXMRV-2 homology



Supplemental Figure 4D: Phylogenetic analysis of the U3-R region



Supplemental Fig. S5: Predicted Regions of RT Template Switching Events and XMRV Sequence Diversity

A

Predicted Regions of RT Template Switching Events

```
      7693
1   TAAAT GATTTTATTCAGTTTCC............CCATAAGGCTTAGCA CGCTA
    TAAAA GATTTTATTCAGTTTCC............CCATAAGGCTTAGCA AGCTA
                                                       7753
                                      7557
2   GATAG TACTTTTATTAATCCTACTC CTCGG
    GATAA TACTTTTATTAATCCTACTC TTCGG
      7536
      7432
3   AACCT GAGACAAAAATTGTTCG............CCATGGTTCACGACC CTGAT
    AACCA GAGACAAAAATTGTTCG............CCATGGTTCACGACC TTGAT
                                                       7506
                                         7097
4   AAACT AAATATAAAAGAGAGCCGGTGTC TTTAA
    AAACC AAATATAAAAGAGAGCCGGTGTC ATTAA
      7073
      3886
5   CATCT TGGCCCCCCATGCGGTAGAAGCACTGGTCAAACAACCCCCTGACCG TTGGC
    CATTC TGGCCCCCCATGCGGTAGAAGCACTGGTCAAACAACCCCCTGACCG CTGGC
                                                         3933
                                           337
6   AGTTA GCTAACTAGATCTGTATCTGGCGGTTCCG TGGA
    AGTTG GCTAACTAGATCTGTATCTGGCGGTTCCG CGGA
      307
```

B

XMRV Sequence Diversity