A RAS-clone12 RAS-clone21 RAS-clone22 RAS-clone28 HERVK-LTR-LE	1:-S0666666666666664ATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 59 1:-S06666666366666693A6TACTAA666AACTCA6A56CT6666666ACTCCTCCATATGCT5 59 1:6566666663663666666AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 1:65666666666666666AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 2:66666666666666666AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 2:66666666666666666AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 2:66666666666666666AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 2:666666666666666664AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 2:66666666666666666664AATACTAA666AACTCA6A56CT666666ATCCTCCATATGCT5 60 2:66666666666666666664AATACTAA666AACTCA6A56CT6666664TCCTCCATATGCT5 60 2:66666666666666666664AATACTAA666AACTCA6A56CT6666664TCCTCCATATGCT5 60 2:66666666666666666664ATACTAA666AACTCA6A56CT6666664TCCTCCATATGCT5 60 2:6666666666666666664ATACTAA666AACTCA6A56CT6666664TCCTCCATATGCT5 60 2:6666666666666666664ATACTAA666AACTCA6A56CT6666664TCCTCCATATGCT5 60 2:6666666666666666664ATACTAA666AACTCA656CT66666647CCTCCATATGCT5 60 2:66666666666666666664ATACTAA666AACTCA6456CT66666647CCTCCATATGCT5 60 2:6666666666666666664ATACTAA666AACTCA6456CT6666647CTCCATATGCT5 60 2:66666666666666666664ATACTAA666AACTCA6456CT6666647CTCCATATGCT5 60 2:66666666666666666666447CTCCATATGCT6666667CTCCATATGCT5 60 2:66666666666666666647CTCCATATGCT66666666666666666666666666666	C RAS- RAS- RAS- RAS- RAS- HERV
RA5-clone12 RA5-clone21 RA5-clone22 RA5-clone28 HERVK-LTR-1e	60:#AACGETGETTCECCGGGTECCCTTTTTTTTTTTTTTTTTT	chr_ RA5- RA5- RA5-
RAS-clone12 RAS-clone21 RAS-clone22 RAS-clone28 HERVK-LTR-16	120.IIITCCAANICICICGECCCACCIIACEIGAAACACCCACAGEIGIGIAGUGGCAACCCA 179 120.IIITCCAANICICICGEICCCACCIIACEIGAAACACCCACAGEIGIGIAGUGGCAACCCA 179 121.IIITCCAANICICICGEICCCACCIIACEIGAAACACCCCACAGEIGIGIAGUGCAACCCA 100 121.IIIICCAANICICICGEICCCACCIIACEIGAAACACCCACAGEIGIAGUGGCAACCCA 100 120.IIIICCAANICICICGEICCCACCIIACEIGAAACACCCACAGEIGIAGUGGGAACCCA 100	RAS- RAS- HERV chr_ RAS-
RA5-clone12 RA5-clone21 RA5-clone22 RA5-clone28 HERVK-LTR-le	180:CCCCTACATCTGGTGCCCAACGTGGAGGCTTTTCTCTAGGGTGAAGGTACGCTCGAG 180:CCCCTACATCTGGTGCCCAACGTGGAGGCTTTTCTCTAGGGTGAAGGTACGCTCGAG 181:CCCCTACATCTGGTGCCCAACGTGGAGGCTTTTCTCTAGGGTGAGGTACGCTCGAG 181:CCCCTACATCTGTGCCCCAACGTGGAGGCTTTTCTCTAGGGTGAGGTACGCTCGAG 237 181:CCCCTACATCTGGCCCAACGTGGAGGCTTTCTCTAGGGTGAGGTACGCTCGAG 237 181:CCCCTACATCTGGCCCAACGTGGAGGCTTTCTCTAGGGTGAGGTACGCTCGAG 237	RA5-0 RA5-0 HERVI
B RA5-clonel RA5-clone3 RA5-clone16	1:GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	RA5 RA5 RA5 RA5 RA5 RA5 HERV
RA5-clonel7 RA5-clonel8 RA5-clonel9 HERVK-LTR-La RA5-clone1	1:	RAS RAS RAS RAS
RAS-clone3 RAS-clone16 RAS-clone17 RAS-clone18 RAS-clone19 HERVK-LTR-L6	30:AACGETGGTTCCCCGGGTCCCCTTCTTCTTTTTTTTTTGTCTGTGTGTCTTTTTC 89 32:AACGETGGTTCCCCGGGTTCCCCCTTATTCTTTCTTCTCTAACTTTGTCTCGTGTCTTTTC 91 32:AACGETGGTTCCCCGGTTCCCCTTATTCTTTCTTTCTAACTTGTCTCGTGTCTTTTC 91 32:AACGETGGTTCCCCGGTTCCCCTTATTCTTTTCTTAACTTGTCTCTGTGTCTTTTC 91 34:AACGETGGTTCCCCGGGTCCCCCTTATTCTTTTCTAACTTGTCTCGTGTCTTTTC 93 34:AACGETGGTTCCCCGGGTCCCCCTTATTCTTTCTCTAACTTGTCTCGTGTGTCTTTTC 93 820007 4941:aacgetggttccccggggtccccttcttttttttttcttabatttgtctctgtgtctttttc 4900	HERV chr_
RAS-clonel RAS-clone3 RAS-clonel6 RAS-clonel7 RAS-clonel9 PAS-clonel9	94:TTTTCCAAATCTCTCBTCCCACCTTACBASAAACACCACACGBTSTGTABGBGCAACCCA 153 90:TTTTCCAAATCTCTCBTCCCACCTTACBGAAACACCCACAGBTSTGTABBBGCAACCCA 149 92:TTTTCCAAATCTCTCBTCCCACCTTACBGAAACACCCACAGBTSTGTABBBGCAACCCA 151 92:TTTTCCAAATCTCTCBTCCCACCTTACGBAAACACCCACAGBTSTGTABGBGCAACCCA 151 92:TTTTCCAAATCTCTCBTCCCACCTTACGBAAACACCCACAGBTSTGTABGBGCAACCCA 151 94:TTTTCCAAATCTCTCBTCCCACCTTACGBAAACACCCACAGBTSTGTABGBGCAACCCA 151 94:TTTTCCAAATCTCTCBTCCCACCTTACGBAAACACCCACAGBTSTGTABGBGCAACCCA 151	
HERVK-LTR-L RA5-clonel RA5-clone3 RA5-clonel6 RA5-clonel7	adder #901:ttttccaaatctctcgtcccaccttacgagaaaacaccacaggtgtgtagggggcaaccca #960 154:CCCCTAC#TCTGGTGCCCAACGTGGAGGCTTTTCTCTAGGGTGAAGGTACGCTCG 208 150:CCCCTAC#TCTGGTGCCCAACGTGGAGGCTTTTCTCTAGGGTGAAGGTACGCTCG 204 151:CCCCTAC#TCTGGTGCCCAACGTGGAAGGCTTTTCTCTAGGGTGAAGGTACGCTCG 206 152:CCCCTAC#TCTGGTGCCCAACGTGGAAGGCTTTTCTCTAGGGTGAAGGTACGCTCG 206 152:CCCCTAC#TCTGGTGCCCAACGTGGAAGGCTTTTCTCTAGGGTGAAGTACGCTCG 206	
RA5-clone18 RA5-clone19 HERVK-LTR-L	152:CCCTAC4/TCT6GT6CCCAACAT6GA6GCTTTTCTCTA68GT6AA6GTAC6CTC6 206 154:CCCCTAC4/TCT6GT6CCCAAC6T6GA6GCTTTTCTCTA68GT6AA6GTAC6CTC6 208 #adem #961:cccctac4/tctggtgcccaacgtggaggctLttcLctagggtgaaggtaggtcgctg #1017	

RA5-clone4 RA5-clone5 RA5-clone6	1:66666666666666664TTTGTCAAAACCCACAGAATGTACAACAACAAAGA6TGACCCCTA 0: 0:	· 60 - 0 - 0
RA5-clone8	0:	- 0
RA5-clone11	0:	- 0
HERVK-LTR-Leader	§749:gaccctgacacatccccctcttcgagaaacacccacagatgatcagtaaatactaaggg	a #808
chr_7_NW_00183907	1:ATTTGTCAAAACCCATAGAATGTACAACACAAAAGAGTGACCCCTA	46
RA5-clone4	61: TATTTGGTTGATAATAATGTGTCAATATCAGTTCATCAGTTGTAATAACTGGACCACAC	120
RA5-clone5	1:GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	22
RA5-clone6	1:GGGGGG	6
RAS-clone8	1;666666	6
RA5-clone11	0:	0
HERVK-LTR-Leader	#809:actcagaggctggcgggatcctccatatgctgaacgctggttccccgggtccccttcttt	#868
chr_7_NW_00183907.	47:TATTTGGTTGATAATAATGTGTCAATATCAGTTCATCAGTTGTAATAACTGGACCACACI	106
RA5-clone4	121:GATGTAAGATATTGATGGTAGAGAGGCTGTGAGCTGGTGGAGAAGGGAGTATGTGGGAA	: 180
RA5-clone5	23:GATGTAAGATATTGATGGTAGAGAGGCTGTGAGCTGGTGGAGAAGGGAGTATGTGGGAA	82
RA5-clone6	7:000000000000000000000000000000000000	66
RA5-clone8	7:GGGGGGGGGGGATTGATGGTAGAGAGGCCGTGAGCTGGTGGAGAAGGGAGTATGTGGGAA	66
RA5-clonell	1:GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	25
HERVK-LTR-Leader	#869:ctttctctatactttgtctctgtgtctttttcttttccaaatctctcgtcccaccttac	g (928
chr 7 NW 00183907	107:GATGTAAGATATTGATGGTAGAGAGGCTGTGAGCTGGTGGAGAAGGGAGTATGTGGGAA	166
RA5-clone4	181:TCTCTTTACTTGCTCAGTTTTGCTGTACACCTAAACTGCTCTGGTGCCCAACGTGGAG	3 240
RA5-clone5	83:TCTCTTTACTTTGCCCAGTTTTGCTGTACACCTAAACTGCTCTGGTGCCCAACGTGGAG	142
RA5-clone6	67:TCTCTTTACTTTGCTCAGTTTTGCTGTACACCTAAACTGCTCTGGTGCCCAACGTGGAG	126
RA5-clone8	67:TCTCTTCACTTTGCTCAGTTTTGCTGTACACCTAAACTGCTCTGGTGCCCAACGTGGAG	126
RAS-clonel1	26:TCTCTTTACTTTGCTCAGTTTTGCTGTACACCTAAACTGCTCTGGTGCCCAACGTGGAG	85
HERVK-LTR-Leader	#929:agaaacacccacaggtgtgtaggggcaacccaccctacatctggtgcccaacgtggag	g #988
chr_7_NW_00183907	167:TCTCTTTACTTTGCTCAGTTTTGCTGTACACCTAAACTGCTCTGGTGCCCAACGTGGAG	226
RA5-clone4	241:CTTTTCTCTAGGGTGAAGGTACGCTCGAG	269
RA5-clone5	143:CTTTTCTCTAGGGTGAAGGTACGCTCGAG	171
RA5-clone6	127:CTTTTCTCTAGGGTGAAGGTACGCTCGAG	155
RA5-clone8	127:CTTTTCTCTAGGGTGAAGGTACGCTCGAG	155
RAS-clonel1	86:CTTTTCTCTAGGGTGAAGGTACGCTCGAG	114
HERVK-LTR-Leader	∲989:cttttctctagggtgaaggtacgctcgag	<i>4</i> 1017
chr_7_NW_00183907.	227 : CTTTTCTCTAGGGTGAAGGTACGCTCGAG	255

Figure W1. Nucleotide sequences of the 5' RACE clones obtained from MeWo cells. (A) Clones indicating transcriptional initiation at nucleotide no. 796 are aligned with the LTR-leader sequences (HERVK-LTR-Leader) of HML-2.HOM (GenBank AF074086.2). Nucleotides identical to the HERVK-LTR-leader sequences are in blue, and mismatches in magenta. The 3' end of the LTR is marked by a line. (B) Clones corresponding to the transcripts starting at no. 826 are shown. The first 6 clones of 11 are shown. (C) Alignment of the 5' RACE products indicating transcriptional initiation by a non-LTR promoter at 7q22. Corresponding chromosomal sequences (NW 001839071) are also aligned. Regions unrelated to HERV-K are in magenta.



Figure W2. Nucleotide sequences of the 5' RACE products obtained from HEK293 cells. (A) The 5' RACE clones starting at no. 460 are aligned with the HERV-K LTR-leader sequences. These clones showed 44 nucleotide mismatches (only 92% sequence identity) compared with the 570 nucleotides of HML-2. HOM. A BLAST search identified the template sequences at a unique locus on chromosome 3 (NT_005612). Nucleotides identical to the HERVK-LTR-leader sequences (GenBank AF074086.2) are in blue and mismatches are in magenta. The 3' end of the LTR is marked by a line. (B) Clones starting at no. 813 or no. 815 are aligned. They show approximately 99% sequence identity to the HERV-K LTR-leader and thought to represent the transcripts from HML-2.HOM and related HERV-K loci.



Figure W3. Transactivation of HERV-K LTR by MITF-M. Luc expression assay with HKLTR-1, Del-7, and 7-LTR in combination with the MITF-M expression vector. Relative luciferase activity (RLA) (mean \pm SD) is presented in relation to the control assay (1.00) with HKLTR-1 and the empty vector (Vec). The ratio (0.5, 1) of the activator plasmid amount to the Luc plasmid amount is given in parentheses. Experiments were performed two times in triplicate.

Chromo- some	Genomic Contig	Position	Inr460	MITF-3	MITF-2	MITF-1	793TATA	Inr826	Identity (≧95%)	Gaps	Score (bits)
7	Template	#1	GGAATGC	CACATG	СТТСТС	CACATC	ΑΑΤΑΑΑΤΑ	GGATCCT			
_	HKLIR-I	<i>".</i>			1				044/045	1/045	
7	NT_079592.2	4636386	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	944/945 (99%)	1/945 (0%)	1696
		4627883	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AGTAAATA	GGATCCT	942/945 (99%)	(0%)	1687
		4619379	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	938/945 (99%)	1/945 (0%)	1669
		16240452	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	934/945 (98%)	1/945 (0%)	1651
		23081303	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	933/945 (98%)	1/945 (0%)	1647
12	NT_029419.11	20865515	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	936/945 (99%)	1/945 (0%)	1660
		20874004	GGAATGT	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	932/945 (98%)	1/945 (0%)	1642
		13992344	GGAATGT	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGCTCCT	929/945 (98%)	1/945 (0%)	1629
		17870521	GGAATGT	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/946 (98%)	1/945 (0%)	1622
		18938415	GGAATGT	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	925/945 (97%)	1/945 (0%)	1609
7	NT_007741.13	2904345	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	944/945 (99%)	1/945 (0%)	1660
7	NT_079596.2	57285861	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	935/945 (98%)	1/945 (0%)	1656
		24287145	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
		22844882	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		25232346	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	926/946 (97%)	2/946 (0%)	1609
		3784161	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	921/946 (97%)	3/946 (0%)	1582
		188057	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	906/946 (95%),	2/946 (0%)	1519
		12386954	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	906/946 (95%)	11/946 (1%)	1517
8	NT_007995.14	7372243	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	934/945 (98%)	1/945 (0%)	1651
		13916038	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	912/945 (96%)	1/945 (0%)	1552
		12972195	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	903/947 (95%)	3/947 (0%)	1499
Y	NT_011903.12	1088725	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	934/945 (98%)	1/945 (0%)	1651
		2722474	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	934/945 (98%)	1/945 (0%)	1651
		3337067	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	934/945 (98%)	1/945 (0%)	1651
19	NT_011109.15	10625035	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	934/945 (98%)	1/945 (0%)	1651
		17365636	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	926/945 (97%)	1/945 (0%)	1615
		10090511	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/959 (97%)	15/959 (1%)	1608
		397710	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/972 (95%)	28/972 (2%)	1581
		24814575	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	914/946 (96%)	2/946 (0%)	1555
		10390031	GGAATGC	CACGTG	C <u>t</u> tgtg	CATAT <u>C</u>	AATAAATA	AGATCCT	909/945 (96%)	7/945 (0%)	1539
		21661083	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	905/945 (95%)	6/945 (0%)	1521
		1449405	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	AGATCCT	904/946 (95%)	2/946 (0%)	1510
		24680600	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	902/945 (95%)	3/945 (0%)	1499
10	NT_008705.15	9157724	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	933/945 (98%)	1/945 (0%)	1647
16	NT_037887.4	5745357	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	933/945 (98%)	1/945 (0%)	1647
		8169731	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	925/945 (97%)	1/945 (0%)	1611
8	NT_008183.18	6798560	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	933/945 (98%)	1/945 (0%)	1647
		9966476	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	911/945 (96%)	11/945 (1%)	1543
4	NT_006216.14	1502918	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	933/945 (98%)	1/945 (0%)	1647
		6812371	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	933/945 (98%)	1/945 (0%)	1647
		9985448	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	911/945 (96%)	11/945 (1%)	1543
16	NT_010498.15	28448457	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	932/945 (98%)	1/945 (0%)	1642
		1512569	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	928/945 (98%)	1/945 (0%)	1624
1	NT_004487.18	6639369	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	932/945 (98%)	1/945	1642

Table W1. Human Genomic LTRs with Less Than 95% Sequence Identity to the HERV-K LTR (HML-2.HOM).

					1			-	000 /045	4 /0.45	
		6087779	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
		6095991	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
		6060651	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	906/946 (95%)	2/946 (0%)	1519
		10227941	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	AGATCCT	903/945 (95%)	1/945 (0%)	1512
2	NT_022184.14	6498779	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	932/945 (98%)	1/945 (0%)	1642
		16268394	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/945 (98%)	1/945 (0%)	1638
		9653238	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
		65304362	GGCATGC	CACGTG	<u>gg</u> catg	CACAT <u>C</u>	AATAAATA	GGATCCT	899/946 (95%)	2/946 (0%)	1487
Х	NT_011786.15	9124856	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/945 (98%)	1/945 (0%)	1638
4	NT_016354.18	86128085	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/945 (98%)	1/945 (0%)	1638
		44812799	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	3/945 (0%)	1626
11	NT_033899.7	5136708	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/945 (98%)	1/945 (0%)	1626
		5128210	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	12/945 (1%)	1629
		22163299	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	909/946 (96%)	3/946 (0%)	1528
		22155106	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	908/946 (95%)	4/946 (0%)	1521
8	NT_023736.16	7343774	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/945 (98%)	1/945 (0%)	1638
		7352269	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	923/945 (97%)	9/945 (0%)	1604
3	NT_029928.12	275097	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	931/945 (98%)	1/945 (0%)	1638
3	NT_022517.17	47241028	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
		14073653	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	928/945 (98%)	1/945 (0%)	1624
		53952548	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	919/945 (97%)	1/945 (0%)	1584
		50497327	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	CATAAATA	GGATCCT	909/945 (96%)	12/945	1535
		23527120	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	900/946 (95%)	8/946 (0%)	1489
7	NT_007933.14	50045767	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
		48604326	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	928/945 (98%)	1/945	1624
		50991790	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	926/945 (97%)	1/945 (0%)	1615
		29571945	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	923/945 (97%)	2/945 (0%)	1599
		25977162	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	906/946 (95%)	2/945 (0%)	1519
		38166108	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	906/946 (95%)	11/946 (1%)	1517
17	NT_010799.14	3765019	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945	1633
		9204418	GGCATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	913/949 (96%)	5/949 (0%)	1544
9	NT_008470.18	31512589	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945	1633
		18678098	GGAATGT	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	919/945 (97%)	9/945	1586
6	NT_007299.12	26911478	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945	1633
		17388676	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	927/943 (98%)	1/943	1626
		16256254	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	925/945 (97%)	9/945 (0%)	1613
		16247792	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	924/945 (97%)	9/945 (0%)	1609
		31703257	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	920/944 (97%)	3/944 (0%)	1588
20	NT_011362.9	5652451	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	930/945 (98%)	1/945 (0%)	1633
12	NT_009775.16	1577352	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		9113298	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	909/946 (96%)	11/946 (1%)	1532
14	NT_026437.11	46446026	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		36491018	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	910/945 (96%)	15/945	1522
		59261730	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	902/946 (95%)	10/946	1503
		19587055	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AGTAAATA	GGATCCT	902/946 (95%)	2/946	1501
		4194701	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ	GGATCCT	900/945 (95%)	1/945	1498
3	NT_005612.15	91776453	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ	GGATCCT	929/945	1/945	1629
		91784665	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	928/945	1/945	1624
		19239229	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ	GGATCCT	927/945	9/945	1622
									(00/0)	(0/0/	

Table W1. (continued).

		19247428	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	927/945 (98%)	9/945 (0%)	1622
		82119455	GGCATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	915/945 (96%)	1/945 (0%)	1566
		93106159	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	910/945 (96%),	2/945 (0%)	1539
		93100773	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	905/946 (95%)	10/946 (1%)	1517
		7918791	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	904/946 (95%)	2/946 (0%	1510
		7914037	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	903/947 (95%	4/947 (0%)	1496
1	NT_004559.13	729687	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		4259163	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	901/946 (95%)	2/946 (0%)	1496
1	NT_032977.8	45814690	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		45820095	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		63715225	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	928/945 (98%)	1/945 (0%)	1624
		22443836	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	920/945 (97%)	2/945 (0%)	1584
		36863376	GGAATGC	CA <u>t</u> gtg	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	917/945 (97%)	9/945 (0%)	1577
		15965693	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	907/943 (96%)	2/943 (0%)	1532
		59505425	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	903/946 (95%)	10/946 (1%)	1508
5	NT_023133.12	895242	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		903454	GGAATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	929/945 (98%)	1/945 (0%)	1629
		23751517	GGCATGC	CA <u>t</u> gtg	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATG	GGATCCT	920/945 (97%)	1/945 (0%)	1588
		14234396	GGAATGT	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	917/945 (97%)	9/945 (0%)	1577
		25063855	GGCATGC	CACGTG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	GGATCCT	898/945 (95%)	5/945 (0%)	1483

Table W1. (continued).

Sequences identical to the corresponding positions of the HERV-K LTR (HML-2.HOM) are in black. Altered sequences that seemed to preserve the proposed functions are in blue. Nucleotide substitutions expected to disrupt the function are in red.

		Inr460	MITF-3	MITF-2	MITF-1	793TATA	Inr826				
Template HKLTR-1	#1	GGAATGC	CACATG	C <u>t</u> tgtg	CACAT <u>C</u>	AATAAATA	Insertion of Inr consensus, PyPyAN(T/A)PyPy	GGATCCT			
<i>M. mulatta</i> genomic reference	Position								Identity (≧80%)	Gaps	Score (bits)
AC210649.4	68871	G——GC	CACATG	CCTATT	CACATT	AATAAATA	CCATT <u>G</u> C	GGGTCCT	800/957 (84%)	30/957 (3%)	996
AC206102. 3	78780	GGTCT	CACAT <u>T</u>	CCTATT	CACAT <u>C</u>	AATAAATA		GGGTCCT	749/922 (82%)	27/922 (2%)	836
AC213330. 3	173369	GGTCT	CACAT <u>t</u>	CCTATT	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ		GGGTCCT	749/922 (82%)	27/922 (2%)	836
AC191812. 4	161342	<u>A</u> gaa <u>g</u> gc	CACATG	CCTATT	CACTTC	AATAAATA		GGGTCCT	774/967 (81%),	47/967 (4%)	836
AC197502. 3	156174	G——GC	<u>t</u> acatg		TTCATC	AA <u>C</u> AAATA		CCATCCT	775/962 (81%)	50/962 (5%)	834
	165784	G——TC	CACATG	CCTATT	CACTTC	A <u>gc</u> aaata		AGGTCCT	756/950 (80%)	46/950 (4%)	821
AC218138. 3	19882	<u>A</u> gaa <u>g</u> gc	CACATG	CTATTG	CACTTC	ΑΑΤΑΑΑΤΑ		CGAGTCC	776/968 (81%)	35/968 (3%)	827
AC209790. 2	102901	<u>A</u> gaa <u>g</u> gc	CCTATT	CCTATT	CACTTC	AATAAATA		CGAGTCC	776/968 (81%)	35/968 (3%)	827
AC201622.7	131721	G——GC	CACATG	CCTATT	CACAT <u>C</u>	AATAAATA			687/831 (83%)	34/831 (4%)	827
AC148674.1	31678	GGCC	CACATG	CCTATA	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ		GGGTCCT	781/980 (80%)	45/980 (4%)	821
AC194579.7	213314	GG-CGGC	CACGT <u>T</u>	CCTATG	CACAT <u>C</u>	AATAAATA	TCAGT <u>G</u> C	AGGTCCT	782/980 (80%)	48/980 (4%)	820
	195164	GG-CGGC	CACGT <u>T</u>	C <u>C</u> TATG	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ	CCAGT <u>G</u> C	GG-TCCT	780/981 (80%)	53/981 (5%)	807
AC191828. 3	148275	GG-CGGC	CACGT <u>T</u>	C <u>C</u> TATG	CACAT <u>C</u>	AATAAATA	TCAGT <u>G</u> C	AGGTCCT	782/980 (80%)	48/980 (4%)	820
AB128049.1	2710708	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	ΑΑΤΑΑΑΤΑ		GGGTCCT	771/970 (80%)	59/970 (6%)	820
	3070098	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	AATAAATA		GGGTCCT	770/967 (80%)	55/967 (5%)	816
	3159304	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	AATAAATA	CCAG <u>C</u> TT	GGGTCCT	766/967 (80%)	54/967 (5%)	801
	2817651	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	AATAAATA	CCAGT <u>G</u> C	GGGTCCT	766/967 (80%)	57/967 (5%)	794
AC199608.6	115869	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	AATA <u>C</u> ATA	CCAGT <u>G</u> C	GGGCCCT	769/967 (80%)	50/967 (5%)	816
AC210647.2	87803	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	ΑΑΤΑΑΑΤΑ		GGGTCCT	770/969 (80%)	60/969 (6%)	814
AC148684. 1	87804	GGAATGT	CATATG	CCTATT	CACTTC	AATAAATA		GGGTCCT	770/969 (80%)	60/969 (6%)	814
AC148680. 1	161314	<u>A</u> gaa <u>g</u> gc	CATATG	CCTATT	CACTTC	ΑΑΤΑΑΑΤΑ		GGGTCCT	770/969 (80%)	60/969 (6%)	814
AC193058.7	104908	G——GC	CACAT <u>C</u>	CCTATT	CACAT <u>C</u>	ΑΑΤΑΑΑΤΑ	CTAGT <u>G</u> C	GGGTC—A	778/981 (80%)	62/981 (6%)	801
AC148690. 1	2395	GGAAT <u>AT</u>	CATATG	CCTATT	CACTTC	ΑΑΤΑΑΑΤΑ	CCAG <u>C</u> TT	GGGTCCT	766/967 (80%)	54/967 (5%)	801
	94162	AGAAGGC	CATATG	CCTATT	CACTTC	AATAAATA	CCAG <u>C</u> TT	GGGTCCT	766/967 (80%)	54/967 (5%)	801
AC148682.1	106447	AGAAGGC	CATATG	CCTATT	CACTTC	AATAAATA	CCAG <u>C</u> TT	GGGTCCT	766/967 (80%)	54/967 (5%)	801
	198215	AGAAGGC	CATATG	CCTATT	CACTTC	AATAAATA	CCAG <u>C</u> TT	GGGTCCT	766/967 (80%)	54/967 (5%)	801
AC225834. 3	157305	AGAAGGC	CATATG	CCTATT	CACTTC	ΑΑΤΑΑΑΤΑ		GGGTCCT	765/965 (80%)	51/965 (5%)	800
AC193521.7	79951	GGCTTCT	C <u>t</u> catg	CCTATA	CACAT <u>C</u>	A <u>G</u> TAAATA	CCAGTCC	GGGTCCT	775/977 (80%)	46/977 (4%)	798
AC214376. 3	140057	AGAAGGC	CACATG	CCTATT	CACTTC	AATAAATA		GGGTCCT	755/951 (80%)	56/951 (5%)	792
AC210118.6	102633	AGAAGCC	CATATG	CTATTA	CACTTC	AATAAATA		GGGTCCT	765/966 (80%)	49/966 (5%)	791
AC148703. 1	4548	AGAAGGC	CATATG	CCTATT	CACTTC	AATAAAT <u>C</u>		GGGTCCT	765/967 (80%)	53/967 (5%)	791
AC148694. 1	141943	AGAAGGC	CATATG	CCTATT	CACTTC	AATAAAT <u>C</u>		GGGTCCT	765/967 (80%)	53/967 (5%)	791
AC191957.7	14747	AGAAGGC	CATATG	CCTATT	CACTTC	AATAAATA		GGGTCCT	733/927 (80%)	49/927 (5%)	756

Table W2. Rhesus Monkey (Macaca mulatta) LTRs with Less Than 80% Sequence Identity to the Human-type HERV-K LTR.

Sequences identical to the corresponding positions of the HERV-K LTR (HML-2.HOM) are in black. Altered sequences that seemed to preserve the proposed functions are in blue. Nucleotide substitutions expected to disrupt the function are in red.