

Dataset S1

CA-NTD amino acid sequences

CERV1 (consensus)

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPGGDGQPTANPDLINATFPLTRPAWDYNT
AEGRRLHLYRQTLMAGLRAAARK

CERV1#23

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPGGDGQPTANPDLINATFPLTRPAWDYNT
AEGRQLHLYRQTLMAGLRAAARK

CERV1#34

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPGGDSQPTANPDLINATFPLTRPAWDYNTA
EGRRLHLYRQTLMAGLRAAARK

CERV1#46

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPEDDGQPTANPDLINATFPLTRPAWDYNTA
EGRRLHLYRQTLMAGLRAAARK

CERV1#137

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPGGDGQPTANPDLINATFPLTRPAWDYNTM
AEGRRLHLYRQTLMAGLRAAARK

CERV1#319

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKSQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVGGDGQPTANPDLINATFPLTRPAWDYNT
AEGRRLHLYRQTLMAGLRAAARK

RhERV1a (consensus)

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINATFPLTRPAWDYNT
AEGRGRLLHLYRQTLMAGLRAAARK

RhERV1a#168

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLSLLDSVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINATFPLTRPAWDYNT
AEGRGRLLRLYRQTLMAGLRAAARK

RhERV1a#181

PLREIGPPDDTGIPRLQYWPFSTSDLYNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVLGDDGQPTANPDLINATFPLTRPAWDYNT
AEGRGRLLCLYRQTLMAGLRAAARK

RhERV1b (consensus)

PLREIGSLDDTGLSRLMYWPFSTSDLYNWKSQNARFSDNPKDLTSLLD SVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINAAFPLTRPRWDYNT
AEGRGRLLIYRQTLMAGLRAAARK

RhERV1b#67

PLREIGSLDDTGLSRLMYWPFSTSDLYNWKSQNARFSDNPKDLTLLLD SVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINAAFPLTRPRWDYNT
AEGRGRLLIYRQTLMAGLRAAARK

RhERV1b#75

PLREIGSLDDTGLSRLMYWPFSTSDLYNWKSQNARFSDNPKDLTSLLD SVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINVAFPLTRPRWDYNT
AEGRGRLLIYRQTLMAGLRAAARK

RhERV1b#77

PLREIGSLDDTGLSRLMYWPFSTSDLYNWKSQNARFSDNPKDLTSLLD SVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINAAFPLTRPRWDYN
MAEGRGRLLIYRQTLMAGLRAAARK

RhERV1b#92

PLREIGSFDDTGLSRLMYWPFSTSDLYN*KSQNARFSDNPKDLTSLLD SVMFTHQPT
WDDCQQLLRILFTTEERERIQVEARKLVPGDDGQPTANPDLINAAFPLTRPRWDYNT
AEGRGRLLIYRQTLMAGLRAAARK

CERV2 (consensus)

PLRTVNRTVQYWPFASDLYNWKTHNPPFSQDPQALTALIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLLEARKHVPGPGGLPTQLPNEIDEGFPLTRPDWDYETASGRESLR
IYRQALLAGLKGAGKR

CERV2#32

PLRTVNRTVQYWPFASDLYNWKTHNPPFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLLEARKNVPGPGGLPTQLPNEIDEGFPLTRLDWDYETATGRESLR
IYRQALLAGLKGAGKR

RhERV2 (consensus)

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLLEARKNVPGPGGLPTQLPNEIDEGFPLTRPDWDYETAPGRESLR
IYRQALLAGLKGAGKR

RhERV2#53

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLLEARKNVPGPGGLPTQLPNEIDEGFPLTRPDWDYETAPGRESLR
IYCQALLAGLKGAGKR

RhERV2#55

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLLEARKNVPGPGGLPTQLPNEIDEGFPLTRPDWDYEMAPGRESL
RIYRQALLAGLKGAGKR

RhERV2#58

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQQVLLLEARKNVPGPGGLPTQLPNKIDEGFPLTRPDWDYETAPGKESL
QIYRQALLAGLKGAGKR

RhERV2#63

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLEARKNVPGPGGLPTQLPNEIDEGFPLTRPDWDYETAPGRESLR
IYRQALLAGLKGAGKC

RhERV2#76

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLEARKNVPGPGGLPTQLPNEIDEGFPLTRPDWDYETAPGRESLR
IYHQALLAGLKGAGKR

RhERV2#91

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLEARKNVPGPGGLPTQLPNEIDEGFPLTRPDWDYETAPGRESLR
IYRQALLAGLKGAGKH

RhERV2#98

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLEARKNVPGPGGLPTQLPNEIDEGFPLTHPDWDYETAPGRESL
RIYRQALLAGLKGAGKR

RhERV2 E91K

PLRTVNRTVQYWPFASDLYNWKTHNPSFSQDPQALTSLIESILLTHQPTWDDCQQ
LLQVLLTTEERQRVLLEARKNVPGPGGLPTQLPNKIDEGFPLTRPDWDYETAPGRESL
RIYRQALLAGLKGAGKR

enMLV(consensus)

PLRMGGDGQLQYWPFSSDLYNWKNNNPSFSEDPGKLTALIESVLITHQPTWDDC
QQLLGTLLTGEEKQRVLLEARKAVRGNDGRPTQLPNEVNAAFPLERPDWDYTTTEGR
NHLVLYRQLLLAGLQNAGRS

enMLV#109

PLRMGGDGQLQYWPFSSDLYNWKNNNPSFSEDPGKLTALIESVLITHQPTWDDC
QQLLGTLLTGEEKQRVLLEARKAVRGNDGRPTQLPNEINAAFPLERPDWDYTTTEGR
NHLVLYRQLLLAGLQNAGRS

enMLV#111

PLRMGGDGQLQYWPFSDDLWNWKNNNPSFSDPGKLTALIESVLITHQPTWDDC
QQLLGTLLTGEEKQRVLLLEARKAVRGNDGRPTQLPNEVNAAFPLERPGWDYTTTEGR
NHLVLYRQLLLGLQAGLNAGRS

enMLV#116

PLRMGGDGQLQYWPFSDDLWNWKNNNPSFSDPGKLTALIESVLITHQPTWDDC
QQLLGTLLTGEEKQRVLLLEARKAVRGNDGRPTQLPNEVNAAFPLERPDWDYTTTEGR
NHLVLYRQLLLGLQAGLNAGRS

enMLV#124

PLRMGRDGQLQYWPFSDDLWNWKNNNPSFSDPGKLTALIESVLITHQPTWDDCQ
QQLLGTLLTGEEKQRVLLLEARKAVRGNDGRPTQLPNEVNAAFPLERPGWDYTTTEGRN
HLVLYRQLLLGLQAGLNAGRS

enMLV#127

PLRMGGDGQLQYWPFSDDLWNWKNNNPSFSDPGKLTALIESVLITHQPTWDDC
QQLLGTLLTGEEKQRVLLLEARKAVRGDDGRPTQLPNEVNAAFPLERPGWDYTTTEGR
NHLVLYRQLLLGLQAGLNAGRS

Complete CA sequences

CERV1(consensus)

PLREIGPPDDTGIPRLQYWPFSDDLWNWKTQSARFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPGDDGQPTANPDLINATFPLTRPAWDYNT
AEGRGRHLHYRQTLMAGLRAAARKPTNLAKVYSILQGKTESPATYLERLMEAFRQYTP
IDPEAPGSQAAVVMSFVNQAAPDIKRKLQKLEDLEGKRIQDLLQIAQRVYNNRDTPE
EKQFKATEKMTKVL

CERV1(ancestral)

PLREIGPPDDTGIPRLQYWPFSDDLWNWKTQSAQFSDNPKDLLALLDSVMFTHQPT
WDDCQQLLRILFTTEERERIEARKLVPGDDGQPTANPDLINATFPLTRPAWDYNT
AEGRGRHLHYRQTLMAGLRAAARKPTNLAKVYSILQGKTESPATYLERLMEAFRQYTP
IDPEAPGSQAAVVMSFVNQAAPDIKRKLQKLEDLEGKRIQDLLQIAQRVYNNRDTPE
EKQFKATEKMTKVL