

LA1-ori	-----GAAAAATTT-----TAAATTTCATATAAAGCTCCCG	29
LAlus4	-----GAAAAATTT-----GAATAATCATATAACTCCCG	29
Labarr1	CACGTAGC <u>TTTATTAATTAAT</u> ATGCTACGTGCCTATCATCTCAAAGTCCTATTGATCCCC	60
	:~::~	
	5' conserved	
LA1-ori	ATGCTAAGATTTTGTTACTAAAAAAGCTCAAGATAAAATCGTCTGATCTATTCTCATTTGT	89
LAlus4	ATGCTTAGATTCGTTACCAAAAAGCTCTCAAGACAAGTCCTCGATTTATTTTCTATTGGT	89
Labarr1	ATGTTTCAGATTTTATTAACAGAAAAACTACG--CAAGCTAATACATCATCCCTATCCAG	117
	*** * ~~~~~ . *** . ~~~~~ : ~~~~~ . * . * . . * . * : * ~~~~~ .	
	Start of Gag and Gag-Pol	
LA1-ori	TCTGATCGCGGAACTTTTGTTGCCATAATAGAGTTCGGACTGATTTCAAGTTTGACAAC	149
LAlus4	TCTGACAAAGGTACTTTTGTGCGCATAATAGGGTGAGGACTGACTTTAAATTTGACAAC	149
Labarr1	TGGGATCAGGGAACTTTACGATATACGGTAGGAATCGTATTGACTTCAAGTATGATGGC	177
	* ~~~~~ . ** ~~~~~ . . ~~~~~ . * ~~~~~ * ~~~~~ * ~~~~~ . ~~~~~ . *	
LA1-ori	TTAGTATTCACCAGAGTTTATGGTGTCTCAAAAATTTACGTTAGTCGGTAACCCGACA	209
LAlus4	TTAGTTTTTAAATCGAGTCTACGGCGTGTCACAAAGTTCACCCTAGTCGGCAACCCGAAA	209
Labarr1	CTAACTTTTTCTAGGCAATTAGGAGTTTCCCAGAAATTTACTTTGCTTGGTAACCCGGAT	237
	** . : ** . . * . : * : ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ . * ~~~~~ * ~~~~~ . :	
LA1-ori	GTCTGCTTTAATGAAGGCAGTTCTTACCTAGAAGGTATTGCTAAAAAGTACCTAACTTTG	269
LAlus4	GTCTGTTTCAATGAGGGAAGTTCGTACTTGGAGGGTATTGCAAGAAATATCTGACCTTG	269
Labarr1	GTCACCAAAGTAGTCGGGCTCTACTTAGATGGTATTGCGAAGAAGTATTTAACTCTT	297
	*** . : * ~~~~~ . * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ *	
LA1-ori	GATGGAGGACTTGCCATTGACAATGTCCTCAATGAGTTGAGATCCACTTGC GGATCCCA	329
LAlus4	GACGGTGGCCTGGCCATTGATAACATCCTCAATGAATTGAAATCAACATGTGGTATACCA	329
Labarr1	GAGGGAGGACTAGCTATTGAGAATGTCATGTCTGAATTACGACGTAACACGGGGATACCT	357
	* ~~~~~ . : ** * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ :	
LA1-ori	GGTAATGCTGTTGCATCTCATGCGTATAATATTACATCCTGGCGTTGGTATGACAATCAC	389
LAlus4	GGTAATGCTGTGCTTCTCATGCATACAACATTACGTCCTGGAGATGGTACGATAATCAC	389
Labarr1	GCCTCAGCCGTATCAGCTCAGCGTTTAACATAGCATCTTGGCGCTGGTACGATAACCAC	417
	* . : ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~::~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ * ~~~~~ *	
LA1-ori	GTGGCACTGTTGATGAACATGTTGCGTGCTTACCACCTACAAGTATTGACCGAACAGGGC	449
LAlus4	GTGGCATTACTCATGAATATGTTACGTGCGTACCATCTTCAAGTGTTATCTGAACAAGGC	449
Labarr1	GTAGCTTTATTAATTAATATGCTACGTGCTATCATCTCAAAGTCCTAGACGAGAGGGGT	477
	* ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ *	
LA1-ori	CAATATAGCGCTGGAGATATCCCTATGTACCATGATGGACATGTCAAAATCAAGCTACCA	509
LAlus4	CAGTATTCTGCCGGTTCATACCTATGTACCACGATGGGATGTCAAAATTAACCTTGAC	509
Labarr1	GAGTTTGGTGGCGGCCACTACCTTAAGTATGACGATGGGCATCTGGCCATTACTGTAGTG	537
	* . : ~~~~~ * ~::~ . : ~~~~~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ :	
LA1-ori	GTGACTATCGATGACACGGCAGGCCAACACAATTTCGTTGGCTAGTGACAGGTCTACT	569
LAlus4	ACTCCTATCAGTGAGGATGATGCTCCTGACTCATTTAAGTGCGCGAGTGACCGGACAACA	569
Labarr1	CCACCTACTGGTGCTTG--GAAACTGACCACTTTAGATGGCATTCAAAGAGGTTGAA	594
	. *** ..** . * . : ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ :	
LA1-ori	GATTCGTATCCTGATTGGGCACAGTTTTCTGAATCATTTCCATCAATCGACGTCCC GTAC	629
LAlus4	GACACTTACCAGACTGGGCTCAGTTCTCAGAGTCCTTCCC GTCTATTGACGTCCC ATAC	629
Labarr1	GTGCTCGCACCAAGAGTGGGCCCACCACTAGAGGCCATGCCTACAATAGAGGTGCCTTAC	654
	* . : ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ *	
LA1-ori	CTAGATGTTAGGCCATTGACCGTAACGGAAGTCAATTTCTGCTTTATGATGATGAGTAAG	689
LAlus4	TTAGATGTCAGGCCCTTACTGTTACCGTAAGTCAATTTCTGTTTATGATGATGAGTAAG	689
Labarr1	ATCGATGTCGCCCCTTAACACTAGGTTGAAGTCAATTTCTGTTTATGATGATGATGTAAG	714
	* ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ * ~::~ *	
LA1-ori	TGGCATAGACGTACTAACTTAGCGATAGACTACGAGGCACCCCAACTAGCTGATAAGTTC	749
LAlus4	TGGACCGTCGGACTAACCTTGCTATTGACTATGAAGCGCCGGCGCTGGCAGACAAATTC	749

Labarr1	TGGTCACGCCAGACTAATCTTGCCATAGACTTCGATTACCTAACCTGACAGACCGGTT	774
LA1-ori	GCTTACCGCCATGCGCTTACTGTTCAAGACGCTGACGAGTGGATAGAAGGCATAGA	809
Lalus4	GCCTACCGTCACGCCATTACGGTCCAGGATGCAGATGAATGGATTGAAGGTGATAGGACA	809
Labarr1	GCCTATAGACACACCGCCAAGGTGACGGAAGCGGACGAGTGTTAATGAACGAAAGCGCT	834
LA1-ori	GATGACCAAGTTCGCCCCCCCTCGTCTAAAGTAATGTTATCGGCACCTTCGTAAGTACGTG	869
Lalus4	GATGATCAGTTTAGACCACCTTCTTCTAAGGTAATGATGTCGGCATTGCGTAAGTATGTT	869
Labarr1	GATGCAGCCTTTCCGGTACTGTCCAGTAAGGCTATGTTGTCTGCGATCAGGAAGTACGTG	894
LA1-ori	AACCATAACAGGCTGTACAATCAGTTTTACTGTCAGCACAACTGTTAGCTCAAATTATG	929
Lalus4	AATCAATAATCGACTTTACAACCACTTCTATACGGCCGCTCAACTATTAGCTCAGATTATG	929
Labarr1	AATCACAAACGAGTGTAACAATCAGTTTGAACATGCGGCGCAGACTATTAGCACGCTGATG	954
LA1-ori	ATGAAACCTGTCCCTAACTGCGCTGAGGGCTACGCTTGGCTGATGCATGACGCATTGGTC	989
Lalus4	ATGAAACCAAGTACCTAACTGCGCTGAAGGGTATGCATGGTTGATGCACGATGCTCTGGTC	989
Labarr1	GTTAAGCCTTTGCCGAATAACGCTGAAGGGCACGCATGGTTACTACATGACCCGGTTGTG	1014
LA1-ori	AATATACCAAATTTGGGTCTATTCGAGGAAGGTACCCCTTTTGTATCAGGTGATGCA	1049
Lalus4	AACTTGCCGAAGTTCGGGTCCGTTAGAGGGCGTTACCCGTTCTTACTTGCAGGTGACGCT	1049
Labarr1	AATATACCTAAGTTCGGGTGAGTTAGGGGCAGATATCCGTTTTTACTCACGGGTGAAGCT	1074
LA1-ori	GCGTTGATTACAGGCTACAGCCCTAGAAGACTGGTCTGCTATCATGGCGAAACCCGAGCTG	1109
Lalus4	GCCCTGATTACAGGCAACAGCACTTGAAGACTGGTCAGCAATTATGGCTAAGCCAGAACTG	1109
Labarr1	TCTTTGGTGCAAGCAACAGCACTAGAAGATTGGTCAGCTATCATGGGTGAAGCCGGAGATA	1134
LA1-ori	GTGTTCACTTACGCGATGCAGGTGTCAAGTACGCTTAAACACCGGACTATACTTACGTCGC	1169
Lalus4	ATTTTACCTACGCTATGCAGGTTGCTGTGCGGTTGAACACTGGCTTATACCTTACGGAGG	1169
Labarr1	GTTTTTACCTATGTCAGTAGATGCTATGACAGTGAATACTGGCTTGACTTGAGGCGC	1194
LA1-ori	GTTAAGAAAACAGGCTTCGGGCACAACTATAGATGACAGCTATGAAGATGGAGCGTTTTTG	1229
Lalus4	GTCAGAAAACGTTTTCGGTACCAGTGTGACGATAGTTATGAAGACGGCGCATTTCTA	1229
Labarr1	ATCAAGAAAACGGGAATGGGCACCAAGGGTTGATGATAGCTACGAAGATGGTGTCTTCCCT	1254
LA1-ori	CAACCGGAGACGTTCTGTTACAGCCGCACTAGCATGTTGTACCGGACAAGATGCGCCCTTA	1289
Lalus4	CAACCTGAGACCTTCGTACAAGCTGCGATAGCTTGTGCACTGGCCAAGATGCACCGCTT	1289
Labarr1	TCACCTGAGACGTTTCATGGCTGACGAGTGCCTTGTGCCACCGAGAGGATGCACCTCTA	1314
LA1-ori	AATGGGATGTCAGATGTGTATGTCACTTATCCAGATCTTCTAGAATTTGATGCTGTTACA	1349
Lalus4	AACGGTATGTCGTATGTGTACGTGACATACCAGATCTCCTTGAGTTGATACTATAACA	1349
Labarr1	AACGGAATGTCAGACGCTACGTAATTTATCCTCAATTTGTGGATCTGCACACTGTGACT	1374
LA1-ori	CAAGTACCCATCACGGTCATTGAGCCCGCTGGCTATAACATTGTTGATGATCATTTAGTG	1409
Lalus4	CGTGTCCCTGTTACGGTACTAGAACCTGAAGGCTATAATATACAGGATGGCGCCTTGGAA	1409
Labarr1	TATATACCTGCCCAAGTTCGTGAAGCGGAAGGGTACAACGCTACTAGCGAGGCTCTTAA	1434
LA1-ori	GTTGTGGGTGATACCTGTGGCATGTTCAACATACATGATATTTCCAGTAGCTGCGTTTGAT	1469
Lalus4	GTGACAGGTGTTCCGATTGCCTGCTCGCCTTACATGATTTTCCCTGTGCGAGCTTTCGAT	1469
Labarr1	GTTTTGGGAGTACCTTTGGCTTGTGCTCGCAGTGTTACTCTACCTATGGCGGCATTTGAT	1494
LA1-ori	ACTGCAATCCTTACTGTGGGAATTTTGTCTTAAGGCTGCTAACAAGTATCTCCGTAAAG	1529
Lalus4	GAGCTTAACCCCTATTCTGGTAGTTTGTATTAAAGCTGCGCTGAAATACCTACGCAAA	1529
Labarr1	GATGCAAAACCATATTCGGGATCGTTCAAGATCGGCAAGCGGAAAGGTATGTGAGAGGC	1554
LA1-ori	GGTGCCGTGTATGATAAACTCGAAGCATGGAAGTTGGCCTGGGCACCTGAGGGTAGCCGGG	1589
Lalus4	GGTGCTCTGTATGCCAACTAGAGGCTGGAAGCTGGCCTGGGCTATGAGGATTGCAGGA	1589
Labarr1	AGGGCGGTTTATTACCACTTCGAGGCTGGAAGCTAGGCTGGGCGCGGAGAAATCGCTGGG	1614

LA1-ori TATGACACTCACTTCAAAGTGTATGGCGATACACACGGCTTAACTAAGTTCTATGCTGAC 1649
 LAlus4 TACGACACTGAATTTAAAGCTTGTGGCAACACGCATGGTTTATCCAAATTTTATGCTGAT 1649
 LAbarr1 TATGATACTCGCGTATCTATACCTGGCGACAATACCGGCTTATCTAAGCTATATGCAGAC 1674
 ** * * * * . . * . : . : . * * * * . * . . * * * * : * * * . * * * * : * *

LA1-ori AACGGTGACACATGGACACACATACCTGAATTTGCTACTGACGGTGACGTGATGGAAGTA 1709
 LAlus4 AATAGCGATAGTTGGACACACATACCGGAATTCGTTAGTGATGAAGAAGTTACGGAAGTG 1709
 LAbarr1 AATGCCGATTCTCGGACTCACATACCGGACTCATTTCTCAATCCTGAAGCAGAATTTATT 1734
 ** . * * : * * * * : * * * * * * * * * * * * . * . * : * * . * . : . : *

LA1-ori TTGCTTACTGCCATCGAACGCAGAGCTAGACATTTCTGTTGAACTACCTAGACTGAATTCA 1769
 LAlus4 TATGTAACATAATTTGAACGTAGAGCCAGACACTTTTGTGAGCTACCAAGACTAAATTCT 1769
 LAbarr1 TATGTGACTGATCTAGAGAGGAGGAAGGCACCTTTATAGACTTGCCGAACATGACCAAT 1794
 * : * * * * . . * * * . * * . * * * * * * * * * * * * . * * * * . * * . : . :

LA1-ori CCAGCATTCTTCAGATCTGTAGAAGTCAGCACCCTATATATGATACTCATGTGCAGGCT 1829
 LAlus4 CCAGCCTTTTAAACAGTAGAGTAAGCACCGCAATTTACGATACTTACGTACAGGCA 1829
 LAbarr1 CCTTCATTGCGAAGATCGATAGAAGTACAGATTACACTTAACGATATGTATGTCGAGTCT 1854
 * * : * * * * * * * * * * * * * * . * : . : * * * * * * * * * * * * :

LA1-ori GGTGCGCATGCGGTGTATCATGCTAGTGAATCAATCTTGATTATGTTAAGCCTGTTTCG 1889
 LAlus4 GGCACATTCTCTGTCTACCATGCAAGCAGGATTAATTTGGATTACGTCAAGCCTGTATCT 1889
 LAbarr1 ACTCCCGGGTCCCGCTCACGGCTATGGTGCTGAGCAGGGACTACGTCGCGCCGATAGCG 1914
 . * * * * * * * * * * * * * * . * * * * * * * * * * * * . * * * : *

LA1-ori ACCGGCATTACAGGTGATCAATGCGGGCGAACTTAAGAACTACTGGGGTAGTGTGCGTCGT 1949
 LAlus4 GCAGGCATACAAGTCATTAACGCTGGTGAACCTAAGGAATTACTGGGGTAGTGTACGCCGT 1949
 LAbarr1 TCTGGTATACAAGTGATTAGTGCCGCTGACTTGAGGAACCTATTGGGGCTCAGTCCGCCGG 1974
 *

$\cap$
-1 Frame-shift site Stem loop for frame-shift ($\Delta G = -14.0$ kcal/mol)

LA1-ori ACTCAGCAGGGGTTTAGGAGTGGTAGGTCTTACGATGCCAGCTGTAATGCCTACCGGAGAA 2009
 LAlus4 ACGCAGCAGGGGTTTAGGAGTGGTAGGTCTTACGATGCCAGCCGTAATGCCTACCGGAGAA 2009
 LAbarr1 TCGAATCGGGTTTATCCTTGGTAGGTTTTACGATGCCAGCAGTAATTCCTACCGAAAGG 2034
 : *

LA1-ori CCTACAGCTGGCGCTGCCACGAAGAGTTGATAGAACAGGCGGACAAATGTTTATGATAG 2069
 LAlus4 CGTACAGCTGGCACTGCCACGAAGACTTGATCGAACAGACGGAATGTTTCAGTCGAG 2069
 LAbarr1 AGTATCGCTGGTCAAGACCTCCAAGAGGCGATAGAAGTACGAGCAAGAGCCTTCAGTAGAA 2094
 . * * * * * * * : . : * * * * * * * * * * * * * * . * * * * * * * *

LA1-ori TAAACGTAATCGAACCTCACACGGACCCCGCCCTACAAGGTACATACTGCAGGAACCAG 2129
 LAlus4 TAAACATAATAGAACCATCCCATGGGCTCGTCCGATGAGATACATATTGAACGAGCCAA 2129
 LAbarr1 TAAATCTTATAGAAAGCGAGCACGGTCATAGGCCACGCAGATACATTGCTTCTGGTAATG 2154
 *

Stop of Gag

LA1-ori GTACGTACCCTGCGTGGATTAGATTCAGGAACAGAGTACAAGCTGTATCGCGTCAGAAAG 2189
 LAlus4 ATACATACCCCGCATGGATAAGGTTCAAGAAATAGAGTACAGGCGGTATCTAGACAGAGGG 2189
 LAbarr1 AAGTAGTCCCTATATGGATCAGATTCGGCCGCCGCATTGAGTGTGTTTCACGTCAGAAAG 2214
 . : . : : * * * . * * * * * * * * * * * * . * . * * * * * : * * * * *

LA1-ori CCACTCACTTCTGTTTGACATCGTACCTGCCGCAGTAATTAGTGATTTTACTACGTCTG 2249
 LAlus4 CAACACATTTCTTATTTGATATAGTACCGGCCAGCAAAATTGCGGACTATACTACTTCAC 2249
 LAbarr1 CAACTTACTTTCTGGCTGACGTTATCCCAGAGAAGAGGCTATAAACTACACCTCTGCTA 2274
 * . * * : * * * * . * * * * * * * * * * * . * * : * * : * * :

LA1-ori ACACGTCTTCGTTTGCATACAAATCGCACACCTACGCTGTAATGTAACAGCATTGAGGT 2309
 LAlus4 AAACGGCTACTTTTGCATACAGGTCGCACACGTATGCATGCACGGTAACGGCACTACGCT 2309
 LAbarr1 ACACAGCCCAAGTGGCCATAAGGAGAAGTACATATACGGCTAGCGGTATTTTATGTCAAAC 2334
 * . * * . * . * * * : : : . : . * * * * * * * * * * * * * * *

LA1-ori TCAGTGACACTTATGCCTTGTACGTACAGACTGATACCAACATGACAATTTTAAGCCAG 2369
 LAlus4 TTGGAGATGACTACGGCTTGTATGTCAGGTAGAGGCTAACATGACATTGTTAAGTCCGG 2369
 LAbarr1 TCGGAGACAACATATGGTTTCTACCTACAGTCGATGAAGATATGACAGTCAAAACGCCGG 2394
 * . * : * * . . * * * * * * * * * * . * * . . * * * * * * * * : : * * * *

LA1-ori CGGCGCGTCGCCAGGCTTCTGCGACGTACTCACAGGTGGCAGGGTTTTGTTATAACACAC 2429
 LAlus4 CTGCACGGCGACAAGCTTCTGCTACGTACTCTCAGGTTGAGGGGTTCTGTTTTAACACAC 2429
 LAbarr1 CCTTACGAAGGTACGCTTCGATGGCTTACTCTAACGTGCGCGGTTTCATGTTATAACACTC 2454
 * . * * . * * * * * * * . *

LA1-ori	CTACCGTTATGGATTTCGCTAGCGCAATATCTTGGACGTAGACGCCAATATACGACCCAAAC	2489
LA1us4	CTACGGTCATGGACACGCTAGCTAACATCTTGGATGTTGACCGCAATATTAGACCCAAAGC	2489
LA1barr1	CTACCATAATGGACACTCTGTGTAATATTTTTGACTACGACCCCGATTCCGCGGCCATCCT	2514
	**** * ***** : * ** * * * * : ***** : * : * : * : *	
LA1-ori	ACTTCAAGGGTTTACGGCTATACACAAGGTCTAAGGTCCTGCTCAACATCATACTCACT	2549
LA1us4	ACTTTAAAGGTCTGCGTACTTATGAGAGGTCTAAGGTCACAGCACAAACACACACTCACT	2549
LA1barr1	CATTTAAAGGCATGCGTCTTATACCAAGTCAAAGGTTACGGGGATGCACCATACCCATT	2574
	..** ** : * : * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	TGCGGCCAGACGAGCTAGTGGAAGCGGCGCAAAGGTCTCGCCTAGACGTAATACTACT	2609
LA1us4	TGCGGCGCTGATGAAGTACTGGAAGCAGCGCGCATGCTATCGCCTAGGCGAAAGTACTACT	2609
LA1barr1	TAAGGCCACAAGATGTTTTGTCGGCTGCCAAGAAGACCGGAAATAGTCATTATTACTACG	2634
	* : ***** : * ** * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	TAATGTGTGTAGTTGAGCTGCTCGCGAACTTACAAGTAGATCTTGAAGCAGCAGTAGCTA	2669
LA1us4	TATTATGCGTCGTAGAGCTATTAGCGGCATGCGAGGTTAACATTGAAGCAGCAGTAGCTA	2669
LA1barr1	TGCAGGTGGCTGTCGAGCGCTAGCATCGGCTGGCGTAACACAGAGGCAACTGTTTCAA	2694
	* : : * ** ***** : * : * : * : * : * : * : * : * : * : *	
LA1-ori	CTATTCTCGCATATGTCCTGACACTAAGTGAAAAATTTGTACCAATTTTCTTGGATTCTA	2729
LA1us4	CAATAATGACGTACGTACTIONTTTGGACGAAAAGTTTATACCTTTGTTTCTAGACTCAC	2729
LA1barr1	CCATAATGTTGTATGTCCTCACATCAAACGACACCACACGGGCATTGTTCTAGATTCTA	2754
	* * : * : * : * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	GAGCAATATGGGTGCGTGAGCCTGGGCGCTGATGCTCTGACTGCACGTCTCAAGGCCAGTA	2789
LA1us4	GGACAATATGGCAGGTAGTAAAGGACCTGAAGAATAACGACAGACTCAAGAAGGCGCA	2789
LA1barr1	AAGCATATATGCGGAGACTGACGGGAAGATTAATCTGACCGACTCAAGCGAGTAA	2814
	..* : ***** : * : * : * : * : * : * : * : * : * : *	
LA1-ori	GTGGGCAGATCAAGAGCATACACACGGCTGATTACGAACCACTCACTGAACATTTCGAGT	2849
LA1us4	GCGGGCAGATTAAAAGTGTTTCATACAGCTGACTATGAGCCACTGACGGAATGTTTGAAT	2849
LA1barr1	GCGGACAGATCAAGTCAATTCACACAAGTGAGCTGGAGCCTTTAACCGAGTTGTTTCGAGC	2874
	* ** * : ***** : * : * : * : * : * : * : * : * : * : *	
LA1-ori	TAGCAGTATTGATGAACCGAGGTGTTGGCCATGTCTCTTGGCAAGCTGAAAAGGATCATC	2909
LA1us4	TGGCTGTGCTCATGAACAGAGGCGTCGGGCATGTATCCTGGAAGACTGAAAGAGAACACA	2909
LA1barr1	TTCAGGTCTTGATGAACCGGGGTATAGGACATGTTGACTGGAATGGGGAAGACATAACA	2934
	* : * * * ***** : * : * : * : * : * : * : * : * : * : *	
LA1-ori	GCTTGAATCCCGACGTGGCTGTAGTTGATCAAGCACGGCTATATTCGTGTGTGCGCGACA	2969
LA1us4	GAGAGAATCCGGATGTAGCAAACGTTAATCAACAAGCACTCTACGCGTGCGTACGTGACA	2969
LA1barr1	GGACCAACCCTAATACTGTGAATGTAGATCAAAAGATGCTCTATAGTGTGTACGCGATA	2994
	* * * * * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	TGTTCCGAAGGATCAAAGCAGACGTATAAATATCCCTTTATGACGTGGGATGACTACACTG	3029
LA1us4	TGTTCCGAGGGGGCAAAGCAGACTTATGATTATCCTTACATGACATGGGATGACTATACGT	3029
LA1barr1	TGTTTGAAGATTCTAAGCAATCGTACAAATACCCGTATATGACATTGGAACGACTATGCGG	3054
	***** * : * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	CAAACAGATGGGAGTGGGTTCCAGGTGGCAGTGTCCACTCTCAATACGAAGAAGACAACG	3089
LA1us4	CAAGTCGGTGGGAATGGGTGCTGGAGGAAGTGTTCACTTCAGTACTCCGAGGATGATG	3089
LA1barr1	ATAGTAGGTGGGAATGGATACCTGGAGGTAGTGTACACTCGCAATACCCCTCAGGATGATG	3114
	..* : * : * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	ATTATATCTATCCTGGTCAGTATACTAGGAACAAGTTCATAACTGTTAACAAAATGCCCA	3149
LA1us4	AATATATATTTCCAGGACAATATACTAGGAACAAGTTTATTACAGTCAACAAAATGCCTA	3149
LA1barr1	AATACATAGTCCTGTGCTCTACACCAGGAATAAGTTTATAACAGTAAACCTCATGCCGA	3174
	* : * * * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	AACACAAAATATCTAGAATGATAGCATACCGCCTGAGGTACGAGCTTGGACGTCGACGA	3209
LA1us4	AGCATAAAGATCGCGCTATGATAGCATCTACGCCTGAAAGTTAGGGCATGGACCTCAACGA	3209
LA1barr1	AGAATAAGTTAGCCTCAATCGGCGCATCGAAACCAGAAGTCAGAGCATGGACATCAACTA	3234
	* : * * * : * : * : * : * : * : * : * : * : * : *	
LA1-ori	AGTACGAATGGGGCAAGCAACGTGCTATCTACGGACGCGATCTACGAAGTACACTGATAA	3269
LA1us4	AGTATGAATGGGGTAAACAAGAGCTATATGGCACAGACTACGGAGCAGCACTTATCA	3269
LA1barr1	AATACGAGTGGGGAAACAGAGAGCTATCTATGGAACATGCTGAGGAGCACTCTGATAA	3294
	* : * * * : ***** : * : * : * : * : * : * : * : * : *	
LA1-ori	CTAACTTTGCAATGTTCAAGTGCGAGGATGTTTCTCACTCACAAGTTCAGTAGGCGACC	3329
LA1us4	CTAACTTCGCTATGTTTCAGATGTGAAGATGTGTTAACACATAAAATTTCTGTAGGTGATC	3329

LA1-ori	AAATACCATA C GAAATAGATGATCCATCATTTTTGGCCAGGGGTAAATGATTATGCTTATA	4229
LAlus4	AAATACCATA C GAAATAGATGATCCATCATTTTTGGCCAGGGGTAAACGATTATGCTTATA	4229
Labarr1	AGATACCTGA A GAAATAGACGATCCGTCATTTTTGGCCAGGTGTTAACGATTATGCCAGGA	4254
	* . * * * * : * . *	
	<i>Stem loop for packaging ($\Delta G = -4.3$ kcal/mol)</i>	
LA1-ori	AAGTCTGGAAAAATTTTCGGAGAACGACTCGAATTTTAATAAGATTAAGATGCCGTAGCTA	4289
LAlus4	AAGTCTGGCAGAATTTTCGGCGAAAGGCTGGAGTTTAAACAAGATTAAGACGCTGTTTCAA	4289
Labarr1	AAGCATACAATATCTTAGGAGAGAACTAGAATTTAAACAAGATTAAGAGTGCAAGTAGCCA	4314
	** * . * . * . * . *	
LA1-ori	GAGGGAGTAGGAGCACTATAGCTCTGAAACGTAAGGCTAGGATAACATCTAAGAAGAATG	4349
LAlus4	AGGGCAGCAGAAACACGATAGCCTTGAAACGTAAGGCGAAGATTTCTGCAGTCAAGAACG	4349
Labarr1	GAGGTAGCAGGCTAACGATTGCCATGAAACGAAGAGCGCGAGTAGGATCGACGCCTACAG	4374
	. . ** * * * * . . * * * * : * * * * * * * * * * * * * * * * * * *	
LA1-ori	AATTGCTAACAAGTCGGAATGGGAAAGGACAATGTACAAAGCCTATAAGGGTTTGCGAG	4409
LAlus4	ACTTCTGCAATAAATCTGAATGGGAGAGAACTATGTACAAAGCTTACAAAGGTTTAGCAG	4409
Labarr1	AGTTTATTAATATCAAGCAATGGGAGCGTGCGATGTACAGAGCTTTTAAAGGGGTTGCGA	4434
	* * * . * * * : . *	
LA1-ori	TCTCATACTATGCTAACCTGAGCAAATTCATGAGTATACCACCAATGGCGAACATTGAAT	4469
LAlus4	TATCTTATTATGCTAACTTGAGCAAATTCATGAGCATACCACCCATGGCAAACATAGAAT	4469
Labarr1	TATCATATTACACCAGCTTGAGCAAATTTATGGCTGTACCACCTATATCAGGGCTTGATC	4494
	* . * * . * * * . *	
LA1-ori	TTGGGCAGGCTAGATATGCTATGCAAGCAGCCCTTGATAGTTCTGATCCACTCCGGGCAT	4529
LAlus4	TCGGACAAGCTAGATTTGCGATGCAGCGGCCCTTAGATAGTTCCGATCCTTTAAGAGCAC	4529
Labarr1	GAGGCGAGGGCGCCATGGCAGTTAACGCTGCTATTGCGAGCGCCGATCCTCTGCGGGCAC	4554
	** * . * . . . *	
LA1-ori	TACAGGTCATACTGT-AA TTGCCAAAAGATAATGGGAATTACCCAT ATGC-----	4579
LAlus4	TACAAATATTCTTA TGA AGTGCTCGAACGAT GAGGGTTTTTACCCAT ATGC-----	4580
Labarr1	TTCAAAATCTAGTA TAAT ACTAGGGGAAGAAG ATGGATTTAATCCAT AAGCATACCCACT	4614
	* : * * . * * * * * : : . * * * * * : : * * * * * * * *	
	<i>Stop of Gag-Pol Stem loop for replication ($\Delta G = -3.3$ kcal/mol)</i>	
LA1-ori	-----	
LAlus4	-----	
Labarr1	ATATCAGG 4622	

Figure S1. Multiple sequence alignment between ScV-LA1-original, ScV-LALus4, and TdV-LAbarr1 (+) strand nucleotide sequences (cDNA). 5'GAAAAA conserved motif (5' conserved), translation initiation (start of Gag and Gag-Pol, or internal ATG in Pol ORF of LA, LALus4 and LAbarr1) and termination (stop of Gag or stop of Gag-Pol) codons, ribosome frameshifting site (-1 frameshift site), frameshifting associated sequence (stem loop for frameshift), packaging signal (stem loop for packaging), and replication signal (stem loop for replication) are indicated and shaded in the nucleotide sequence. An AU-rich 15-nucleotide stretch located in the 5' untranslated terminal region is underlined. Λ , ribosomal frameshift. Asterisks (*), colons (:), and dots (.) indicate identical nucleotide positions, transitions, and transversions, respectively.

[illegible]

Figure S2. Comparison of the amino acid sequences of Gag-Pol encoded by ScV-LA1-original, ScV-LAlus4, and TdV-LABarr1 dsRNA genomes. The separation between Gag and Pol domains is indicated (Gag ◀▶ Pol). The H154 residue required for 5'cap-snatching is black shaded. The stretch of variable amino acid sequence located downstream from H154 is double underlined. The four crucial residues for cap recognition (Tyr-150, Asp-152, Tyr-452, and Tyr-538) are grey shaded. A variable region of 44 amino acids in the N-terminal region of Pol that is likely to separate Gag and Pol is indicated above the sequence. The highly conserved RdRp domain located in the central third of Pol is underlined, the four consensus motifs (A-D) conserved in RNA-dependent RNA polymerases from totiviruses are indicated above the sequence, and the conserved amino acids for each motif are grey shaded. Methionine (M) from in-frame re-initiation codons

(AUG) in the N-terminal region of Pol are shown in bold. Asterisks (*) indicate identical amino acids; colons (:) and single dots (.) indicate conserved and semi-conserved amino acids, respectively.

TdV-LA1082	100	61	62	62	62	62	62	62	62	62	62	61	63	62	62	62	62	62	63
ScV-LA1-original	61	100	73	74	73	73	73	74	74	74	74	74	74	74	73	74	73	73	74
ScV-LA2-8F13	62	73	100	74	74	74	74	74	74	77	75	74	74	74	74	75	74	75	74
ScV-LA2-EX1125	62	74	74	100	78	78	78	78	78	75	76	76	76	76	76	76	75	76	76
ScV-LA1-EX231	62	73	74	78	98	100	100	98	99	75	77	75	76	76	76	76	76	76	76
ScV-LA10560	62	73	74	78	98	100	100	98	99	75	77	75	76	76	76	76	76	76	76
ScV-LA10561	62	74	74	78	98	98	98	100	99	75	77	75	76	76	76	76	77	76	76
ScV-LA10562	62	74	74	78	98	99	99	99	100	75	76	75	76	76	76	76	76	76	76
SpV-LA45	62	74	76	75	75	75	75	75	75	100	77	77	77	76	76	76	77	76	77
SkV-LAFM1183	62	74	75	76	76	77	77	77	76	77	100	79	79	80	79	79	79	79	79
SpV-LA4650	61	74	74	76	75	75	75	75	75	77	79	100	90	90	90	90	82	82	83
SpV-LA62	63	74	74	76	76	76	76	76	76	77	79	90	100	92	90	91	82	82	83
SpV-LA1143	62	74	74	76	76	76	76	76	76	76	80	90	92	100	90	92	82	82	82
SuV-LA10560	62	73	74	76	76	76	76	76	76	76	79	90	90	90	100	92	82	82	82
SpV-LA74	62	74	74	76	76	76	76	76	76	76	79	90	91	92	92	100	93	82	82
SpV-LA1939	62	73	75	76	76	76	76	76	76	77	79	90	91	92	92	93	100	82	82
SpV-LA28	62	73	74	76	76	76	76	77	76	76	79	82	82	82	82	82	100	90	90
SpV-LA66	62	74	75	75	76	76	76	76	76	77	79	82	82	82	82	82	90	100	90
SpV-LA21	63	74	74	76	76	76	76	76	76	77	79	83	83	82	82	82	90	90	100

Colours: *S. cerevisiae* cluster Close to *S. cerevisiae* cluster
S. paradoxus cluster Close to *S. paradoxus* cluster
Equally distant from both clusters (74-75% identity)
Equally distant from both clusters (73-74% identity)
Equally distant from both clusters (61-63% identity)

Figure S3. Percentage identity matrix for canonical V-LA nucleotide sequences. Each value was rounded to the nearest whole number.

TdV-LAbarr1	100	83	82	84	85	85	85	85	84	83	83	84	84	84	83	84	84	84	84	85
SkV-LA1082	83	100	93	93	93	93	93	93	93	93	93	92	93	93	93	93	93	93	92	92
ScV-LA1-original	82	93	100	95	95	95	95	95	95	94	94	94	94	94	93	94	94	94	94	94
ScV-LA2-8F13	84	93	95	100	96	96	96	96	96	94	95	94	95	95	94	95	95	95	94	95
ScV-LA1-EX231	85	93	95	96	100	100	100	100	100	95	94	94	95	95	94	95	95	95	96	96
ScV-LAlus-EX1160	85	93	95	96	100	100	100	100	100	95	94	94	95	95	94	95	95	95	96	96
ScV-LA2-EX1125	85	93	95	96	100	100	100	100	100	95	94	94	95	95	94	95	95	95	96	96
ScV-LAlus1-EX436	85	93	95	96	100	100	100	100	100	95	94	94	95	95	94	95	95	95	96	96
LAlus4-EX229	84	93	95	96	100	100	100	100	100	94	94	94	94	94	94	94	94	95	95	95
SpV-LA45	83	93	94	94	95	95	95	95	95	100	96	96	99	95	96	95	95	96	96	96
SkV-LAFM1183	83	93	94	95	94	94	94	94	94	96	100	97	97	97	99	97	97	97	96	97
SpV-LA28	84	92	94	94	94	94	94	94	94	96	97	100	99	99	99	98	98	98	98	98
SpV-LA66	84	93	94	95	95	95	95	95	94	96	97	99	100	99	98	98	98	99	99	98
SpV-LA21	84	93	94	95	94	95	95	95	94	95	97	99	99	100	99	99	99	99	99	99
SpV-LA1939	83	93	93	94	94	94	94	94	94	96	99	99	98	99	100	99	99	99	98	98
SpV-LA62	84	93	94	95	95	95	95	95	94	95	97	98	98	99	99	100	100	100	99	99
SpV-LA4650	84	93	94	95	95	95	95	95	94	95	97	98	98	99	99	100	100	100	99	99
SpV-LA74	84	93	94	95	95	95	95	95	95	96	97	98	99	99	99	100	100	100	99	99
SpV-LA1143	84	92	94	94	96	96	96	96	95	96	96	98	99	99	98	99	99	99	100	99
SuV-LA10560	85	92	94	95	96	96	96	96	95	96	97	98	98	99	98	99	99	99	99	100

Colours: *S. cerevisiae* cluster Close to *S. cerevisiae* cluster
S. paradoxus cluster Close to *S. paradoxus* cluster
Equally distant from both clusters (93-95% identity)
Equally distant from both clusters (92-93% identity)
Equally distant from both clusters (82-85% identity)

Figure S4. Percentage identity matrix for highly-conserved RdRp-domain amino-acid sequences. Each value was rounded to the nearest whole number except for those between ScV-LA1-EX231, ScV-LAlusA-EX1160, ScV-LA2-EX1125, and ScV-LAlus1-EX436 of the *S. cerevisiae*-cluster, and SpV-LA62-vs-SpV-LA4650 of the *S. paradoxus*-cluster, which initially were 100%.

TdV-LAbarr1	100	47	45	47	45	45	44	44	46	45	44	46	44	43	44	46	46	46	46	46
SkV-LA1082	47	100	68	75	75	76	75	73	74	75	72	75	76	74	74	79	79	79	79	78
SpV-LA45	45	68	100	85	82	78	77	79	78	79	78	81	79	74	77	76	76	75	75	76
SkV-LAFM1183	47	75	85	100	87	86	85	85	83	85	82	87	84	76	76	79	79	80	80	79
SpV-LA62	45	75	82	87	100	94	92	94	95	95	89	92	92	75	78	79	79	79	79	78
SuV-LA10560	44	76	78	86	94	100	93	95	93	95	86	90	90	76	76	78	78	78	78	77
SpV-LA74	44	75	77	85	92	93	100	95	94	95	86	90	90	78	75	78	78	78	78	77
SpV-LA4650	44	73	79	85	94	95	95	100	95	98	88	90	92	76	76	78	78	79	79	78
SpV-LA1143	46	74	78	83	95	93	94	95	100	98	87	89	91	75	78	78	78	78	78	78
SpV-LA1939	45	75	79	85	95	95	95	98	98	100	87	89	91	76	78	79	79	79	79	78
SpV-LA66	44	72	78	82	89	86	86	89	87	87	100	92	93	74	73	76	76	76	76	75
SpV-LA28	46	75	81	87	92	90	90	90	89	89	92	100	95	78	77	79	79	79	79	78
SpV-LA21	44	76	79	84	92	90	90	92	91	91	93	95	100	78	78	82	82	82	86	81
ScV-LA1-original	43	74	74	76	75	76	75	76	75	76	74	78	78	100	76	80	80	80	80	81
ScV-LA2-8F13	44	74	77	76	78	76	75	76	78	78	73	77	78	76	100	85	85	83	83	84
ScV-LAlus4-EX229	46	79	76	79	79	78	78	78	78	79	76	79	82	80	85	100	100	98	98	99
ScV-LAlus1-EX436	46	79	76	79	79	78	78	78	78	79	76	79	82	80	85	100	100	98	98	99
ScV-LA1-EX231	46	79	75	80	79	78	78	79	78	79	76	79	82	80	83	98	98	100	100	99
ScV-LAlus-EX1160	46	79	75	80	79	78	78	79	78	79	76	79	82	80	83	98	98	100	100	99
ScV-LA2-EX1125	46	78	76	79	78	77	77	78	78	78	75	78	81	81	84	99	99	99	99	100

Colours: S. cerevisiae cluster Close to S. cerevisiae cluster
S. paradoxus cluster Close to S. paradoxus cluster
Equally distant from both clusters (74-85% identity)
Equally distant from both clusters (72-79% identity)
Equally distant from both clusters (43-47% identity)

Figure S5. Percentage identity matrix for Gag amino-acid sequences of LA viruses. Each value was rounded to the nearest whole number.

TdV-LAbarr1	100	20	34	32	32	32	32	32	30	25	30	27	23	25	25	25	23	27	25	27
ScV-LA1-original	20	100	57	64	64	64	64	64	68	61	64	64	64	66	66	64	68	64	57	55
SkV-LA1082	34	57	100	68	68	68	68	68	59	57	55	57	57	57	57	57	59	57	50	57
ScV-LA1-EX231	32	64	68	100	100	100	100	100	84	68	68	70	66	70	70	66	70	68	64	68
ScV-LA1us-EX1160	32	64	68	100	100	100	100	100	84	68	68	70	66	70	70	66	70	68	64	68
ScV-LA2-EX1125	32	64	68	100	100	100	100	100	84	68	68	70	66	70	70	66	70	68	64	68
ScV-LA1us4-EX229	32	64	68	100	100	100	100	100	84	68	68	70	66	70	70	66	70	68	64	68
ScV-LA1us1-EX436	32	64	68	100	100	100	100	100	84	68	68	70	66	70	70	66	70	68	64	68
ScV-LA2-8F13	30	68	59	84	84	84	84	84	100	68	73	70	68	73	73	66	70	68	68	66
SpV-LA74	25	61	57	68	68	68	68	68	68	100	91	91	91	95	95	80	86	82	68	77
SpV-LA1143	30	64	55	68	68	68	68	68	73	91	100	95	91	95	95	82	89	84	70	75
SpV-LA62	27	64	57	70	70	70	70	70	70	91	95	100	91	95	95	86	89	86	73	80
SuV-LA10560	23	64	57	66	66	66	66	66	68	91	91	91	100	95	95	80	86	82	68	77
SpV-LA4650	25	66	57	70	70	70	70	70	73	95	95	95	95	100	100	84	91	86	73	80
SpV-LA1939	25	66	57	70	70	70	70	70	73	95	95	95	95	100	100	84	91	86	73	80
SpV-LA66	25	64	57	66	66	66	66	66	66	80	82	86	80	84	84	100	93	91	70	77
SpV-LA21	23	68	59	70	70	70	70	70	70	86	88	89	86	91	91	93	100	91	68	75
SpV-LA28	27	64	57	68	68	68	68	68	68	82	84	86	82	86	86	91	91	100	73	80
SpV-LA45	25	57	50	64	64	64	64	64	64	68	70	73	68	73	73	70	68	73	100	80
SkV-LAFM1183	27	55	57	68	68	68	68	68	66	77	75	80	77	80	80	77	75	80	80	100

Colours: S. cerevisiae cluster Close to *S. cerevisiae* cluster
S. paradoxus cluster Close to *S. paradoxus* cluster
Equally distant from both clusters (64-68% identity)
Equally distant from both clusters (50-68% identity)
Equally distant from both clusters (23-32% identity)

Figure S6. Percentage identity matrix for the variable hydrophobic 44-amino-acid stretch of Pol sequences. Each value was rounded to the nearest whole number, except for the *S. cerevisiae*-cluster and SpV-LA4650-vs-SpV-LA1939 values that initially were 100%.

TdV-LAbarr1	100	17	17	11	22	17	17	17	17	28	22	28	33	22	22	22	22	22	11
SkV-LA1082	17	100	68	63	63	47	53	53	58	58	53	53	53	58	53	53	53	53	58
ScV-LA1-original	17	68	100	58	47	42	53	53	47	47	42	42	42	53	42	42	42	42	47
SpV-LA45	11	63	58	100	47	42	47	47	42	47	37	37	37	47	63	63	63	63	58
SpV-LA62	22	63	47	47	100	79	79	79	89	84	79	74	68	63	68	68	68	68	68
SpV-LA4650	17	47	42	42	79	100	84	84	89	84	63	74	68	58	58	58	58	58	58
SuV-LA10560	17	53	53	47	79	84	100	100	89	95	68	79	74	58	63	63	63	63	63
SpV-LA1939	17	53	53	47	79	84	100	100	89	95	68	79	74	58	63	63	63	63	63
SpV-LA1143	17	58	47	42	89	89	89	89	100	95	74	84	79	58	58	58	58	58	63
SpV-LA74	17	58	47	47	84	84	95	95	95	100	74	84	79	58	63	63	63	63	68
SpV-LA66	28	53	42	37	79	63	68	68	74	74	100	74	84	58	58	58	58	58	58
SpV-LA28	22	53	42	37	74	74	79	79	84	84	74	100	89	58	47	47	47	47	53
SpV-LA21	28	53	42	37	68	68	74	74	79	79	84	89	100	58	47	47	47	47	53
ScV-LA2-8F13	33	58	53	47	63	58	58	58	58	58	58	58	58	100	74	74	74	74	63
ScV-LA1-EX231	22	53	42	63	68	58	63	64	58	63	58	47	47	74	100	100	100	100	74
ScV-LA1us-EX1160	22	53	42	63	68	58	63	63	58	63	58	47	47	74	100	100	100	100	74
ScV-LA2-EX1125	22	53	42	63	68	58	63	63	58	63	58	47	47	74	100	100	100	100	74
ScV-LA1us4-EX229	22	53	42	63	68	58	63	63	58	63	58	47	47	74	100	100	100	100	74
ScV-LA1us1-EX436	22	53	42	63	68	58	63	63	58	63	58	47	47	74	100	100	100	100	74
SkV-LAFM1183	11	58	47	58	68	58	63	63	63	68	58	53	53	63	74	74	74	74	100

Colours: S. cerevisiae cluster Close to *S. cerevisiae* cluster
S. paradoxus cluster Close to *S. paradoxus* cluster
Equally distant from both clusters (53-58% identity)
Equally distant from both clusters (42-53% identity)
Equally distant from both clusters (17-22% identity)

Figure S7. Percentage identity matrix for the variable 19-amino-acid stretch located downstream of H153/154 of Gag sequences. Each value was rounded to the nearest whole number, except for the *S. cerevisiae*-cluster that initially was 100%.