Supplemental Information

Table S1. Estimated insertion time of endogenous lokiretroviruses

Species	Order	Genome	5'LTR-3'LTR Distance	Insertion Time/Mya
Betta splendens	Anabantiformes	GCF_900634795.2	0.001	0.23
Seriola rivoliana	Carangiformes	GCA_002994505.1	0.006	1.37
Astyanax mexicanus	Characiformes	GCF_000372685.2	0.011	2.75
Denticeps clupeoides	Clupeiformes	GCF_900700375.1	0	0
Labeo rohita	Cyprinodontiformes	GCA_000776015.1	0.013	2.95
Symphodus melops	Labriformes	GCA_002819105.1	0	0
Perca flavescens	Perciformes	GCF_004354835.1	0.009	2.05
Salmo trutta	Salmoniformes	GCF_901001165.1	0.003	0.68
Nanorana parkeri	Anura	GCF_000935625.1	0.162	36.82
Rana catesbeiana	Anura	GCA_002284835.2	0.018	4.09
Xenopus tropicalis	Anura	GCF_000004195.3	0.003	0.68

Table S2. Information of genomes used in this study

Organism Groups	Organism Name	Order	Assembly	Level	Size(Mb)	GC%	Scaffolds
Other animals (36)	*Asymmetron lucayanum*	Amphioxiformes	GCA_001663935.1	Scaffold	460.59	40.4	212366
	Branchiostoma belcheri	Amphioxiformes	GCF_001625305.1	Scaffold	426.12	41.4	2308
	Branchiostoma belcheri	Amphioxiformes	GCA_001625405.1	Scaffold	707.9	41.3	10354
	Branchiostoma belcheri	Amphioxiformes	GCA_900128855.1	Scaffold	417.04	39	2307
	Branchiostoma floridae	Amphioxiformes	GCF_000003815.1	Scaffold	521.89	41.8	398
	Branchiostoma lanceolatum	Amphioxiformes	GCA_900088365.1	Scaffold	495.29	39.8	10247
	Eucidaris tribuloides	Cidaroida	GCA_001188425.1	Scaffold	2187.26	37.2	637071
	Bathochordaeus stygius	Copelata	GCA_004367955.1	Scaffold	396.47	36.6	467514
	Fritillaria borealis	Copelata	GCA_004368075.1	Scaffold	143.09	40.5	142328
	Mesochordaeus erythrocephalus	Copelata	GCA_004367975.1	Scaffold	874.02	36.4	745784
	Oikopleura albicans	Copelata	GCA_004367875.1	Scaffold	365.96	33.9	92906
	Oikopleura dioica	Copelata	GCA_000209535.1	Scaffold	70.47	40.2	1260
	Oikopleura dioica	Copelata	GCA_000209555.1	Scaffold	45.14	40.3	4196
	Oikopleura longicauda	Copelata	GCA_004367895.1	Contig	308.74	37.3	178440
	Oikopleura vanhoeffeni	Copelata	GCA_004367855.1	Scaffold	643.66	30.2	79827
	Hemicentrotus pulcherrimus	Echinoida	GCA_003118195.1	Scaffold	568.91	30.7	16251
	Strongylocentrotus purpuratus	Echinoida	GCF_000002235.4	Scaffold	921.86	37.4	871
	Phallusia mammillata	Enterogona	GCA_003260075.1	Scaffold	233.92	36.1	11003
	Ciona intestinalis	Enterogona	GCF_000224145.3	Scaffold	115.23	36.02	1280
	Ciona intestinalis	Enterogona	GCA_000183065.1	Scaffold	116.73	36.1	2501
	Ciona savignyi	Enterogona	GCA_000149265.1	Scaffold	587.35	37.5	34009
	Ptychodera flava	Enteropneusta	GCA_001465055.1	Scaffold	1228.69	38.4	218255
	Saccoglossus kowalevskii	Enteropneusta	GCF_000003605.2	Scaffold	775.84	38.1	54120
	Ophionereis fasciata	Ophiurida	GCA_900067615.1	Scaffold	1184.53	32.6	3968282
	Ophiothrix spiculata	Ophiurida	GCA_000969725.1	Scaffold	2764.32	35.1	75696
	Botryllus schlosseri	Pleurogona	GCA_000444245.1	Scaffold	579.63	40.6	120139
	Salpa thompsoni	Salpida	GCA_001749815.1	Scaffold	318.75	34.2	478281
	Apostichopus japonicus	Synallactida	GCA_002754855.1	Scaffold	804.62	36.8	3278

	Apostichopus japonicus	Synallactida	GCA_001866495.1	Scaffold	664.37	35.9	132607
	Apostichopus parvimensis	Synallactida	GCA_000934455.1	Scaffold	873.09	38.7	21559
	Lytechinus variegatus	Temnopleuridea	GCA_000239495.2	Scaffold	1061.2	36.9	322794
	Acanthaster planci	Valvatida	GCF_001949145.1	Scaffold	383.86	41.5	1766
	Acanthaster planci	Valvatida	GCA_000950615.1	Scaffold	2.58	41.5	2
	Acanthaster planci	Valvatida	GCA_001949165.1	Scaffold	383.52	41.5	3274
	Patiria miniata	Valvatida	GCA_000285935.1	Scaffold	811.03	40.5	60183
	Patiriella regularis	Valvatida	GCA_900067625.1	Scaffold	949.33	37	3006458
Fishes (4)	_Eptatretus burgeri_	Myxiniformes	GCA_900186335.2	Scaffold	2608.38	-	10846
	_Lethenteron camtschaticum_*	Petromyzontiformes	GCA_000466285.1	Scaffold	1030.66	48.1	86125
	_Petromyzon marinus_*	Petromyzontiformes	GCA_002833325.1	Scaffold	1130.42	44.4	12061
	_Petromyzon marinus_*	Petromyzontiformes	GCA_000148955.1	Scaffold	885.53	46.8	25005
Amphibians (7)	_Xenopus tropicalis_*	Anura	GCF_000004195.3	Chromosome	1440.4	40.5404	6822
	_Nanorana parkeri_*	Anura	GCF_000935625.1	Scaffold	2053.87	42.7	25188
	_Xenopus laevis_*	Anura	GCF_001663975.1	Chromosome	2718.43	40.2325	108033
	Rhinella marina	Anura	GCA_900303285.1	Contig	2551.76	-	31391
	Rana catesbeiana *	Anura	GCA_002284835.2	Scaffold	6250.35	43.6998	1544635
	Pyxicephalus adspersus	Anura	GCA_004786255.1	Chromosome	1563.37	39.2274	5411
	_Rhinatrema bivittatum_*	Gymnophiona	GCF_901001135.1	Chromosome	5319.24	44.44	1330
reptiles (43)	_Alligator mississippiensis_	Crocodylia	GCF_000281125.3	Scaffold	2161.73	44.4	7094
	Alligator sinensis	Crocodylia	GCF_000455745.1	Scaffold	2270.57	44.6	9317
	Crocodylus porosus	Crocodylia	GCF_001723895.1	Scaffold	2049.54	44.2	70
	Sphenodon punctatus	Sphenodontia	GCA_003113815.1	Scaffold	4272.21	-	16536
	Anolis carolinensis	Squamata	GCF_000090745.1	Chromosome	1799.14	40.8238	6646
	Crotalus horridus	Squamata	GCA_001625485.1	Scaffold	1520.33	34.3	186068
	Crotalus pyrrhus	Squamata	GCA_000737285.1	Scaffold	1126.79	38.5	473380
	Crotalus viridis viridis	Squamata	GCA_003400415.2	Chromosome	1340.2	39.3623	7043
	Emydocephalus ijimae	Squamata	GCA_004319985.1	Scaffold	1625.2	40.3	157858
	Gekko japonicus	Squamata	GCF_001447785.1	Scaffold	2490.27	45.5	191500
	Hydrophis cyanocinctus	Squamata	GCA_004023725.1	Scaffold	1389.86	37.6	546690
	Hydrophis hardwickii	Squamata	GCA_004023765.1	Scaffold	1296.39	37.2	616046
	Hydrophis melanocephalus	Squamata	GCA_004320005.1	Scaffold	1402.64	34.8	122022
	Lacerta bilineata	Squamata	GCA_900245895.1	Contig	1418.15	-	7510

	Lacerta viridis	Squamata	GCA_900245905.1	Contig	1439.84	43.7	4737
	Laticauda colubrina	Squamata	GCA_004320045.1	Scaffold	2024.69	35.6	62906
	Laticauda laticaudata	Squamata	GCA_004320025.1	Scaffold	1558.71	40.1	83587
	Notechis scutatus	Squamata	GCF_900518725.1	Scaffold	1665.53	40.2	52414
	Ophiophagus hannah	Squamata	GCA_000516915.1	Scaffold	1594.07	40.6	296399
	Pantherophis guttatus	Squamata	GCA_001185365.1	Scaffold	1404.22	38.3	883920
	Paroedura picta	Squamata	GCA_003118565.1	Scaffold	1694.16	40.7	110906
	Podarcis muralis	Squamata	GCF_004329235.1	Chromosome	1511.02	44.2081	2161
	Pogona vitticeps	Squamata	GCF_900067755.1	Scaffold	1716.68	42.1	13749
	Protobothrops flavoviridis	Squamata	GCA_003402635.1	Scaffold	1413.2	38.2	84502
	Protobothrops mucrosquamatus	Squamata	GCF_001527695.2	Scaffold	1673.88	40.6	52280
	Pseudonaja textilis	Squamata	GCF_900518735.1	Scaffold	1590.04	40.1	28550
	Python bivittatus	Squamata	GCF_000186305.1	Scaffold	1435.05	39.7	39113
	Salvator merianae	Squamata	GCA_003586115.2	Scaffold	2068.17	41.5	4512
	Thamnophis sirtalis	Squamata	GCF_001077635.1	Scaffold	1424.9	41.8	7930
	Thermophis baileyi	Squamata	GCA_003457575.1	Scaffold	1747.68	-	20729
	Varanus komodoensis	Squamata	GCA_004798865.1	Scaffold	1507.95	43.6	1411
	Vipera berus berus	Squamata	GCA_000800605.1	Scaffold	1532.39	41.3	28883
	Apalone spinifera	Testudines	GCA_000385615.1	Scaffold	1931.08	42.8	286620
	*Chelonia mydas**	Testudines	GCF_000344595.1	Scaffold	2208.41	43.7	140023
	*Chelonoidis abingdonii**	Testudines	GCA_003597395.1	Scaffold	2300.75	45.4	10623
	*Chrysemys picta bellii**	Testudines	GCF_000241765.3	Chromosome	2365.77	44.564	78631
	*Cuora amboinensis**	Testudines	GCA_004028625.2	Scaffold	2214.83	43.9	89628
	*Cuora mccordi**	Testudines	GCA_003846335.1	Scaffold	2390.37	44.6	100020
	*Gopherus agassizii**	Testudines	GCA_002896415.1	Scaffold	2184.97	43	172559
	*Malaclemys terrapin terrapin**	Testudines	GCA_001728815.2	Contig	2439.75	44.5	21684
	Pelodiscus sinensis	Testudines	GCF_000230535.1	Scaffold	2202.48	44.4999	19904
	*Platysternon megacephalum**	Testudines	GCA_003942145.1	Scaffold	2319.09	43.7	359693
	*Terrapene carolina triunguis**	Testudines	GCF_002925995.2	Scaffold	2571.27	44.6	52260
Birds (146)	*Accipiter nisus*	Accipitriformes	GCA_004320145.1	Scaffold	1190.65	41.8	44975
	Anser brachyrhynchus	Anseriforme	GCA_002592135.1	Scaffold	1116.99	41.1	2723
	Anser cygnoides domesticus	Anseriforme	GCF_000971095.1	Scaffold	1119.15	41.5001	7593
	Anser indicus	Anseriforme	GCA_006229135.1	Scaffold	1143.09	40.8	10527
	Anas platyrhynchos	Anseriformes	GCF_003850225.1	Chromosome	1126.18	41.6676	2150
	Anas zonorhyncha	Anseriformes	GCA_002224875.1	Scaffold	1310.8	37.7	20089

Calypte anna	Apodiformes	GCF_000699085.1	Scaffold	1105.68	41.3	54736
Chaetura pelagica	Apodiformes	GCF_000747805.1	Scaffold	1119.19	41.6	19072
Apteryx australis mantelli	Apterygiformes	GCF_001039765.1	Scaffold	1523.97	42.5	24720
Apteryx haastii	Apterygiformes	GCA_003342985.1	Scaffold	1221.44	40.8	4183
Apteryx owenii	Apterygiformes	GCA_003342965.1	Scaffold	1231.28	40.9	5096
Apteryx rowi	Apterygiformes	GCF_003343035.1	Scaffold	1228.9	41.6	4120
Buceros rhinoceros silvestris	Bucerotiformes	GCF_000710305.1	Scaffold	1065.78	42.6	62257
Antrostomus carolinensis	Caprimulgiformes	GCF_000700745.1	Scaffold	1119.68	40.8	70122
Cariama cristata	Cariamiformes	GCF_000690535.1	Scaffold	1132.25	41.2	53474
Casuarius casuarius	Casuariiformes	GCA_003342895.1	Scaffold	1209	41.4	3184
Dromaius novaehollandiae	Casuariiformes	GCF_003342905.1	Scaffold	1192.25	41.7	2778
Calidris pugnax	Charadriiformes	GCF_001431845.1	Scaffold	1229.09	42.7	3753
Calidris pygmaea	Charadriiformes	GCA_003697955.1	Scaffold	1178.37	44.2	29819
Charadrius vociferus	Charadriiformes	GCF_000708025.1	Scaffold	1219.86	42	15167
Himantopus himantopus leucocephalus	Charadriiformes	GCA_003993805.1	Scaffold	1116.81	41.6	1442
Limosa lapponica baueri	Charadriiformes	GCA_002844005.1	Scaffold	1034.77	42.3	32319
Recurvirostra avosetta	Charadriiformes	GCA_004023745.1	Scaffold	1193.28	41.9	35873
Scolopax mira	Charadriiformes	GCA_004320125.1	Scaffold	1091.29	41.6	40750
Uria lomvia	Charadriiformes	GCA_002289315.1	Scaffold	1179.36	41.5	9327
Cathartes aura	Ciconiiformes	GCA_000699945.1	Scaffold	1152.57	41.1	104141
Ciconia boyciana	Ciconiiformes	GCA_002002965.1	Scaffold	1364.95	42.1	505419
Pterocles gutturalis	Ciconiiformes	GCF_000699245.1	Scaffold	1069.32	41.4	58607
Colius striatus	Coliiformes	GCF_000690715.1	Scaffold	1075.93	40.9	70188
Columba livia	Columbiformes	GCF_000337935.1	Scaffold	1107.99	41.6001	14923
Patagioenas fasciata monilis	Columbiformes	GCA_002029285.1	Scaffold	1089.15	40.9	9784
Streptopelia turtur	Columbiformes	GCA_901699155.1	Chromosome	1178.93	41.84	357
Leptosomus discolor	Coraciiformes	GCF_000691785.1	Scaffold	1136.24	41.8	57160
Merops nubicus	Coraciiformes	GCF_000691845.1	Scaffold	1062.96	41.7	53499
Cuculus canorus	Cuculiformes	GCF_000709325.1	Scaffold	1153.89	41.7	14930
Aquila chrysaetos canadensis	Falconiformes	GCF_000766835.1	Scaffold	1192.74	41.9001	1142
Falco cherrug	Falconiformes	GCF_000337975.1	Scaffold	1174.81	41.8	5863
Falco peregrinus	Falconiformes	GCF_000337955.1	Scaffold	1171.97	41.8	7021
Haliaeetus albicilla	Falconiformes	GCF_000691405.1	Scaffold	1133.55	40.9	50905
Haliaeetus leucocephalus	Falconiformes	GCF_000737465.1	Scaffold	1178.41	41.8	1023
Bambusicola thoracicus	Galliformes	GCA_002909625.1	Scaffold	1032.3	0	163748

Callipepla squamata	Galliformes	GCA_002218305.1	Scaffold	1045.28	40.6	34302
Chrysolophus pictus	Galliformes	GCA_003413605.1	Scaffold	1023.99	41	14294
Colinus virginianus	Galliformes	GCA_000599465.2	Scaffold	1254.15	38.6	42369
Coturnix japonica	Galliformes	GCF_001577835.1	Chromosome	927.657	41.3684	2531
Gallus gallus	Galliformes	GCF_000002315.6	Chromosome	1065.37	42.3032	525
Lagopus muta japonica	Galliformes	GCA_004320205.1	Scaffold	1002.56	41.4	62806
Lyrurus tetrix tetrix	Galliformes	GCA_000586395.1	Contig	657.025	40.9	971998
Meleagris gallopavo	Galliformes	GCF_000146605.2	Chromosome	1128.34	41.7219	233806
Numida meleagris	Galliformes	GCF_002078875.1	Chromosome	1043.26	41.8975	2739
Pavo cristatus	Galliformes	GCA_005519975.1	Scaffold	1027.51	-	179332
Phasianus colchicus	Galliformes	GCA_004143745.1	Scaffold	1017.51	40.9	39677
Syrmaticus mikado	Galliformes	GCA_003435085.1	Scaffold	1066.43	40.5	71281
Tympanuchus cupido pinnatus	Galliformes	GCA_001870855.1	Scaffold	983.777	-	12186
Gavia stellata	Gaviiformes	GCF_000690875.1	Scaffold	1129.69	41.1001	61831
Balearica regulorum gibbericeps	Gruiformes	GCF_000709895.1	Scaffold	1127.62	41.2001	53491
Chlamydotis macqueenii	Gruiformes	GCF_000695195.1	Scaffold	1086.57	41.1	59693
Chlamydotis undulata undulata	Gruiformes	GCA_003400225.1	Scaffold	1307.64	39	80662
Eurypyga helias	Gruiformes	GCF_000690775.1	Scaffold	1088.02	42.3	62699
Gallirallus okinawae	Gruiformes	GCA_002003005.1	Contig	1114.53	42.7	768680
Grus japonensis	Gruiformes	GCA_002002985.1	Scaffold	1265.63	41.6	357545
Grus nigricollis	Gruiformes	GCA_004360235.1	Scaffold	1230.91	42.7	1812
Mesitornis unicolor	Gruiformes	GCF_000695765.1	Scaffold	1087.29	41.3	67520
Tauraco erythrolophus	Musophagiformes	GCF_000709365.1	Scaffold	1155.54	41.6	59587
Opisthocomus hoazin	Opisthocomiformes	GCF_000692075.1	Scaffold	1203.71	42.7	10256
Acanthisitta chloris	Passeriformes	GCF_000695815.1	Scaffold	1035.88	41.6	53875
Acridotheres javanicus	Passeriformes	GCA_002849675.1	Scaffold	1051.58	39.9	4312
Cicinnurus regius	Passeriformes	GCA_003713305.1	Scaffold	1059.54	41.8	4390
Corapipo altera	Passeriformes	GCF_003945725.1	Scaffold	1095.75	41.9	9068
Corvus brachyrhynchos	Passeriformes	GCF_000691975.1	Scaffold	1091.31	42	10547
Corvus cornix cornix	Passeriformes	GCF_000738735.2	Scaffold	1042.28	41.7	113
Corvus hawaiiensis	Passeriformes	GCA_003402825.1	Contig	1497.61	42.9	2752
Cyanistes caeruleus	Passeriformes	GCF_002901205.1	Scaffold	1186.98	41.8	29297
Diphyllodes magnificus	Passeriformes	GCA_003713285.1	Scaffold	1065.98	41.5	3098
Empidonax traillii	Passeriformes	GCF_003031625.1	Scaffold	1123.49	42.4	7791
Eopsaltria australis	Passeriformes	GCA_003426825.1	Scaffold	1228.34	43	20702

Erythrura gouldiae	Passeriformes	GCA_003676055.1	Scaffold	1070.79	41	18372
Ficedula albicollis	Passeriformes	GCF_000247815.1	Chromosome	1118.34	44.2999	21836
Geospiza fortis	Passeriformes	GCF_000277835.1	Scaffold	1065.29	41.7	27239
Hemignathus virens	Passeriformes	GCA_003286495.1	Scaffold	1086.51	40.1	19546
Hirundo rustica rustica	Passeriformes	GCA_003692655.1	Scaffold	1213.74	41.6	364
Junco hyemalis	Passeriformes	GCA_003829775.1	Scaffold	958.315	40.7	4457
Lepidothrix coronata	Passeriformes	GCF_001604755.1	Scaffold	1079.58	41.9	4612
Lonchura striata domestica	Passeriformes	GCF_002197715.1	Scaffold	1060.17	41.8001	3014
Manacus vitellinus	Passeriformes	GCF_001715985.2	Scaffold	1213.15	42.3	15315
Mixornis gularis	Passeriformes	GCA_003546035.1	Scaffold	1036.1	-	9542
Neopelma chrysocephalum	Passeriformes	GCF_003984885.1	Scaffold	1142.8	42.4	42063
Paradisaea raggiana	Passeriformes	GCA_003713265.1	Scaffold	1037.51	41.9	6544
Paradisaea rubra	Passeriformes	GCA_003713215.1	Scaffold	1071.26	41.6	4376
Parotia lawesii	Passeriformes	GCA_003713295.1	Scaffold	1056.92	42	3976
Parus major	Passeriformes	GCF_001522545.3	Chromosome	1020.31	41.5208	1675
Passer domesticus	Passeriformes	GCA_001700915.1	Chromosome	1042.72	41.1833	2571
Phylloscopus plumbeitarsus	Passeriformes	GCA_001655115.1	Scaffold	997.399	41.5	308480
Phylloscopus trochiloides viridanus	Passeriformes	GCA_001655095.1	Scaffold	1003.33	41.4	158605
Phylloscopus trochilus acredula	Passeriformes	GCA_002305835.1	Scaffold	1066.52	41.2	5995
Pipra filicauda	Passeriformes	GCF_003945595.1	Scaffold	1110.96	42.3	3744
Pseudopodoces humilis	Passeriformes	GCF_000331425.1	Scaffold	1043	41.8001	5406
Saxicola maurus maurus	Passeriformes	GCA_900205225.1	Scaffold	1020.37	39.6	2819
Serinus canaria	Passeriformes	GCF_000534875.1	Scaffold	1152.1	42.6	304400
Setophaga coronata coronata	Passeriformes	GCA_001746935.1	Scaffold	1171.93	37.3	154686
Sporophila hypoxantha	Passeriformes	GCA_002167245.1	Scaffold	1067.76	40.6	5095
Sturnus vulgaris	Passeriformes	GCF_001447265.1	Scaffold	1036.76	41.7	2361
Taeniopygia guttata	Passeriformes	GCF_000151805.1	Chromosome	1232.14	41.4526	37422
Zosterops lateralis melanops	Passeriformes	GCA_001281735.1	Scaffold	1036	41.3	2933
Zonotrichia albicollis	Passeriformes	GCF_000385455.1	Scaffold	1052.6	41.8	6018
Egretta garzetta	Pelecaniformes	GCF_000687185.1	Scaffold	1206.5	42.5	11791
Nannopterum auritus	Pelecaniformes	GCA_002173455.1	Scaffold	1246.05	42.1	27523
Nannopterum brasilianus	Pelecaniformes	GCA_002174335.1	Scaffold	1346.19	42.2	266364
Nannopterum harrisi	Pelecaniformes	GCA_002173475.1	Scaffold	1202.99	42.1	2651
Nipponia nippon	Pelecaniformes	GCF_000708225.1	Scaffold	1223.86	42.0001	59555
Pelecanus crispus	Pelecaniformes	GCF_000687375.1	Scaffold	1160.92	41.4	63982

Phaethon lepturus	Pelecaniformes	GCF_000687285.1	Scaffold	1152.96	41.5	66785
Phalacrocorax carbo	Pelecaniformes	GCF_000708925.1	Scaffold	1138.97	41.3	64312
Urile pelagicus	Pelecaniformes	GCA_002173435.1	Scaffold	1210.66	-	20951
Phoenicopterus ruber ruber	Phoenicopteriformes	GCA_000687265.1	Scaffold	1132.18	41.9	76189
Dendrocopos noguchii	Piciformes	GCA_004320165.1	Scaffold	1134.56	44.8	154720
Picoides pubescens	Piciformes	GCF_000699005.1	Scaffold	1167.32	44.6	31254
Podiceps cristatus	Podicipediformes	GCA_000699545.1	Scaffold	1134.92	41.5	82923
Fulmarus glacialis	Procellariiformes	GCF_000690835.1	Scaffold	1141.4	41.2	57389
Agapornis roseicollis	Psittaciformes	GCA_002631895.1	Scaffold	1121.79	41	52900
Amazona aestiva	Psittaciformes	GCA_001420675.1	Scaffold	1129.54	42.2	3232
Amazona collaria	Psittaciformes	GCA_003947215.1	Scaffold	1258.72	-	10638
Amazona vittata	Psittaciformes	GCA_000332375.1	Scaffold	1175.4	41.7	182974
Ara macao	Psittaciformes	GCA_000400695.1	Scaffold	1204.7	42.5001	192790
Melopsittacus undulatus	Psittaciformes	GCF_000238935.1	Scaffold	1117.37	41.4001	25212
Nestor notabilis	Psittaciformes	GCF_000696875.1	Scaffold	1053.56	41.1	42180
Psittacula krameri	Psittaciformes	GCA_002870145.1	Scaffold	1417.81	38.1	167560
Strigops habroptila	Psittaciformes	GCA_004027225.1	Chromosome	1165.64	42.1216	100
Pterocnemia pennata	Rheiformes	GCA_003342835.1	Scaffold	1168.29	40.3	3019
Rhea americana	Rheiformes	GCA_003343005.1	Scaffold	1160.84	40.2	2162
Aptenodytes forsteri	Sphenisciformes	GCF_000699145.1	Scaffold	1254.35	42	10672
Pygoscelis adeliae	Sphenisciformes	GCF_000699105.1	Scaffold	1216.62	41.8001	19265
Pygoscelis antarcticus	Sphenisciformes	GCA_003264595.1	Scaffold	1209.83	39.6	1783
Pygoscelis papua	Sphenisciformes	GCA_003264615.1	Scaffold	1209.91	39.6	1783
Spheniscus humboldti	Sphenisciformes	GCA_003264545.1	Scaffold	1209.86	39.6	1783
Spheniscus magellanicus	Sphenisciformes	GCA_003264715.1	Scaffold	1209.86	39.6	1783
Spheniscus mendiculus	Sphenisciformes	GCA_003264655.1	Scaffold	1209.85	39.6	1783
Athene cunicularia	Strigiformes	GCF_003259725.1	Scaffold	1157.07	41	445
Bubo blakistoni	Strigiformes	GCA_004320225.1	Scaffold	1193.86	41.7	46772
Strix occidentalis caurina	Strigiformes	GCA_002372975.1	Scaffold	1255.54	41.3	8108
Tyto alba	Strigiformes	GCF_000687205.1	Scaffold	1120.14	40.2	62122
Struthio camelus australis	Struthioniformes	GCF_000698965.1	Scaffold	1225.04	41.3	6915
Crypturellus cinnamomeus	Tinamiformes	GCA_003342915.1	Scaffold	1122.05	41.7	8086
Eudromia elegans	Tinamiformes	GCA_003342815.1	Scaffold	955.515	41.8	2364
Nothoprocta perdicaria	Tinamiformes	GCF_003342845.1	Scaffold	965.904	42.1	2217
Tinamus guttatus	Tinamiformes	GCF_000705375.1	Scaffold	1047.06	41.5	82514

	Apaloderma vittatum	Trogoniformes	GCF_000703405.1	Scaffold	1070.84	41.4	54728
Fishes (226)	*Acipenser ruthenus**	Acipenseriformes	GCA_004119895.1	Scaffold	1732.55	34.4	215913
	Anabas testudineus	Anabantiformes	GCF_900324465.1	Scaffold	569.722	40.5001	70
	*Betta splendens**	Anabantiformes	GCF_900634795.2	Chromosome	441.389	45.2975	70
	Channa argus var. Kimnra	Anabantiformes	GCA_004786185.1	Chromosome	644.13	39.8556	562
	*Helostoma temminkii**	Anabantiformes	GCA_900302695.1	Scaffold	599.808	-	21407
	*Anguilla anguilla**	Anguilliformes	GCA_000695075.1	Scaffold	1018.7	42.9	501148
	*Anguilla japonica**	Anguilliformes	GCA_003597225.1	Scaffold	966.917	-	83292
	*Anguilla rostrata**	Anguilliformes	GCA_001606085.1	Scaffold	1413.03	41	79209
	*Guentherus altivela**	Ateleopodiformes	GCA_900312595.1	Scaffold	539.599	42.2	189411
	Parasudis fraserbrunneri	Aulopiformes	GCA_900302295.1	Scaffold	707.987	-	156389
	*Chatrabus melanurus**	Batrachoidiformes	GCA_900302635.1	Scaffold	1128.44	41.4	238795
	*Oryzias javanicus**	Beloniformes	GCA_003999625.1	Chromosome	809.68	38.9689	254
	Oryzias latipes	Beloniformes	GCF_002234675.1	Chromosome	734.057	40.8475	25
	*Oryzias melastigma**	Beloniformes	GCF_002922805.1	Scaffold	779.47	39.6	8603
	*Acanthochaenus luetkenii**	Beryciformes	GCA_900312575.1	Scaffold	545.759	42.6	91087
	*Beryx splendens**	Beryciformes	GCA_900312565.1	Scaffold	533.268	43.3	117400
	Monocentris japonicus	Beryciformes	GCA_900323365.1	Scaffold	556.024	-	52108
	*Rondeletia loricata**	Beryciformes	GCA_900302605.1	Scaffold	568.598	-	103827
	*Echeneis naucrates**	Carangiformes	GCF_900963305.1	Chromosome	544.213	41.43	37
	Selene dorsalis	Carangiformes	GCA_900303245.1	Scaffold	528.779	41.7	36113
	*Seriola dumerili**	Carangiformes	GCF_002260705.1	Scaffold	677.686	40.9002	34656
	*Seriola lalandi dorsalis**	Carangiformes	GCF_002814215.1	Scaffold	732.51	40.9002	99598
	*Seriola quinqueradiata**	Carangiformes	GCA_002217815.1	Scaffold	639.27	40.9	384
	*Seriola rivoliana**	Carangiformes	GCA_002994505.1	Scaffold	666.142	40.5002	1343
	*Trachinotus ovatus**	Carangiformes	GCA_900607315.1	Scaffold	648.062	41.1	138
	Scyliorhinus torazame	Carcharhiniformes	GCA_003427355.1	Scaffold	4470.98	38.5	458049
	*Maccullochella peelii**	Centrarchiformes	GCA_002120245.1	Scaffold	633.241	-	18198
	*Macquaria australasica**	Centrarchiformes	GCA_005408345.1	Scaffold	675.976	41	2962
	*Micropterus floridanus**	Centrarchiformes	GCA_002592385.1	Contig	1001.52	40.7	249768
	*Oplegnathus fasciatus**	Centrarchiformes	GCA_003416845.1	Scaffold	766.301	38.7	4149
	*Astyanax mexicanus**	Characiformes	GCF_000372685.2	Chromosome	1335.24	38.409	2415
	*Pygocentrus nattereri**	Characiformes	GCF_001682695.1	Scaffold	1285.35	40.6001	283518
	Callorhinchus milii	Chimaeriformes	GCF_000165045.1	Scaffold	974.499	42.5998	21204
	Amphilophus citrinellus	Cichliformes	GCA_000751415.1	Scaffold	844.903	41.4	6637

Astatotilapia calliptera	Cichliformes	GCF_900246225.1	Chromosome	880.446	41.0816	249
Haplochromis burtoni	Cichliformes	GCF_000239415.1	Scaffold	831.412	41.9	8001
Labeotropheus fuelleborni	Cichliformes	GCA_000150875.1	Scaffold	70.8584	42.1	58245
Maylandia zebra	Cichliformes	GCF_000238955.4	Chromosome	957.485	41.1537	1690
Mchenga conophoros	Cichliformes	GCA_000150855.1	Scaffold	73.4256	41.8	61923
Melanochromis auratus	Cichliformes	GCA_000150895.1	Scaffold	68.2386	41.5	63297
Neolamprologus brichardi	Cichliformes	GCF_000239395.1	Scaffold	847.91	42.0001	9099
Oreochromis aureus	Cichliformes	GCA_005870065.1	Scaffold	918.94	36.3	12951
Oreochromis niloticus	Cichliformes	GCF_001858045.2	Chromosome	1005.68	40.7187	2460
Pundamilia nyererei	Cichliformes	GCF_000239375.1	Scaffold	830.133	41.9	7236
Rhamphochromis esox	Cichliformes	GCA_000150935.1	Scaffold	71.2951	42.3	55751
Simochromis diagramma	Cichliformes	GCA_900408965.1	Scaffold	848.827	40.9	823
*Clupea harengus**	Clupeiformes	GCF_000966335.1	Scaffold	807.712	44.5001	6915
*Denticeps clupeoides**	Clupeiformes	GCF_900700375.1	Chromosome	567.401	43.744	460
Sardina pilchardus	Clupeiformes	GCA_900499035.1	Scaffold	949.617	-	117259
Tenualosa ilisha	Clupeiformes	GCA_003651195.1	Scaffold	815.648	43.3	124209
*Latimeria chalumnae**	Coelacanthiformes	GCF_000225785.1	Scaffold	2860.59	43	22819
*Carassius auratus**	Cypriniformes	GCF_003368295.1	Chromosome	1820.64	37.4838	6216
*Cirrhinus molitorella**	Cypriniformes	GCA_004028445.1	Scaffold	920.273	34.3	478313
*Cyprinus carpio**	Cypriniformes	GCF_000951615.1	Chromosome	1713.66	37.3325	9378
*Danionella dracula**	Cypriniformes	GCA_900490495.1	Scaffold	665.208	37.5	996
*Hypophthalmichthys molitrix**	Cypriniformes	GCA_004764525.1	Scaffold	1104.68	31.4	107095
*Hypophthalmichthys nobilis**	Cypriniformes	GCA_004193235.1	Scaffold	1012.06	34.7	121326
*Labeo rohita**	Cypriniformes	GCA_004120215.1	Scaffold	1484.73	35.0001	13623
*Leuciscus waleckii**	Cypriniformes	GCA_900092035.1	Scaffold	752.539	37.4	4888
*Oxygymnocypris stewartii**	Cypriniformes	GCA_003573665.1	Scaffold	1849.22	-	26281
*Pimephales promelas**	Cypriniformes	GCA_000700825.1	Scaffold	1219.33	41.7	73057
*Poropuntius huangchuchieni**	Cypriniformes	GCA_004124795.1	Scaffold	760.177	35.4	625277
*Sinocyclocheilus anshuiensis**	Cypriniformes	GCF_001515605.1	Scaffold	1632.72	38.0001	85682
*Sinocyclocheilus grahami**	Cypriniformes	GCF_001515645.1	Scaffold	1750.29	38.7	31277
*Sinocyclocheilus rhinocerous**	Cypriniformes	GCF_001515625.1	Scaffold	1655.79	38.1	164173
*Triplophysa siluroides**	Cypriniformes	GCA_006030095.1	Scaffold	583.43	39.5	1002
*Anabarilius grahami**	Cypriniformes	GCA_003731715.1	Scaffold	991.887	36.7	80398
*Austrofundulus limnaeus**	Cyprinodontiformes	GCF_001266775.1	Scaffold	866.963	41.1	29785
*Cyprinodon nevadensis pectoralis**	Cyprinodontiformes	GCA_000776015.1	Scaffold	1011.85	39	96516

*Cyprinodon variegatus**	Cyprinodontiformes	GCF_000732505.1	Scaffold	1035.18	39.5001	9259
*Fundulus heteroclitus**	Cyprinodontiformes	GCF_000826765.1	Scaffold	1021.9	41.2	10180
Gambusia affinis	Cyprinodontiformes	GCA_003097735.1	Scaffold	598.663	38.2	2943
*Kryptolebias marmoratus**	Cyprinodontiformes	GCF_001649575.1	Scaffold	680.367	40.6001	3073
*Nothobranchius furzeri**	Cyprinodontiformes	GCF_001465895.1	Chromosome	1242.52	44.8104	6013
Poecilia latipinna	Cyprinodontiformes	GCF_001443285.1	Scaffold	815.145	40.8	17988
Poecilia mexicana	Cyprinodontiformes	GCF_001443325.1	Scaffold	801.711	40.7	18105
Poecilia reticulata	Cyprinodontiformes	GCF_000633615.1	Chromosome	731.622	40.2816	3029
Xiphophorus couchianus	Cyprinodontiformes	GCF_001444195.1	Chromosome	688.542	39.0399	68
Xiphophorus hellerii	Cyprinodontiformes	GCA_003331165.1	Chromosome	733.127	39.1072	85
Xiphophorus maculatus	Cyprinodontiformes	GCF_002775205.1	Chromosome	704.321	39.0265	102
Poecilia formosa	Cyprinodontiformes	GCF_000485575.1	Scaffold	748.923	39.6	3985
Esox lucius	Esociformes	GCF_004634155.1	Chromosome	940.907	42.2905	811
Arctogadus glacialis	Gadiformes	GCA_900303235.1	Scaffold	428.792	44.5	139389
Bagarius yarrelli	Gadiformes	GCA_005784505.1	Scaffold	570.81	38	541
Bathygadus melanobranchus	Gadiformes	GCA_900302375.1	Scaffold	431.203	-	92290
Boreogadus saida	Gadiformes	GCA_900302515.1	Scaffold	412.07	-	137701
Bregmaceros cantori	Gadiformes	GCA_900302395.1	Scaffold	1144.1	42.9	258566
Brosme brosme	Gadiformes	GCA_900302425.1	Scaffold	412.731	-	114910
Coryphaenoides rupestris	Gadiformes	GCA_002895965.1	Scaffold	829.209	-	47680
Gadiculus argenteus	Gadiformes	GCA_900302595.1	Scaffold	396.767	-	123363
Gadus chalcogrammus	Gadiformes	GCA_900302575.1	Scaffold	448.868	44.6	130159
Gadus morhua	Gadiformes	GCA_000231765.1	Scaffold	824.311	46.3	427427
Laemonema laureysi	Gadiformes	GCA_900303225.1	Scaffold	306.495	-	86525
Lota lota	Gadiformes	GCA_900302385.1	Scaffold	397.499	-	106616
Macrourus berglax	Gadiformes	GCA_900302365.1	Scaffold	399.876	-	118318
Malacocephalus occidentalis	Gadiformes	GCA_900312585.1	Scaffold	350.34	-	95829
Melanogrammus aeglefinus	Gadiformes	GCA_900291075.1	Scaffold	652.791	44.9	8420
Melanonus zugmayeri	Gadiformes	GCA_900302345.1	Scaffold	432.903	-	82409
Merlangius merlangus	Gadiformes	GCA_900323355.1	Scaffold	423.942	44.4	122642
Merluccius capensis	Gadiformes	GCA_900312945.1	Scaffold	414.317	-	110925
Merluccius merluccius	Gadiformes	GCA_900312545.1	Scaffold	401.035	-	102914
Merluccius polli	Gadiformes	GCA_900312625.1	Scaffold	401.149	-	113894
Molva molva	Gadiformes	GCA_900323295.1	Scaffold	437.481	-	111875
Mora moro	Gadiformes	GCA_900303205.1	Scaffold	344.961	-	100621

Muraenolepis marmoratus	Gadiformes	GCA_900302325.1	Scaffold	416.391	-	138928
Phycis blennoides	Gadiformes	GCA_900302315.1	Scaffold	416.767	-	62684
Phycis phycis	Gadiformes	GCA_900302335.1	Scaffold	346.335	-	100771
Pollachius virens	Gadiformes	GCA_900312635.1	Scaffold	394.928	-	116705
Trachyrincus murrayi	Gadiformes	GCA_900323305.1	Scaffold	452.417	-	40927
Trachyrincus scabrus	Gadiformes	GCA_900303215.1	Scaffold	369.862	-	80958
Trisopterus minutus	Gadiformes	GCA_900302415.1	Scaffold	334.717	-	106116
*Boleophthalmus pectinirostris**	Gobiiformes	GCF_000788275.1	Scaffold	955.752	40.1001	16620
*Lesueurigobius sanzi**	Gobiiformes	GCA_900303255.1	Scaffold	810.626	-	130360
Periophthalmus magnuspinnatus	Gobiiformes	GCA_000787105.1	Scaffold	701.697	40	26060
Scartelaos histophorus	Gobiiformes	GCA_000787155.1	Scaffold	695.009	39.1	156044
Periophthalmodon schlosseri	Gobiiformes	GCA_000787095.1	Scaffold	679.761	40.2	46662
*Electrophorus electricus**	Gymnotiformes	GCF_003665695.1	Scaffold	551.881	42.8	8786
*Holocentrus rufus**	Holocentriformes	GCA_900302615.1	Scaffold	649.757	-	58113
*Myripristis jacobus**	Holocentriformes	GCA_900302555.1	Scaffold	720.397	-	64974
Neoniphon sammara	Holocentriformes	GCA_900302535.1	Scaffold	659.226	-	56655
*Labrus bergylta**	Labriformes	GCF_900080235.1	Scaffold	805.481	40.9	13466
*Symphodus melops**	Labriformes	GCA_002819105.1	Contig	614.569	41.6	5060
*Lampris guttatus**	Lampridiformes	GCA_900302545.1	Scaffold	849.278	44.3	208230
*Regalecus glesne**	Lampridiformes	GCA_900302585.1	Scaffold	656.004	-	105196
Antennarius striatus	Lophiiformes	GCA_900303275.1	Scaffold	441.857	40.8	70398
*Planiliza haematocheilus**	Mugiliformes	GCA_005024645.1	Contig	747.343	42.4	1453
*Benthosema glaciale**	Myctophiformes	GCA_900323375.1	Scaffold	676.314	-	143923
*Brotula barbata**	Ophidiiformes	GCA_900303265.1	Scaffold	485.061	-	29854
Carapus acus	Ophidiiformes	GCA_900312935.1	Scaffold	387.834	-	46699
*Lamprogrammus exutus**	Ophidiiformes	GCA_900312555.1	Scaffold	492.85	-	120937
Chiloscyllium plagiosum	Orectolobiformes	GCA_004010195.1	Chromosome	3776.55	42.0371	601049
Chiloscyllium punctatum	Orectolobiformes	GCA_003427335.1	Scaffold	3375.7	37.6	280241
Rhincodon typus	Orectolobiformes	GCF_001642345.1	Scaffold	2931.6	41.8	57334
Osmerus eperlanus	Osmeriformes	GCA_900302275.1	Scaffold	342.759	-	73274
Arapaima gigas	Osteoglossiformes	GCA_900497675.1	Scaffold	661.279	-	5301
*Paramormyrops kingsleyae**	Osteoglossiformes	GCF_002872115.1	Scaffold	799.421	44.4	4667
Scleropages formosus	Osteoglossiformes	GCF_900964775.1	Chromosome	784.563	44.1597	72
*Lateolabrax maculatus**	Pempheriformes	GCA_004023545.1	Chromosome	519.239	41.0772	24
*Gouania willdenowi**	Perciformes	GCF_900634775.1	Chromosome	937.151	38.4531	441

Parablennius parvicornis	Perciformes	GCA_900302745.1	Scaffold	599.249	-	66539
*Acanthochromis polyacanthus**	Perciformes	GCF_002109545.1	Scaffold	991.585	41.5	30414
Amphiprion ocellaris	Perciformes	GCF_002776465.1	Scaffold	880.721	39.5001	6405
Amphiprion percula	Perciformes	GCA_003047355.2	Chromosome	908.956	39.5241	366
*Anarrhichthys ocellatus**	Perciformes	GCA_004355925.1	Scaffold	612.774	42.6	10816
Anoplopoma fimbria	Perciformes	GCA_000499045.1	Contig	699.326	40.3	208506
Chaenocephalus aceratus	Perciformes	GCA_900302675.1	Scaffold	623.452	-	114963
*Chromis chromis**	Perciformes	GCA_900302755.1	Scaffold	834.429	-	104482
Collichthys lucidus	Perciformes	GCA_004119915.1	Chromosome	877.615	41.8051	854
Cottoperca gobio	Perciformes	GCF_900634415.1	Chromosome	609.392	40.9976	322
Cottus rhenanus	Perciformes	GCA_001455555.1	Scaffold	563.609	36.8	164693
*Dicentrarchus labrax**	Perciformes	GCA_000689215.1	Scaffold	675.917	40.4	25
*Epinephelus lanceolatus**	Perciformes	GCA_005281545.1	Chromosome	1087.4	0.58246	4200
*Epinephelus moara**	Perciformes	GCA_006386435.1	Chromosome	1030.46	41.53	4562
Gasterosteus aculeatus	Perciformes	GCA_006229165.1	Scaffold	467.45	43.5	10242
*Hypoplectrus puella**	Perciformes	GCA_900610375.1	Scaffold	612.29	39.5	14375
Larimichthys crocea	Perciformes	GCF_000972845.2	Chromosome	657.94	41.3589	9998
Lates calcarifer	Perciformes	GCF_001640805.1	Scaffold	668.481	40.8001	3808
Liparis tanakae	Perciformes	GCA_006348945.1	Scaffold	498.98	40.7	27878
Miichthys miiuy	Perciformes	GCA_001593715.1	Scaffold	619.301	39.3	6294
Morone chrysops	Perciformes	GCA_003610055.1	Scaffold	620.984	40	84096
*Morone saxatilis**	Perciformes	GCA_004916995.1	Scaffold	598.11	39.8	629
Myoxocephalus scorpius	Perciformes	GCA_900312955.1	Scaffold	520.316	41.9	85863
Nibea albiflora	Perciformes	GCA_900327885.1	Scaffold	574.466	-	11977
Notothenia coriiceps	Perciformes	GCF_000735185.1	Scaffold	636.614	40.8002	38657
Ophiodon elongatus	Perciformes	GCA_004358465.1	Scaffold	635.568	42.3	18379
Parambassis ranga	Perciformes	GCF_900634625.1	Chromosome	551.013	42.5376	156
*Perca flavescens**	Perciformes	GCF_004354835.1	Chromosome	877.456	40.8426	268
*Perca fluviatilis**	Perciformes	GCA_003412525.1	Scaffold	958.225	45.1	31105
*Pseudochromis fuscus**	Perciformes	GCA_900323345.1	Scaffold	657.041	-	52042
Pungitius pungitius	Perciformes	GCA_003399555.1	Scaffold	441.09	37.9	7847
*Sebastes aleutianus**	Perciformes	GCA_001910805.2	Scaffold	899.65	40.9	10489
*Sebastes koreanus**	Perciformes	GCA_004335335.1	Contig	725.092	40.6	147157
*Sebastes minor**	Perciformes	GCA_001910765.2	Scaffold	681.653	40.8	166448
*Sebastes nigrocinctus**	Perciformes	GCA_000475235.3	Scaffold	746.045	40.8	15872

*Sebastes norvegicus**	Perciformes	GCA_900302655.1	Scaffold	717.741	-	75627
*Sebastes nudus**	Perciformes	GCA_004335365.1	Contig	724.045	40.6	180312
*Sebastes rubrivinctus**	Perciformes	GCA_000475215.1	Scaffold	756.297	40.7	68206
*Sebastes schlegelii**	Perciformes	GCA_004335315.1	Contig	728.477	40.6	146105
*Sebastes steindachneri**	Perciformes	GCA_001910785.2	Scaffold	648.011	40.7	279232
*Stegastes partitus**	Perciformes	GCF_000690725.1	Scaffold	800.492	42.1	5818
Pseudoliparis sp. Yap Trench	Perciformes	GCA_004335475.1	Scaffold	840.641	43.8	3134
Percopsis transmontana	Percopsiformes	GCA_900302285.1	Scaffold	458.089	-	53197
*Typhlichthys subterraneus**	Percopsiformes	GCA_900302405.1	Scaffold	555.56	41.6	84841
*Cynoglossus semilaevis**	Pleuronectiformes	GCF_000523025.1	Chromosome	470.199	41.2788	31181
Paralichthys olivaceus	Pleuronectiformes	GCF_001970005.1	Scaffold	643.911	43.2001	9525
*Pseudopleuronectes yokohamae**	Pleuronectiformes	GCA_000787555.1	Contig	547.831	42	525502
*Reinhardtius hippoglossoides**	Pleuronectiformes	GCA_006182925.1	Scaffold	677.54	42.7	4,453
*Scophthalmus maximus**	Pleuronectiformes	GCA_003186165.1	Chromosome	524.979	43.431	22
Polymixia japonica	Polymixiiformes	GCA_900302305.1	Scaffold	554.896	42.6	92198
Erpetoichthys calabaricus	Polypteriformes	GCF_900747795.1	Chromosome	3811.04	39.4067	1885
Leucoraja erinacea	Rajiformes	GCA_000238235.1	Contig	1555.46	40.3	2962366
*Hucho hucho**	Salmoniformes	GCA_003317085.1	Scaffold	2487.55	32.9	71639
*Oncorhynchus kisutch**	Salmoniformes	GCF_002021735.1	Chromosome	2369.93	43.6225	22813
*Oncorhynchus mykiss**	Salmoniformes	GCF_002163495.1	Chromosome	2179	43.4875	139800
*Oncorhynchus nerka**	Salmoniformes	GCF_006149115.1	Chromosome	1927.14	43.63	38027
*Oncorhynchus tshawytscha**	Salmoniformes	GCF_002872995.1	Chromosome	2425.71	43.7341	15946
*Salmo salar**	Salmoniformes	GCF_000233375.1	Chromosome	2966.89	43.8912	241573
*Salmo trutta**	Salmoniformes	GCF_901001165.1	Chromosome	2371.88	43.57	1441
*Salvelinus alpinus**	Salmoniformes	GCF_002910315.2	Chromosome	2169.55	43.6053	16702
*Thymallus thymallus**	Salmoniformes	GCA_004348285.1	Chromosome	1564.83	41.9651	3831
*Pampus argenteus**	Scombriformes	GCA_000697985.1	Scaffold	350.449	38.2	298139
*Thunnus albacares**	Scombriformes	GCA_900302625.1	Scaffold	728.212	-	38995
*Thunnus orientalis**	Scombriformes	GCA_000418415.1	Contig	684.497	39.7	133062
*Thunnus thynnus**	Scombriformes	GCA_003231725.1	Scaffold	648.209	39.3	354425
*Lepisosteus oculatus**	Semionotiformes	GCF_000242695.1	Chromosome	945.878	40.4083	2106
*Ageneiosus marmoratus**	Siluriformes	GCA_003347165.1	Scaffold	1030	26.5	16063
*Clarias batrachus**	Siluriformes	GCA_003987875.1	Scaffold	821.75	-	10042
*Ictalurus punctatus**	Siluriformes	GCF_001660625.1	Chromosome	783.275	39.8391	9975
*Pangasianodon hypophthalmus**	Siluriformes	GCF_003671635.1	Scaffold	715.76	39.0001	567

	*Tachysurus fulvidraco**	Siluriformes	GCF_003724035.1	Scaffold	713.811	39.6001	663
	*Pagrus major**	Spariformes	GCA_002897255.1	Scaffold	875.465	38	886260
	*Sparus aurata**	Spariformes	GCA_003309015.1	Chromosome	830.382	41.5049	34642
	*Spondyliosoma cantharus**	Spariformes	GCA_900302685.1	Scaffold	680.472	-	47064
	*Borostomias antarcticus**	Stomiiformes	GCA_900323325.1	Scaffold	430.363	-	104734
	*Stylephorus chordatus**	Stylephoriformes	GCA_900312615.1	Scaffold	488.489	-	128468
	*Mastacembelus armatus**	Synbranchiformes	GCF_900324485.1	Scaffold	593.082	40.9001	126
	*Monopterus albus**	Synbranchiformes	GCF_001952655.1	Scaffold	684.144	41.5001	20622
	Hippocampus comes	Syngnathiformes	GCF_001891065.1	Scaffold	493.776	43.6998	37377
	Syngnathus acus	Syngnathiformes	GCA_901709675.1	Chromosome	324.33	43.45	87
	*Mola mola**	Tetraodontiformes	GCA_001698575.1	Scaffold	639.452	41.2	5552
	Takifugu bimaculatus	Tetraodontiformes	GCA_004026145.1	Chromosome	371.676	45.4074	22
	Takifugu flavidus	Tetraodontiformes	GCA_003711565.1	Chromosome	366.288	45.623	867
	Takifugu rubripes	Tetraodontiformes	GCF_000180615.1	Chromosome	391.485	45.8414	7091
	Tetraodon nigroviridis	Tetraodontiformes	GCA_000180735.1	Scaffold	342.403	46.6	25773
	*Cyttopsis rosea**	Zeiformes	GCA_900302355.1	Scaffold	546.506	43	111974
	*Zeus faber**	Zeiformes	GCA_900323335.1	Scaffold	610.433	-	135758
mammals (336)	*Chrysochloris asiatica*	Afrosoricida	GCF_000296735.1	Scaffold	4210.11	41.8	20500
	Echinops telfairi	Afrosoricida	GCF_000313985.1	Scaffold	2947.02	43.6	8402
	Microgale talazaci	Afrosoricida	GCA_004026705.1	Scaffold	3236.15	43	788895
	Acinonyx jubatus	Carnivora	GCF_003709585.1	Scaffold	2384.85	41.6	3220
	Ailuropoda melanoleuca	Carnivora	GCF_000004335.2	Scaffold	2299.51	41.7	81467
	Ailurus fulgens styani	Carnivora	GCA_002007465.1	Scaffold	2342.94	-	11589
	Callorhinus ursinus	Carnivora	GCF_003265705.1	Scaffold	2706.87	41.6	14230
	Canis lupus dingo	Carnivora	GCF_003254725.1	Scaffold	2439.83	41.4	2444
	Canis lupus familiaris	Carnivora	GCF_000002285.3	Chromosome	2410.98	41.3022	3310
	Cryptoprocta ferox	Carnivora	GCA_004023885.1	Scaffold	2458.85	41.5	464036
	Enhydra lutris kenyoni	Carnivora	GCF_002288905.1	Scaffold	2455.28	41.6	6771
	Eumetopias jubatus	Carnivora	GCF_004028035.1	Scaffold	2418.26	41.4	7472
	Felis catus	Carnivora	GCF_000181335.3	Chromosome	2521.86	41.8806	4525
	Felis nigripes	Carnivora	GCA_004023925.1	Scaffold	2428.42	41.9	590124
	Gulo gulo	Carnivora	GCA_900006375.2	Scaffold	2423.18	37.6	47417
	Helogale parvula	Carnivora	GCA_004023845.1	Scaffold	2392.47	41.8	189370
	Hyaena hyaena	Carnivora	GCA_003009895.1	Scaffold	2374.72	41.3	5760
	Leptonychotes weddellii	Carnivora	GCF_000349705.1	Scaffold	3156.9	43.8	16711

Lycaon pictus	Carnivora	GCA_001887905.1	Chromosome	2358.14	41.1817	803
Lynx pardinus	Carnivora	GCA_900661375.1	Scaffold	2413.21	46.3	41700
Mellivora capensis	Carnivora	GCA_004024625.1	Scaffold	3091.17	42	2068312
Mirounga angustirostris	Carnivora	GCA_004023865.1	Scaffold	2407.32	41.5	321256
Mungos mungo	Carnivora	GCA_004023785.1	Scaffold	2448.21	42	312364
Mustela putorius furo	Carnivora	GCF_000215625.1	Scaffold	2410.88	41.9	7783
Neomonachus schauinslandi	Carnivora	GCF_002201575.1	Scaffold	2400.93	41.6	7873
Neovison vison	Carnivora	GCA_900108605.1	Scaffold	2447.19	-	7175
Odobenus rosmarus divergens	Carnivora	GCF_000321225.1	Scaffold	2400.15	41.7	3893
Panthera onca	Carnivora	GCA_004023805.1	Scaffold	2504.53	41.7	438792
Panthera pardus	Carnivora	GCF_001857705.1	Scaffold	2578.02	41.9	50377
Panthera tigris altaica	Carnivora	GCF_000464555.1	Scaffold	2391.08	41.5	1479
Paradoxurus hermaphroditus	Carnivora	GCA_004024585.1	Scaffold	2477.23	42.5	438125
Phoca vitulina	Carnivora	GCA_004348235.1	Scaffold	2363.59	41.4	5542
Prionailurus bengalensis euptilurus	Carnivora	GCA_005406085.1	Scaffold	2435.35	41.1	65356
Pteronura brasiliensis	Carnivora	GCA_004024605.1	Scaffold	2602.28	41.7	644728
Puma concolor	Carnivora	GCF_003327715.1	Scaffold	2432.97	41.8	2175
Spilogale gracilis	Carnivora	GCA_004023965.1	Scaffold	2463.32	41.6	422009
Suricata suricatta	Carnivora	GCA_004023905.1	Scaffold	2471.81	41.9	349868
Taxidea taxus jeffersonii	Carnivora	GCA_003697995.1	Scaffold	2416.01	41.2	145161
Ursus americanus	Carnivora	GCA_003344425.1	Scaffold	2588.39	39.5	111495
Ursus arctos horribilis	Carnivora	GCF_003584765.1	Scaffold	2328.66	41.9	6673
Ursus maritimus	Carnivora	GCF_000687225.1	Scaffold	2301.38	41.7	23819
Vulpes lagopus	Carnivora	GCA_004023825.1	Scaffold	2688.56	41.9	1263345
Vulpes vulpes	Carnivora	GCF_003160815.1	Scaffold	2421.57	41.3	82424
Zalophus californianus	Carnivora	GCF_900631625.1	Scaffold	2372.37	41.4	10423
Aepyceros melampus	Cetartiodactyla	GCA_006408695.1	Scaffold	2631.3	41.4	24102
Alcelaphus buselaphus	Cetartiodactyla	GCA_006408545.1	Scaffold	3882.34	9.3	539659
Ammotragus lervia	Cetartiodactyla	GCA_002201775.1	Scaffold	2650.84	40.3	77320
Axis porcinus	Cetartiodactyla	GCA_003798545.1	Scaffold	2676.16	41	136093
Balaenoptera acutorostrata scammoni	Cetartiodactyla	GCF_000493695.1	Scaffold	2431.69	41.4	10776
Balaenoptera bonaerensis	Cetartiodactyla	GCA_000978805.1	Scaffold	2234.64	40.7	421444
Beatragus hunteri	Cetartiodactyla	GCA_004027495.1	Scaffold	2697.26	42.3	535889
Bison bison bison	Cetartiodactyla	GCF_000754665.1	Scaffold	2828.03	42.2	128431
Bos grunniens	Cetartiodactyla	GCA_005887515.1	Chromosome	2832.78	41.966	414

Bos indicus	Cetartiodactyla	GCF_000247795.1	Chromosome	2673.97	42.3689	32
Bos mutus	Cetartiodactyla	GCF_000298355.1	Scaffold	2645.16	42	41192
Bos taurus	Cetartiodactyla	GCF_002263795.1	Chromosome	2715.85	41.9362	2211
Bubalus bubalis	Cetartiodactyla	GCF_003121395.1	Chromosome	2655.78	41.814	509
Camelus bactrianus	Cetartiodactyla	GCF_000767855.1	Scaffold	1992.66	41.4	35455
Camelus dromedarius	Cetartiodactyla	GCF_000767585.1	Scaffold	2004.06	41.3	32573
Camelus ferus	Cetartiodactyla	GCF_000311805.1	Scaffold	2009.19	41.3	13334
Capra aegagrus	Cetartiodactyla	GCA_000978405.1	Chromosome	2828.87	42.4124	89498
Capra hircus	Cetartiodactyla	GCF_001704415.1	Chromosome	2922.81	43.1505	29907
Capra ibex	Cetartiodactyla	GCA_006410555.1	Scaffold	2701.36	42.2	55914
Capra sibirica	Cetartiodactyla	GCA_003182615.2	Scaffold	2733.03	42.2	85609
Capreolus capreolus	Cetartiodactyla	GCA_000751575.1	Scaffold	2785.38	41.8	3088511
Catagonus wagneri	Cetartiodactyla	GCA_004024745.1	Scaffold	2593.59	42.3	1230162
Cephalophus harveyi	Cetartiodactyla	GCA_006410635.1	Scaffold	2816.64	41.1	85160
Cervus elaphus hippelaphus	Cetartiodactyla	GCA_002197005.1	Chromosome	3438.62	44.8578	11479
Connochaetes taurinus	Cetartiodactyla	GCA_006408615.1	Scaffold	2644.04	41.1	78350
Damaliscus lunatus	Cetartiodactyla	GCA_006408505.1	Scaffold	3127.76	39	92109
Delphinapterus leucas	Cetartiodactyla	GCF_002288925.1	Scaffold	2358.52	41.3	6972
Elaphurus davidianus	Cetartiodactyla	GCA_002443075.1	Scaffold	2584.69	40.5	194812
Eschrichtius robustus	Cetartiodactyla	GCA_002189225.1	Scaffold	2849.45	30.7	57203
Eubalaena japonica	Cetartiodactyla	GCA_004363455.1	Scaffold	2693.58	41.1	1353963
Eudorcas thomsonii	Cetartiodactyla	GCA_006408755.1	Scaffold	2904.27	41.7	123308
Giraffa camelopardalis	Cetartiodactyla	GCA_006408565.1	Scaffold	2473.18	40.2	78945
Giraffa tippelskirchi	Cetartiodactyla	GCA_001651235.1	Scaffold	2705.07	41.4	513177
Hemitragus hylocrius	Cetartiodactyla	GCA_004026825.1	Scaffold	2715.04	42.4	469163
Hippopotamus amphibius	Cetartiodactyla	GCA_002995585.1	Scaffold	2425.75	38.8	128772
Hydropotes inermis	Cetartiodactyla	GCA_006459105.1	Scaffold	2530.18	40.7	22246
Inia geoffrensis	Cetartiodactyla	GCA_004363515.1	Scaffold	2596.28	42.1	1213610
Kobus ellipsiprymnus	Cetartiodactyla	GCA_006410655.1	Scaffold	2895.34	39.6	88848
Kogia breviceps	Cetartiodactyla	GCA_004363705.1	Scaffold	2759.7	41.9	1252072
Lagenorhynchus obliquidens	Cetartiodactyla	GCF_003676395.1	Scaffold	2334.47	41.3	5422
Lipotes vexillifer	Cetartiodactyla	GCF_000442215.1	Scaffold	2429.21	41.4	30713
Litocranius walleri	Cetartiodactyla	GCA_006410535.1	Scaffold	2979.34	40.4	79667
Madoqua kirkii	Cetartiodactyla	GCA_006408675.1	Scaffold	2648.13	42	216116
Megaptera novaeangliae	Cetartiodactyla	GCA_004329385.1	Scaffold	2265.79	38.6	2558

Mesoplodon bidens	Cetartiodactyla	GCA_004027085.1	Scaffold	2797.69	41.7	1801720
Monodon monoceros	Cetartiodactyla	GCF_005190385.1	Scaffold	2355.57	41	6972
Moschus berezovskii	Cetartiodactyla	GCA_006459085.1	Scaffold	2818.01	39.8	77520
Moschus chrysogaster	Cetartiodactyla	GCA_006461725.1	Scaffold	4972.48	19.4	127198
Moschus moschiferus	Cetartiodactyla	GCA_004024705.1	Scaffold	2966.36	42.3	1175703
Muntiacus crinifrons	Cetartiodactyla	GCA_006408485.1	Scaffold	2681.2	41.5	21054
Muntiacus muntjak	Cetartiodactyla	GCA_006409035.1	Scaffold	2703.71	40.1	383561
Muntiacus reevesi	Cetartiodactyla	GCA_006408525.1	Scaffold	2601.89	40.9	200261
Nanger granti	Cetartiodactyla	GCA_006408635.1	Scaffold	3032.78	35.7	62189
Neophocaena asiaeorientalis asiaeorientalis	Cetartiodactyla	GCF_003031525.1	Scaffold	2284.63	41.3	13699
Neotragus moschatus	Cetartiodactyla	GCA_006410615.1	Scaffold	2675.73	40.9	218204
Neotragus pygmaeus	Cetartiodactyla	GCA_006410875.1	Scaffold	2895.87	40.5	354089
Odocoileus hemionus hemionus	Cetartiodactyla	GCA_004115125.1	Scaffold	2343.7	36.9	16208
Odocoileus virginianus texanus	Cetartiodactyla	GCF_002102435.1	Scaffold	2380.51	41.6	17026
Okapia johnstoni	Cetartiodactyla	GCA_001660835.1	Scaffold	2878.13	41.5	837179
Orcinus orca	Cetartiodactyla	GCF_000331955.2	Scaffold	2372.92	41.7	1668
Oreotragus oreotragus	Cetartiodactyla	GCA_006410675.1	Scaffold	3247.63	37.7	24892
Oryx gazella	Cetartiodactyla	GCA_003945745.1	Scaffold	2762.92	39	31177
Ourebia ourebi	Cetartiodactyla	GCA_006417275.1	Scaffold	2524.87	42	3508661
Ovis ammon	Cetartiodactyla	GCA_003121645.1	Scaffold	2643.43	41.6	41955
Ovis aries	Cetartiodactyla	GCF_002742125.1	Chromosome	2869.91	41.9088	2641
Ovis canadensis	Cetartiodactyla	GCA_004026945.1	Scaffold	2863.66	43.2	1048136
Pantholops hodgsonii	Cetartiodactyla	GCF_000400835.1	Scaffold	2696.89	42.4	15059
Philantomba maxwellii	Cetartiodactyla	GCA_006410695.1	Scaffold	3145.94	30.5	25633
Phocoena phocoena	Cetartiodactyla	GCA_003071005.1	Scaffold	2441.14	37.2	142029
Physeter catodon	Cetartiodactyla	GCF_002837175.2	Chromosome	2512.15	41.7225	14677
Platanista minor	Cetartiodactyla	GCA_004363435.1	Scaffold	2671.07	41.7	1098790
Pontoporia blainvillei	Cetartiodactyla	GCA_004363935.1	Scaffold	1685.1	46.7	1885058
Procapra przewalskii	Cetartiodactyla	GCA_006410515.1	Scaffold	2692.99	40.7	274130
Przewalskium albirostris	Cetartiodactyla	GCA_006408465.1	Scaffold	2692.23	41.1	171874
Pseudois nayaur	Cetartiodactyla	GCA_003182575.1	Scaffold	2584.17	42.1	214772
Rangifer tarandus	Cetartiodactyla	GCA_004026565.1	Scaffold	2897.3	42.1	1360739
Raphicerus campestris	Cetartiodactyla	GCA_006410735.1	Scaffold	3137.86	34.8	57517
Redunca redunca	Cetartiodactyla	GCA_006410935.1	Scaffold	2724.98	41.3	534502
Saiga tatarica	Cetartiodactyla	GCA_004024985.1	Scaffold	2876.19	42.5	1772199

Sousa chinensis	Cetartiodactyla	GCA_003521335.2	Scaffold	2339.09	40.9	20903
Sus scrofa	Cetartiodactyla	GCF_000003025.6	Chromosome	2501.91	41.971	706
Sylvicapra grimmia	Cetartiodactyla	GCA_006408735.1	Scaffold	3145.09	34.8	22324
Syncerus caffer	Cetartiodactyla	GCA_006408785.1	Scaffold	2929.43	40.9	148371
Tragelaphus buxtoni	Cetartiodactyla	GCA_006411685.1	Scaffold	3267.49	36.2	2910462
Tragelaphus eurycerus	Cetartiodactyla	GCA_006410755.1	Scaffold	2962.04	36.2	1539703
Tragelaphus imberbis	Cetartiodactyla	GCA_006410775.1	Scaffold	4038.24	32.5	272415
Tragelaphus scriptus	Cetartiodactyla	GCA_006410495.1	Scaffold	2873.33	40.2	237860
Tragelaphus spekii	Cetartiodactyla	GCA_006411015.1	Scaffold	3725.95	36.2	1769205
Tragelaphus strepsiceros	Cetartiodactyla	GCA_006410795.1	Scaffold	2867.75	40.7	320432
Tragulus javanicus	Cetartiodactyla	GCA_004024965.1	Scaffold	2582.63	42.7	1156990
Tragulus kanchil	Cetartiodactyla	GCA_006408655.1	Scaffold	3053.42	33.9	185856
Tursiops aduncus	Cetartiodactyla	GCA_003227395.1	Scaffold	2503.93	-	16249
Tursiops truncatus	Cetartiodactyla	GCF_001922835.1	Scaffold	2132.52	40.9	2648
Vicugna pacos	Cetartiodactyla	GCF_000164845.2	Scaffold	2172.21	41.7	276725
Ziphius cavirostris	Cetartiodactyla	GCA_004364475.1	Scaffold	3150.42	42	3758276
Antidorcas marsupialis	Cetartiodactyla	GCA_006408585.1	Scaffold	3024.61	38.9	94522
Antilocapra americana	Cetartiodactyla	GCA_004027515.1	Scaffold	2899.77	42	1643131
Artibeus jamaicensis	Chiroptera	GCA_004027435.1	Scaffold	2424.78	42.2	794376
Carollia perspicillata	Chiroptera	GCA_004027735.1	Scaffold	2689.41	42.7	1925339
Craseonycteris thonglongyai	Chiroptera	GCA_004027555.1	Scaffold	2272.09	41	1224256
Desmodus rotundus	Chiroptera	GCF_002940915.1	Scaffold	2063.81	42.3	29801
Eidolon helvum	Chiroptera	GCA_000465285.1	Scaffold	1837.75	39.2	133538
Eonycteris spelaea	Chiroptera	GCA_003508835.1	Contig	1966.86	40.3	4469
Eptesicus fuscus	Chiroptera	GCF_000308155.1	Scaffold	2026.63	43.5	6789
Hipposideros armiger	Chiroptera	GCF_001890085.1	Scaffold	2236.58	42.2	7571
Hipposideros galeritus	Chiroptera	GCA_004027415.1	Scaffold	2440.75	41.3	840200
Lasiurus borealis	Chiroptera	GCA_004026805.1	Scaffold	2857.59	42.7	518900
Macroglossus sobrinus	Chiroptera	GCA_004027375.1	Scaffold	1897.64	39.6	171985
Megaderma lyra	Chiroptera	GCA_004026885.1	Scaffold	2620.95	40.8	1902801
Micronycteris hirsuta	Chiroptera	GCA_004026765.1	Scaffold	2314.65	42.4	550090
Miniopterus natalensis	Chiroptera	GCF_001595765.1	Scaffold	1803.1	42.4	1269
Miniopterus schreibersii	Chiroptera	GCA_004026525.1	Scaffold	1775.85	42.3	177620
Mormoops blainvillei	Chiroptera	GCA_004026545.1	Scaffold	2111.75	41.7	205259
Murina aurata feae	Chiroptera	GCA_004026665.1	Scaffold	2331.55	43.8	880177

Myotis brandtii	Chiroptera	GCF_000412655.1	Scaffold	2107.24	42.9	169750
Myotis davidii	Chiroptera	GCF_000327345.1	Scaffold	2059.8	43.1	101769
Myotis lucifugus	Chiroptera	GCF_000147115.1	Scaffold	2034.58	42.7	11654
Myotis myotis	Chiroptera	GCA_004026985.1	Scaffold	2294.47	43.2	1045197
Noctilio leporinus	Chiroptera	GCA_004026585.1	Scaffold	2098.5	40.9	298222
Phyllostomus discolor	Chiroptera	GCF_004126475.1	Chromosome	2117.76	42.4541	141
Pipistrellus pipistrellus	Chiroptera	GCA_004026625.1	Scaffold	2090.67	42.9	726104
Pteronotus parnellii	Chiroptera	GCA_000465405.1	Scaffold	1960.32	40.8	177401
Pteropus alecto	Chiroptera	GCF_000325575.1	Scaffold	1985.98	39.9	65598
Pteropus vampyrus	Chiroptera	GCF_000151845.1	Scaffold	2198.28	40.5	36094
Rhinolophus ferrumequinum	Chiroptera	GCA_004115265.2	Chromosome	2075.79	40.5699	135
Rousettus aegyptiacus	Chiroptera	GCF_001466805.2	Scaffold	1910.25	40	2490
Tadarida brasiliensis	Chiroptera	GCA_004025005.1	Scaffold	2709.93	42.4	1067615
Tonatia saurophila	Chiroptera	GCA_004024845.1	Scaffold	2105.89	42	249810
Anoura caudifer	Chiroptera	GCA_004027475.1	Scaffold	2206.59	41.7	337255
Chaetophractus vellerosus	Cingulata	GCA_004027955.1	Scaffold	5335.6	39.7	7677333
Dasypus novemcinctus	Cingulata	GCF_000208655.1	Scaffold	3631.52	41.5	46559
Tolypeutes matacus	Cingulata	GCA_004025125.1	Scaffold	4115.03	39.2	2925507
Sarcophilus harrisii	Dasyuromorphia	GCF_000189315.1	Scaffold	3174.69	37	35974
Galeopterus variegatus	Dermoptera	GCF_000696425.1	Scaffold	3187.66	41.2	179514
Monodelphis domestica	Didelphimorphia	GCF_000002295.2	Chromosome	3598.44	38.1446	5223
Notamacropus eugenii	Diprotodontia	GCA_000004035.1	Scaffold	3075.18	40.4	277711
Phascolarctos cinereus	Diprotodontia	GCF_002099425.1	Contig	3192.58	39.1	1907
Vombatus ursinus	Diprotodontia	GCF_900497805.2	Scaffold	3486.6	38.9	15416
Condylura cristata	Eulipotyphla	GCF_000260355.1	Scaffold	1769.66	41.9	2040
Crocidura indochinensis	Eulipotyphla	GCA_004027635.1	Scaffold	3082.22	39.8	2415412
Erinaceus europaeus	Eulipotyphla	GCF_000296755.1	Scaffold	2715.72	42.4999	5803
Scalopus aquaticus	Eulipotyphla	GCA_004024925.1	Scaffold	2068.87	42	640296
Solenodon paradoxus	Eulipotyphla	GCA_004363575.1	Scaffold	2109.88	40	176618
Sorex araneus	Eulipotyphla	GCF_000181275.1	Scaffold	2423.16	43.4	12845
Uropsilus gracilis	Eulipotyphla	GCA_004024945.1	Scaffold	2112.71	41.1	723835
Heterohyrax brucei	Hyracoidea	GCA_004026845.1	Scaffold	3618.15	40.9	967933
Procavia capensis	Hyracoidea	GCA_000152225.2	Scaffold	3602.18	41.8	65694
Lepus americanus	Lagomorpha	GCA_004026855.1	Scaffold	3123.27	44.1	1592829
Ochotona princeps	Lagomorpha	GCF_000292845.1	Scaffold	2229.84	44	10421

Oryctolagus cuniculus	Lagomorpha	GCF_000003625.3	Chromosome	2737.46	44.0526	3318
Elephantulus edwardii	Macroscelidea	GCF_000299155.1	Scaffold	3843.98	41.5	8768
Ornithorhynchus anatinus	Monotremata	GCF_004115215.1	Chromosome	1858.55	46.1699	305
Ceratotherium simum simum	Perissodactyla	GCF_000283155.1	Scaffold	2464.37	41.2	3087
Dicerorhinus sumatrensis sumatrensis	Perissodactyla	GCA_002844835.1	Scaffold	2949.86	39.2	1061589
Diceros bicornis	Perissodactyla	GCA_004027315.1	Scaffold	3020.7	41.2	2022658
Equus asinus	Perissodactyla	GCF_001305755.1	Scaffold	2391.05	41.4	2167
Equus caballus	Perissodactyla	GCF_002863925.1	Chromosome	2506.97	41.5271	4701
Equus przewalskii	Perissodactyla	GCF_000696695.1	Scaffold	2395.95	41.3	53097
Tapirus indicus	Perissodactyla	GCA_004024905.1	Scaffold	2425.2	40.9	280135
Tapirus terrestris	Perissodactyla	GCA_004025025.1	Scaffold	2571.46	41.1	803506
Manis javanica	Pholidota	GCF_001685135.1	Scaffold	2547.4	41.5	80670
Manis pentadactyla	Pholidota	GCA_000738955.1	Scaffold	2204.73	41.6	92772
Manis tricuspis	Pholidota	GCA_004765945.1	Scaffold	3013.4	40.8	2116242
Bradypus variegatus	Pilosa	GCA_004027775.1	Scaffold	1502.48	42.8	1771805
Choloepus didactylus	Pilosa	GCA_004027855.1	Scaffold	3390.66	39.8	2532333
Choloepus hoffmanni	Pilosa	GCA_000164785.2	Scaffold	3286.01	40	269084
Myrmecophaga tridactyla	Pilosa	GCA_004026745.1	Scaffold	3547.27	39.5	1621408
Tamandua tetradactyla	Pilosa	GCA_004025105.1	Scaffold	3849.37	39.5	2144204
Cercocebus atys	Primates	GCF_000955945.1	Scaffold	2848.25	41.1	11433
Cercopithecus neglectus	Primates	GCA_004027615.1	Scaffold	3278.86	41.2	2119200
Chlorocebus sabaeus	Primates	GCF_000409795.2	Chromosome	2789.66	40.9303	2022
Colobus angolensis palliatus	Primates	GCF_000951035.1	Scaffold	2970.12	41.6	13124
Erythrocebus patas	Primates	GCA_004027335.1	Scaffold	3490.49	41.2	2563960
Gorilla gorilla gorilla	Primates	GCF_000151905.2	Chromosome	3063.36	41.2224	40730
Homo sapiens	Primates	GCF_000001405.39	Chromosome	3272.09	41.48	919
Macaca fascicularis	Primates	GCF_000364345.1	Chromosome	2946.84	41.3398	7625
Macaca fuscata fuscata	Primates	GCA_003118495.1	Contig	2930.71	40.9	90905
Macaca mulatta	Primates	GCF_003339765.1	Chromosome	2971.33	41.1015	2979
Macaca nemestrina	Primates	GCF_000956065.1	Scaffold	2948.7	41.3	9733
Mandrillus leucophaeus	Primates	GCF_000951045.1	Scaffold	3061.99	41.6	12821
Mandrillus sphinx	Primates	GCA_004802615.1	Scaffold	2860.58	39.6	53550
Nasalis larvatus	Primates	GCA_000772465.1	Chromosome	3011.97	42.2536	319549
Nomascus leucogenys	Primates	GCF_000146795.2	Chromosome	2962.06	41.3958	17524
Pan paniscus	Primates	GCF_000258655.2	Chromosome	3286.64	42.3185	10984

Pan troglodytes	Primates	GCF_002880755.1	Chromosome	3050.4	40.8082	4432
Papio anubis	Primates	GCF_000264685.3	Chromosome	2959.37	41.0402	71135
Piliocolobus tephrosceles	Primates	GCF_002776525.2	Scaffold	3038.01	40.9	47449
Pongo abelii	Primates	GCF_002880775.1	Chromosome	3065.05	40.8623	5300
Pygathrix nemaeus	Primates	GCA_004024825.1	Scaffold	3217.15	41	1431093
Rhinopithecus bieti	Primates	GCF_001698545.1	Scaffold	2977.07	41.5	105032
Rhinopithecus roxellana	Primates	GCF_000769185.1	Scaffold	2899.55	41	135512
Semnopithecus entellus	Primates	GCA_004025065.1	Scaffold	3033.04	41.3	1063749
Theropithecus gelada	Primates	GCF_003255815.1	Chromosome	2889.63	41.1126	15308
Alouatta palliata	Primates	GCA_004027835.1	Scaffold	3033.62	40.8	1136565
Aotus nancymaae	Primates	GCF_000952055.2	Scaffold	2861.68	41.1	28922
Ateles geoffroyi	Primates	GCA_004024785.1	Scaffold	2897.03	40.7	869075
Callithrix jacchus	Primates	GCF_000004665.1	Chromosome	2914.96	41.3414	16399
Cebus albifrons	Primates	GCA_004027755.1	Scaffold	3058.67	41.1	1539166
Cebus capucinus imitator	Primates	GCF_001604975.1	Scaffold	2717.7	41	7156
Pithecia pithecia	Primates	GCA_004026645.1	Scaffold	2933.56	40.6	892330
Plecturocebus donacophilus	Primates	GCA_004027715.1	Scaffold	2959.76	41	1033983
Saguinus imperator	Primates	GCA_004024885.1	Scaffold	3493.72	40.8	1666189
Saimiri boliviensis boliviensis	Primates	GCF_000235385.1	Scaffold	2608.59	41.1	2686
Cheirogaleus medius	Primates	GCA_004024725.1	Scaffold	2393.09	40.5	547751
Daubentonia madagascariensis	Primates	GCA_004027145.1	Scaffold	2498.42	40	342451
Eulemur flavifrons	Primates	GCA_001262665.1	Scaffold	2115.57	40.5	38367
Eulemur fulvus	Primates	GCA_004027275.1	Scaffold	2748.89	41.4	1693913
Eulemur macaco	Primates	GCA_001262655.1	Scaffold	2119.88	38.3	26772
Indri indri	Primates	GCA_004363605.1	Scaffold	2701.85	41.3	1243540
Lemur catta	Primates	GCA_004024665.1	Scaffold	2298.45	41.1	575427
Microcebus murinus	Primates	GCF_000165445.2	Chromosome	2487.41	41.3423	7678
Mirza coquereli	Primates	GCA_004024645.1	Scaffold	2340.87	40.8	440879
Nycticebus coucang	Primates	GCA_004027815.1	Scaffold	3294.56	40.7	1937143
Otolemur garnettii	Primates	GCF_000181295.1	Scaffold	2519.72	41.5	7793
Prolemur simus	Primates	GCA_003258685.1	Scaffold	2411.59	37.8	128596
Propithecus coquereli	Primates	GCF_000956105.1	Scaffold	2798.15	43.2	22539
Carlito syrichta	Primates	GCF_000164805.1	Scaffold	3453.86	41	337189
Loxodonta africana	Proboscidea	GCF_000001905.1	Scaffold	3196.74	40.9	2352
Acomys cahirinus	Rodentia	GCA_004027535.1	Scaffold	2306.07	42.7	371342

Allactaga bullata	Rodentia	GCA_004027895.1	Scaffold	3093.58	41.9	1284382
Apodemus speciosus	Rodentia	GCA_002335545.1	Scaffold	3494.29	24.2	336124
Apodemus sylvaticus	Rodentia	GCA_001305905.1	Scaffold	3758.14	22.6	559629
Capromys pilorides	Rodentia	GCA_004027915.1	Scaffold	3023.85	42	2352358
Castor canadensis	Rodentia	GCF_001984765.1	Scaffold	2518.31	39.7	21157
Cavia aperea	Rodentia	GCA_000688575.1	Scaffold	2716.4	42.2	3131
Cavia porcellus	Rodentia	GCF_000151735.1	Scaffold	2723.22	40.1	3144
Cavia tschudii	Rodentia	GCA_004027695.1	Scaffold	3077.35	40.1	1368179
Chinchilla lanigera	Rodentia	GCF_000276665.1	Scaffold	2390.87	41.4	2839
Cricetomys gambianus	Rodentia	GCA_004027575.1	Scaffold	2397.72	42.1	1102952
Cricetulus griseus	Rodentia	GCF_000223135.1	Scaffold	2399.79	41.6	109152
Ctenodactylus gundi	Rodentia	GCA_004027205.1	Scaffold	2322.47	42.7	860073
Ctenomys sociabilis	Rodentia	GCA_004027165.1	Scaffold	2834.81	40.7	807571
Cuniculus paca	Rodentia	GCA_004365215.1	Scaffold	3399.58	40.8	3768450
Dasyprocta punctata	Rodentia	GCA_004363535.1	Scaffold	3012.1	40.8	986664
Dinomys branickii	Rodentia	GCA_004027595.1	Scaffold	2688.56	41.8	1158438
Dipodomys ordii	Rodentia	GCF_000151885.1	Scaffold	2236.37	42.6	65193
Dipodomys stephensi	Rodentia	GCA_004024685.1	Scaffold	2346.42	41.9	1269643
Dolichotis patagonum	Rodentia	GCA_004027295.1	Scaffold	3289.61	39.8	1905764
Ellobius lutescens	Rodentia	GCA_001685075.1	Scaffold	2353.19	39.9	58423
Ellobius talpinus	Rodentia	GCA_001685095.1	Scaffold	2265.97	41.3	350460
Fukomys damarensis	Rodentia	GCF_000743615.1	Scaffold	2333.89	40.5	74730
Glis glis	Rodentia	GCA_004027185.1	Scaffold	2462.09	42.3	433430
Grammomys surdaster	Rodentia	GCF_004785775.1	Scaffold	2412.66	41.3	23656
Graphiurus murinus	Rodentia	GCA_004027655.1	Scaffold	2815.01	41.5	1267939
Heterocephalus glaber	Rodentia	GCF_000247695.1	Scaffold	2618.2	41.2	4229
Hydrochoerus hydrochaeris	Rodentia	GCA_004027455.1	Scaffold	2734.59	39.8	655342
Hystrix cristata	Rodentia	GCA_004026905.1	Scaffold	2470.92	41.5	1136598
Ictidomys tridecemlineatus	Rodentia	GCF_000236235.1	Scaffold	2478.39	40.5	12483
Jaculus jaculus	Rodentia	GCF_000280705.1	Scaffold	2835.25	42.7	10898
Marmota flaviventris	Rodentia	GCF_003676075.1	Scaffold	2582.14	40.1	32915
Marmota himalayana	Rodentia	GCA_005280165.1	Scaffold	2471.7	38	4363
Marmota marmota marmota	Rodentia	GCF_001458135.1	Scaffold	2510.59	40.2	14543
Meriones unguiculatus	Rodentia	GCF_002204375.1	Scaffold	2523.12	42.5	68793
Mesocricetus auratus	Rodentia	GCF_000349665.1	Scaffold	2504.93	43.2	21484

Microtus agrestis	Rodentia	GCA_001305995.1	Scaffold	3124.14	26.4	230202
Microtus ochrogaster	Rodentia	GCF_000317375.1	Chromosome	2287.34	42.8312	6450
Mus caroli	Rodentia	GCF_900094665.1	Chromosome	2553.13	42.6193	3163
Mus musculus	Rodentia	GCF_000001635.26	Chromosome	2818.97	41.95	336
Mus pahari	Rodentia	GCF_900095145.1	Chromosome	2475.01	42.6694	2581
Mus spicilegus	Rodentia	GCA_003336285.1	Scaffold	2496.08	-	37536
Mus spretus	Rodentia	GCA_001624865.1	Chromosome	2625.59	42.3292	5404
Muscardinus avellanarius	Rodentia	GCA_004027005.1	Scaffold	2527.15	41.6	466805
Myocastor coypus	Rodentia	GCA_004027025.1	Scaffold	2903.18	41.7	1669667
Myodes glareolus	Rodentia	GCA_004368595.1	Scaffold	2396.23	39.8	50926
Nannospalax galili	Rodentia	GCF_000622305.1	Scaffold	3061.42	41.6	154976
Neotoma lepida	Rodentia	GCA_001675575.1	Scaffold	2353.86	36.2	119373
Octodon degus	Rodentia	GCF_000260255.1	Scaffold	2995.89	42.5	7135
Octomys mimax	Rodentia	GCA_002564305.1	Scaffold	2591.29	41.2	1257804
Ondatra zibethicus	Rodentia	GCA_004026605.1	Scaffold	2562.75	41.6	565356
Onychomys torridus	Rodentia	GCA_004026725.1	Scaffold	3202.22	41.6	2261735
Perognathus longimembris pacificus	Rodentia	GCA_004363475.1	Scaffold	2601.7	41.8	2409818
Peromyscus leucopus	Rodentia	GCF_004664715.1	Scaffold	2474.07	42.2	1763
Peromyscus maniculatus bairdii	Rodentia	GCF_000500345.1	Scaffold	2630.54	42.7	30921
Peromyscus polionotus subgriseus	Rodentia	GCA_003704135.2	Chromosome	2645.11	42.9297	5292
Petromus typicus	Rodentia	GCA_004026965.1	Scaffold	2389.37	41.8	910679
Phodopus sungorus	Rodentia	GCA_001707965.1	Contig	2003.23	41.3	1462128
Psammomys obesus	Rodentia	GCA_002215935.1	Contig	2311.47	42	96427
Rattus norvegicus	Rodentia	GCF_000001895.5	Chromosome	2870.18	42.3282	1395
Rhizomys pruinosus	Rodentia	GCA_004026225.1	Scaffold	3937.27	41.7	4150536
Sigmodon hispidus	Rodentia	GCA_004025045.1	Scaffold	2730.6	39.8	883546
Spermophilus dauricus	Rodentia	GCA_002406435.1	Scaffold	3106.27	40.2	360804
Thryonomys swinderianus	Rodentia	GCA_004025085.1	Scaffold	2662.15	42.8	1889641
Tympanoctomys barrerae	Rodentia	GCA_002564285.1	Scaffold	2651.63	41.2	1558800
Urocitellus parryii	Rodentia	GCF_003426925.1	Scaffold	2520.51	41	15884
Xerus inauris	Rodentia	GCA_004024805.1	Scaffold	2601.42	40	577370
Zapus hudsonius	Rodentia	GCA_004024765.1	Scaffold	2611.19	42.3	1158100
Aplodontia rufa	Rodentia	GCA_004027875.1	Scaffold	3005.54	40.6	1279526
Tupaia belangeri	Scandentia	GCA_000181375.1	Contig	2137.23	41.4	841006
Tupaia chinensis	Scandentia	GCF_000334495.1	Scaffold	2846.58	42	50750

Tupaia tana	Scandentia	GCA_004365275.1	Scaffold	3648.91	42	3883620
Trichechus manatus latirostris	Sirenia	GCF_000243295.1	Scaffold	3103.81	41.6	6323
Orycteropus afer afer	Tubulidentata	GCF_000298275.1	Scaffold	4444.08	42.0995	22509

*indicates the presence of endogenous lokiretroviruses in corresponding species.

Table S3. Information of queries used in this study

Name	Abbreviation	Accession No.
Friend murine leukemia virus	MLV	NP_040333.1
Walleye dermal sarcoma virus	WDSV	NP_045937.2
Mouse mammary tumor virus	MMTV	NP_056880.1
Human endogenous retroviruses type L	HERV-L	Ref. 1
Eastern chimpanzee simian foamy virus	SFVcpz	YP_009508551.1

References

1. L. Vargiu, P. Rodriguez-Tomé, G. O. Sperber, M. Cadeddu, N. Grandi, V. Blikstad, E. Tramontano, J. Blomberg, Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* **13**, 7 (2016).

Table S4. Information of sequences used for phylogenetic analyses

Type/Group	Name	Accession No.
Bacterium	*Mycolicibacterium gilvum* RNase H	WP_011893771.1
Bacterium	Mycobacteriaceae RNase H	WP_011560725.1
Bacterium	*Micromonospora aurantiaca* RNase H	WP_013287786.1
Bacterium	*Buchnera aphidicola* RNase H	WP_011091358.1
Bacterium	*Shewanella denitrificans* RNase H	WP_011496458.1
Bacterium	*Helicobacter pylori* RNase H	WP_001155465.1
Belpaoviridae	Kobel	Gypsy Database 2.0
Belpaoviridae	Tamy	Gypsy Database 2.0
Belpaoviridae	Zebel	Gypsy Database 2.0
Belpaoviridae	Bel	Gypsy Database 2.0
Belpaoviridae	Purbel	Gypsy Database 2.0
Belpaoviridae	Cer10-1	Gypsy Database 2.0
Belpaoviridae	ApBIN1h	Gypsy Database 2.0
Belpaoviridae	ApBIN1j	Gypsy Database 2.0
Caulimovirus	CaMV	Gypsy Database 2.0
Caulimovirus	SbCMV	Gypsy Database 2.0
Caulimovirus	ComYMV	Gypsy Database 2.0
Caulimovirus	RTBV	Gypsy Database 2.0
Caulimovirus	CSVMV	Gypsy Database 2.0
Caulimovirus	PVCV	Gypsy Database 2.0
Cgin1	HsCGIN1	Gypsy Database 2.0
Cgin1	MmCGIN1	Gypsy Database 2.0
DIRS	*Caenorhabditis briggsae* AF16 DIRS	CAP25970.2
DIRS	*Panagrellus redivivus* DIRS	CAA43185.1
Eukaryote	*Arabidopsis thaliana* RNase H	AAF24611.1
Eukaryote	*Oryza sativa* RNase H	AAK54296.1
Eukaryote	*Ostreococcus tauri* RNase H	XP_003080907.1
Eukaryote	*Ignicoccus hospitalis* RNase H	WP_012123294.1
Eukaryote	*Vulcanisaeta moutnovskia* RNase H	WP_013604220.1
Eukaryote	*Ferroglobus placidus* RNase H	WP_012965086.1
Eukaryote	*Pisum sativum* RNase H	AAQ82037
Eukaryote	*Haloarcula marismortui* RNase H	AAV48010.1
Eukaryote	*Natrinema pellirubrum* RNase H	WP_006182114.1
Eukaryote	*Halorhabdus utahensis* RNase H	WP_015788958.1
Eukaryote	*Halorubrum lacusprofundi* RNase H	WP_012659552.1
Eukaryote	*Candida orthopsilosis* RNase H	XP_003871485.1
Eukaryote	*Komagataella phaffii* GS115 RNase H	XP_002490024.1
Eukaryote	*Trypanosoma brucei* RNase H	AAC47537.1
Eukaryote	*Homo sapiens* RNase H	AAV38476.1
Eukaryote	*Caenorhabditis elegans* RNase H	NP_001040786.1
Gin1	HsGIN1	Gypsy Database 2.0
Gin2	DrGIN2-2	Gypsy Database 2.0
Gina	HmGINA1	Gypsy Database 2.0
Ginger1	HmGINGER1	Gypsy Database 2.0
Ginger2	HmGINGER2-1	Gypsy Database 2.0
Ginger2	FcSCAND3	Gypsy Database 2.0
Ginger2	HsSCAND3	Gypsy Database 2.0
Ginger2	MeSCAND3	Gypsy Database 2.0
Ginger2	HsKRBA2	Gypsy Database 2.0
Ginger2	BtKRBA2	Gypsy Database 2.0
Ginger2	LaKRBA2	Gypsy Database 2.0
Ginger2	ApGINGER2a	Gypsy Database 2.0

Ginger2	BmGINGER2a	Gypsy Database 2.0
Ginny	HmGINNY1	Gypsy Database 2.0
Gino	HmGINO1	Gypsy Database 2.0
IS3/IS481	ISSc	Gypsy Database 2.0
IS3/IS482	ISAf	Gypsy Database 2.0
IS3/IS483	ISLa	Gypsy Database 2.0
IS3/IS484	ISTv1	Gypsy Database 2.0
IS3/IS485	ISEc	Gypsy Database 2.0
Maverick/Polinton	DpCIN1	Gypsy Database 2.0
Maverick/Polinton	DrCIN1	Gypsy Database 2.0
Maverick/Polinton	SmCIN1	Gypsy Database 2.0
Metaviridae	Maggy	Gypsy Database 2.0
Metaviridae	Pyggy	Gypsy Database 2.0
Metaviridae	Skippy	Gypsy Database 2.0
Metaviridae	TF1	Gypsy Database 2.0
Metaviridae	Amn-san	Gypsy Database 2.0
Metaviridae	Sushi-ichi	Gypsy Database 2.0
Metaviridae	Del	Gypsy Database 2.0
Metaviridae	Galadriel	Gypsy Database 2.0
Metaviridae	CRM	Gypsy Database 2.0
Metaviridae	Reina	Gypsy Database 2.0
Metaviridae	Mdg3	Gypsy Database 2.0
Metaviridae	Kabuki	Gypsy Database 2.0
Metaviridae	RetroSor1	Gypsy Database 2.0
Metaviridae	Tat4-1	Gypsy Database 2.0
Metaviridae	Athila4-1	Gypsy Database 2.0
Metaviridae	Diaspora	Gypsy Database 2.0
Metaviridae	Cer4	Gypsy Database 2.0
Metaviridae	SPM	Gypsy Database 2.0
Metaviridae	Mag	Gypsy Database 2.0
Metaviridae	Tor1	Gypsy Database 2.0
Metaviridae	Tor2	Gypsy Database 2.0
Metaviridae	SURL	Gypsy Database 2.0
Metaviridae	Cer3	Gypsy Database 2.0
Metaviridae	Gmr1	Gypsy Database 2.0
Metaviridae	Woot	Gypsy Database 2.0
Metaviridae	Mdg1	Gypsy Database 2.0
Metaviridae	Tor4a	Gypsy Database 2.0
Metaviridae	Cigr-1	Gypsy Database 2.0
Metaviridae	Tom	Gypsy Database 2.0
Metaviridae	Zam	Gypsy Database 2.0
Metaviridae	Gypsy	Gypsy Database 2.0
Metaviridae	ApGIN7a	Gypsy Database 2.0
Metaviridae	ApGIN4c	Gypsy Database 2.0
Metaviridae	ApGIN5c	Gypsy Database 2.0
Metaviridae	ApGIN3c	Gypsy Database 2.0
Non-LTR	TRAS1	BAA07467.1
Non-LTR	CATS	BAE44464.1
Non-LTR	You	CAC16871.1
Non-LTR	I	AAA70222.2
Non-LTR	Lian-Aa1	AAB65093.1
Non-LTR	CgT1	AAA85636.1
Non-LTR	MGR583	AAB71689.1
Non-LTR	Mars1	CAA67543.1
Pseudoviridae	1731	Gypsy Database 2.0
Pseudoviridae	pCretro6	Gypsy Database 2.0

Pseudoviridae	SIRE1-4	Gypsy Database 2.0
Pseudoviridae	Vitico1-1	Gypsy Database 2.0
Pseudoviridae	Hydra1-2	Gypsy Database 2.0
Pseudoviridae	Copia	Gypsy Database 2.0
Pseudoviridae	Tnt-1	Gypsy Database 2.0
Pseudoviridae	Koala	Gypsy Database 2.0
Pseudoviridae	Ty1B	Gypsy Database 2.0
Pseudoviridae	Ty4	Gypsy Database 2.0
Pseudoviridae	CoDi6.3	Gypsy Database 2.0
Pseudoviridae	Zeco1	Gypsy Database 2.0
Pseudoviridae	CoDi6.7	Gypsy Database 2.0
Pseudoviridae	CoDi5.6	Gypsy Database 2.0
Retrovirus	SRV-1	M11841.1
Retrovirus	JSRV	NC_001494.1
Retrovirus	MMTV	NC_001503.1
Retrovirus	LDV	U09568.1
Retrovirus	RSV	NC_001407.1
Retrovirus	HIV-1	NC_001802.1
Retrovirus	FIV	NC_001482.1
Retrovirus	EIAV	NC_001450.1
Retrovirus	HTLV-2	NC_001488.1
Retrovirus	BLV	NC_001414.1
Retrovirus	MuLV	NC_001362.1
Retrovirus	FLV	NC_001940.1
Retrovirus	KoRV	AF151794.2
Retrovirus	REV	NC_006934.1
Retrovirus	WDSV	NC_001867.1
Retrovirus	Xen-1	AJ506107.1
Retrovirus	MuERV-L	Y12713.1
Retrovirus	HERV-L	Ref. 1
Retrovirus	SnRV	NC_001724.1
Retrovirus	BFV	NC_001831.1
Retrovirus	EFV	NC_002201.1
Retrovirus	FFV	NC_039242.1
Retrovirus	SFV	EU010385.1
Retrovirus	CoEFV	Ref. 2
Retrovirus	AciFLERV	Ref. 3
Retrovirus	AliFLERV	Ref. 3
Retrovirus	NviFLERV-1	Ref. 3
Retrovirus	HreFLERV	Ref. 3
Retrovirus	CmiFLERV	Ref. 3
Retrovirus	CmiFLERV	Ref. 3
Retrovirus	XtrFLERV	Ref. 3
Retrovirus	ZFERV-2	CAAK05053864.1

References

1. Vargiu L, Rodriguez-Tomé P, Sperber GO, Cadeddu M, Grandi N, Blikstad V, Tramontano E, Blomberg J. 2016. Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* **13:**7.
2. Han GZ, Worobey M. 2012. An endogenous foamy-like viral element in the coelacanth genome. *PLoS Pathog.* **8**:e1002790.
3. Aiewsakun P, Katzourakis A. 2017. Marine origin of retroviruses in the early Palaeozoic Era. *Nat Commun.* 8:13954.

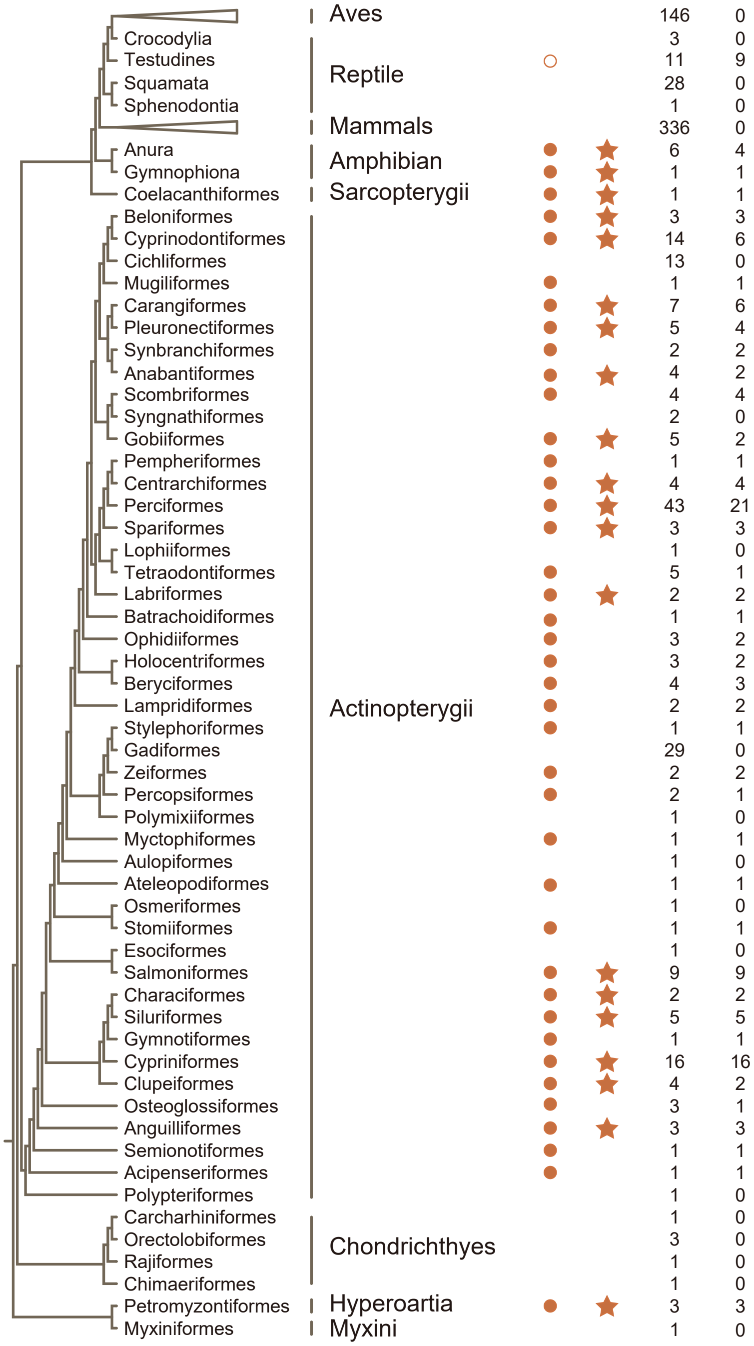

Fig. S1. Distribution of endogenous lokiretrovirus elements. Filled orange circle indicates the presence of endogenous lokiretrovirus elements in the corresponding order. Empty orange circle indicates endogenous lokiretroviruses in Testudines were highly degraded. Filled orange star indicates that at least one lokiretrovirus consensus sequence has been reconstructed in the corresponding order. The first column of numbers represents the number of genomes used to screen lokiretrovirus. The second column of numbers represent the number of genomes where lokiretroviruses were identified.

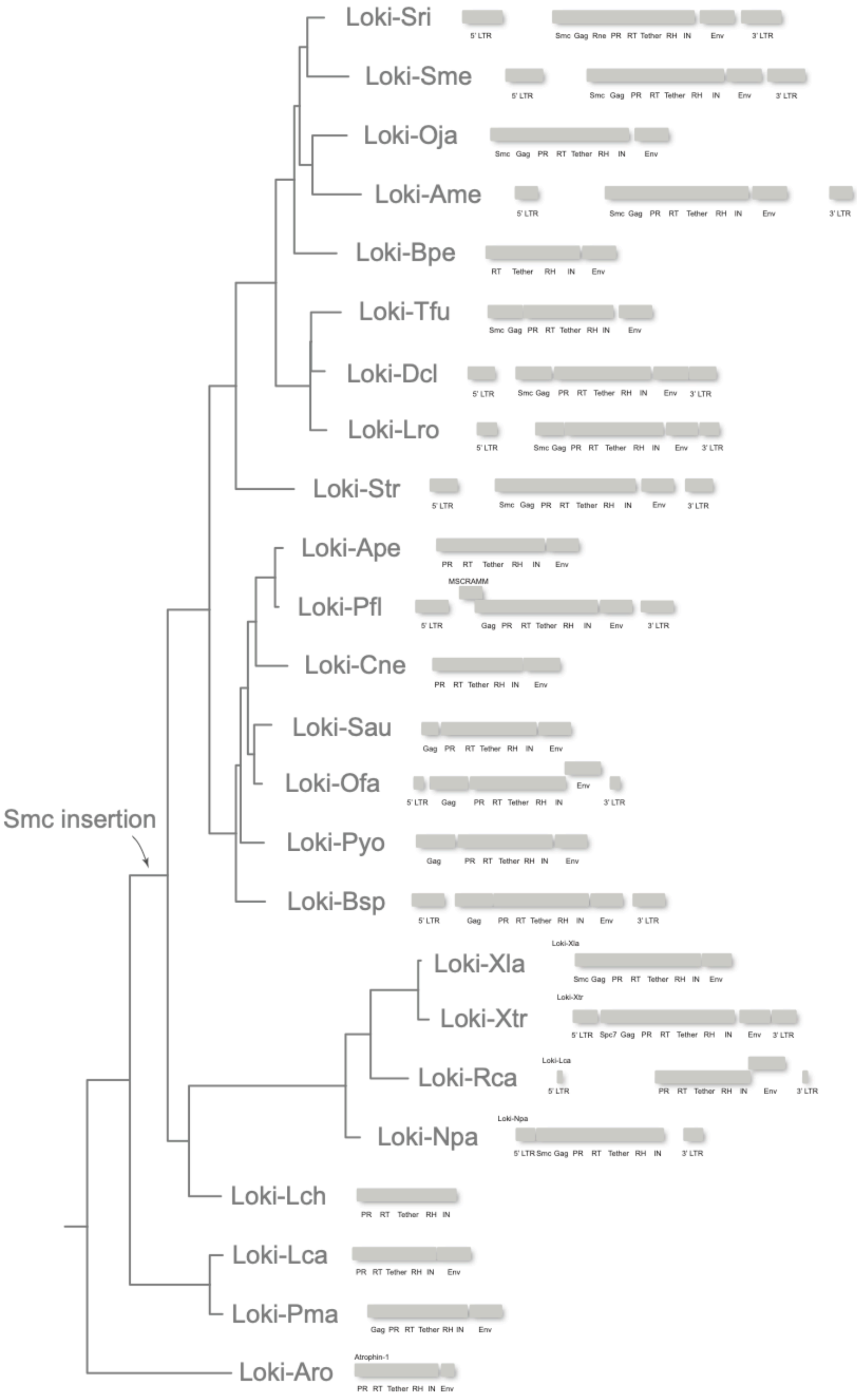

Fig. S2. Genome structures of 24 lokiretrovirus consensus sequences.

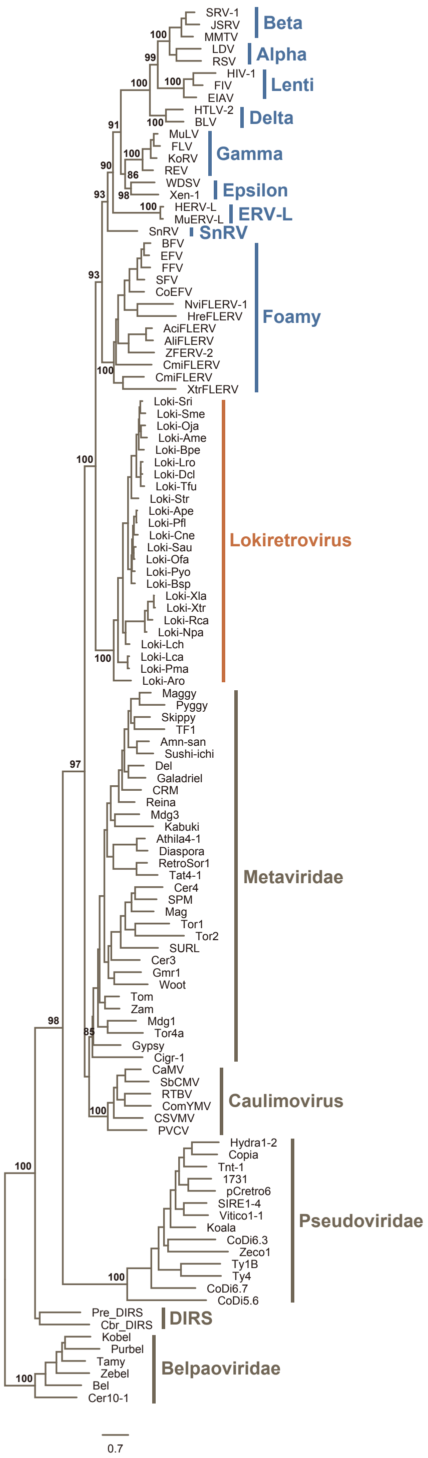

Fig. S3. Phylogenetic relationship of **lokiretroviruses, retroviruses, DIRS, and other orterviruses.** The supports for selected nodes were evaluated using ultrafast bootstrap approximation method with 1,000 replicates.

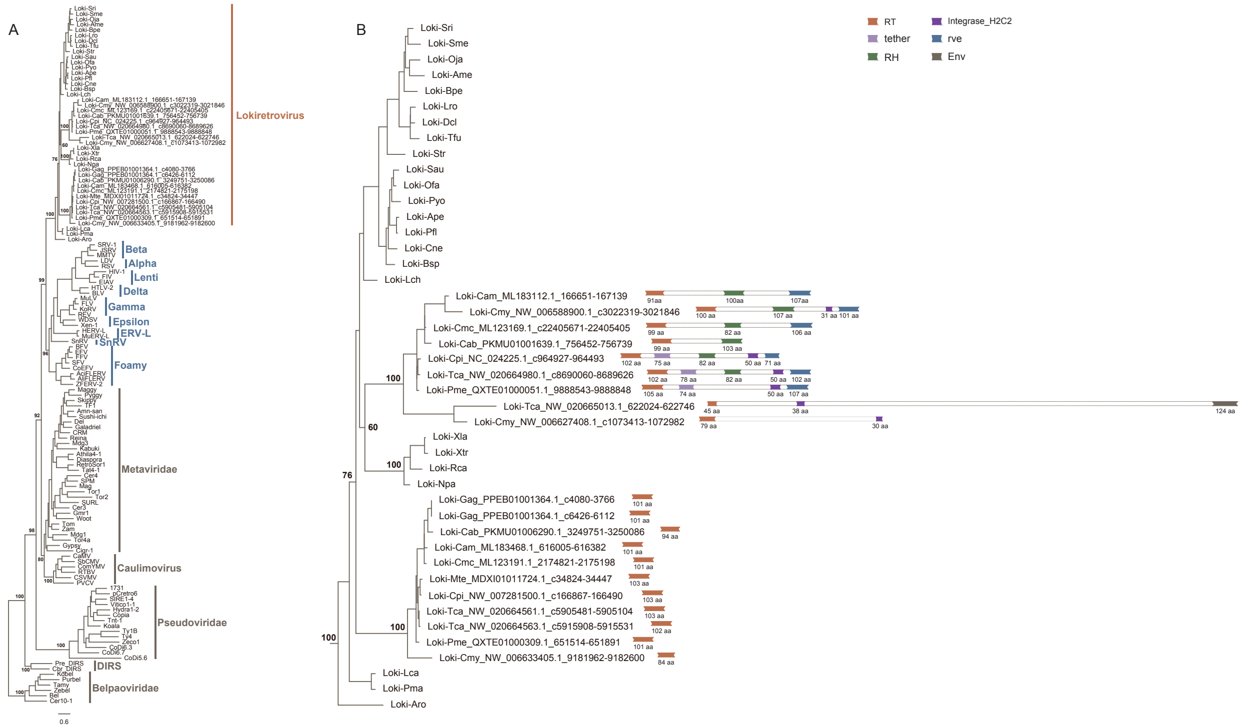

Fig. S4. Lokiretroviruses in reptiles. (**A**) Phyogenetic relationship among reptile lokiretroviruses and other reverse-transcribing viruses. (**B**) The length of domains within reptile lokiretroviruses. All the domains are very short in length, indicating reptile lokiretroviruses are highly degraded.

Dataset S1. Reconstructed consensus sequences of lokiretroviruses

```
>Loki-Bsp consensus sequence
TCAGTGAAGTGTTGGGGGACCTCATATGTATAATCTGTTATTAACAGAATTATTATT
AAGGTCCCACATGCTCAGACAAAGGAAATACTGCGGCCTGCATGTAAAGTCCTCAT
CCTGATAGTAGTGAATGAGAACAAAGAGATTTCAAATCTCGGCGCCAAAACCAGGC
GCCAGGCTGGGGAACCACACCTGGCTCCCCGCTGTGAAGGCCGAAACCCCGCTGAC
ACAAGCGACAAACGGAGGAGAAACTTGGACGTCAGCCATCCTGATTGGACCAAACA
ATAGGATGGGCGTGACCAAAAACACGGTTATAAAAGTCGCCGTGACCCAAGACTCG
TCCCTTGTCCTCTTCGATCACCCGTTACCCCTTGTTCTCCCCCTTGTTCTCTTCGATCT
CCCGTTTCCCCTTCCTCTCTTCGACCCTCTCGCTGCCCCACCTGGAGCTTCCCTCGGC
ACCAGCACGCCTGAAGCCTCCACCGCAGCCAGCAGCTGAACCAACGGAGACGTCGT
CACAGAGAACGGGCCTGATCACCCATCTACTTCCAGCCATACGGCAGGAAACCGCC
AGCTTAACCGCCTCACTCGCCGCCGCATCAGCTGATCACGACGCGATCACCGCGACG
CACACCTGGACCGGACCGGACCGCAACAGAGCAAGCACGCACGCACGCACGCACG
CACGCAACGCCGGAGCTTCTCACCAGGATTGAGCAGTAAGGCAGAGGCATTCTTAA
GTCTGAGCAGAGTAGAACTCTTAGGGTATCAGTAATTGTTTTATTGTTGTGTTTTCTT
TTCCTGCACACACGATCTGACCGCTGTGTATTTCCGCGATCACGTGACTGTAACGCT
GTTACAGATCTACGTGAATCTAAGAGGTGCACATTGAAGTGTTGTAACCTGTGATTT
CGCTAGCTTAAGTGCTTTATTGTGTTACTGTGTAGTGCTCTGTTTCTGCTGCGACCAC
GCCAAGCGCACGTTTCAACCACTGCACGCTTCCGCGATCACGTGACTATATCGCAGA
ATATAGCCCCACGTGTTTGTGTGACCACGAGCCTTATCCACGCTAGTCCCACTTCAC
ACTCCTTTACACGCTTCTTTCCCTGTGAGCGCCAGACACGCGCGCTCCATTCAAAGG
GCTCACCGGCACACGCGCGCAGCTGTGCCACCTCACCACGCCCCTAAAGGGGGGCG
CTCTCCTCCTCCTCCTCTTCTCTCTTTCTCTCTCTCTCTCACCCTCACACACATATA
CACACACACACACACACACACACACACCCCTCTTATCTCAGGTAACGCGCATACAC
ATACAGAACGCTAGACGAACACACCATATAATTAAGTTGTGATTTGCTGAGTTATTC
AAGATAGTGTTGACACTAGAGTTAATATTAACTGTTTAATAAAGCGTTCATAATTTA
ACTAGTCGTTGTTTGTGATTATTTGTATATGGTTTTGCAGCACTAGTGAATTGAGTCG
GCTATGGTACGGAATTCACCCTTCAACTAACTTAAATACAAATGAGACTGTATTGGC
TATCCCAAATCGGTTATTGGTCCCTGGAAACAGGGTGGTGCCCCGTAAATTAAGTAA
TTATTAATTCAATTAATAATTATTAAATCAATAATAGTAAGTGTTTCACTTGCCAAAC
TATAGACATTTGTCATATAGTTATATATATTTAAGCTAATAACCAAGATTTGTGAATC
TAAATTATAAACATAATTGGTATCTCCACTCAGGAGCTATAAATCCAATTATTATAA
TTTGTGCTCCTCGCATATCCCCAACATAATTGGTGCCCCGTGTGAGGCCGACAAAGA
TTATTCACCGTGCTGTAAACCTGTACAAATAATCTACTGGTGGATGCAAGAATCCAC
AACTCTGTGTGTAGAGACCAGGTTGCCTATTGGTAAAGCTAGTCGCCGAAAGGAACT
TATTGAGTTTGGAAGCTGTGTTTCACCTAGTGTTGATGTGCTAAGTGTGGGTGTTGCA
TAAGTTACTGAAGCACCAGCATCACTAGCATTGCTCTGCTACCATCTTAGCTACCAG
CATTAGCCAACAGCTAATATTAACTTAGATGCCAGCGTAGCATTAGCAACAGTTGCT
ATTAGCTTAGCAGCCAGCTTAAGCGGTAGCCTAGCGATAAGCTACAGTGGGTCAGC
AGTAGTTGTCACCCTAGCCCAGCCACTGTGGTATTAGCTTAGCCTCAGCCATTGAAT
AATACCAGTGCATATATATACACATATTGTTTTGCTCAGTGAGCATTCAAGTAAGTC
GCAAGTAAACACTACTTGTGCGATACTATTCTAGCTGTTGTGAAAACCTTACAAATA
AGGATGAGTAACAAGGGCTCCCCTGGGTCAGGAAGTCTGGACTTCCAAGGTGTTGTT
GCAGGACTTGTCAGAATATCCACGAAAGAGCTAGAGCGCCTTAACAAGTATACTGT
```

CGCAGAGTGTGAGGGAGAAATGCAGAGGCTGATGTCGGTCGTGAGTCAACCCACGG
GAGAAACCCGCCTGACTCAAGAGGTAGTGGGCCAGCTGGCATTGCTATGCCATCAG
AAAGCCCAACTGGAGTCCGCAGAATCTGCTGAGAAGATCGCGGCCCTACACAGTAG
CTACTTAGCTGAGCAAAAGACCGTCACCCATCTAGAGCATAGCTTAGAGGTAATGA
AAAGGGAACTGGAACAACTCGAGGAGAGCTATAAGAGGTTGCAGCGAGACCATGA
CATACCCCCTAGGTCCCGAACAGTTCACCAGGAGGAGGTAGGAGATAGGAATGAAA
ATCAGTGCAGACAAATAGGTATGGCGGCTCAGTCAAGTGACCAAGAGCTGGACAAA
CCTGGTCCTGGGAGTGAGACTTTCCACAGCCTAGAGGCCCCAAAGTCAGTCCCCACA
AACCGGACCAATCTAGCACAAAGAGAGCCCCGTGGGTCCTTTGGAGAAGCGACCCC
AGCTCACGCCCACAAGGGCACTCATTCAAACTCTTACCTCACTACTGACACCCCAGT
CTACTACGCATCGACCATAGCCCATGGTAGCGTAGTCAGGGGAACAGGAAGACTGA
ACAGGGACAGGTATGCACCCATGGATACGCCAATCCACAGAAATGTACTCCCGAGT
TACCCCCAGCCAGAGCCACACCTCGAAGACCAGAGATGGCAAGGTCCTAGTAGAGA
AACCTACTTTAAGCCTGAGCCCACTAATGGTAGATATGAGCCACTGCCACCCAGTGG
GGGATACCTTCCTCGCAACGCCCCCCCCCTACGCCCTCACCCAGTCCCTTACTACGG
GGGCGACTCAAGCTCAGAGGACTCTGACATTGTGAACCCTGTCCCAAGGCGGGAGA
TACGCACACGCCAACTCGAATCTATGGCTAAGGACATTGAGCGCTTTGATCCCAGTG
ACAGGGAGTCCACCATAGATGACTACCTGCGGGAAGTTGAGCGGTGCCTCCTGGAC
TTGCCTGAAGCAACAGCTCGCGAAAAGCTGAAGCTGATCTGGAAGACCACGGCTAG
AAGTGTGCATAGTTTTATAGAAACCCTTCCCTGCAGCATACGTGACAGTTACCCAAG
GCTCTGTGATGCCCTTAGGGAAGAGTTCTCGCCCTTTACTGACTCTGCCGCTGCCAC
CCTTGGAGCATTTGAGGTCCTGCACAAAAGGGACGAATCTCCAAAGGAATATTATA
GGCGCCTCCGAGCAGCCTACTTCCAGGGTCACAATGCCCCAGGTCTAGAGGAAGAC
AGCGCATTCCGATCCCTCTTCCTGCATAACCTCCACGAGAGCATTCGGTATGAGGTG
ACAATGCATAGCCAATCAGGTCGACTCACTATGCAGGAACTCCGAAAATACGCGCA
AATGGCCTGGCAAACCCGCGCACGACCAAGGCGTCGCCCTGACCGGGACGCCAGAG
TCCTGGGCATCCAAACTACGGACGACTTGGCGTTAGAAGGCCCAGAATTACAGCGG
GGAAGGGGAGAAAATAGAGTACAACCCCATGCCACAGAGTTTTCCCATAGGCCCGG
GAGTCGACCCAGACAAGGTGGGGAACACACCCACCGGGCCTCAGAGAACAACCAA
CAGTATGACACTGCACCACAGAGATATAAGGGACGGAGTGAACAAAAGTCAGCTCC
GAAAGGGAAGTACAGTAGAAAGGCTGATAACCATCTCAGTGCAGAACTGAGAGCA
GCCATTGCGCAAGGCATAGCTGATGCTCTGGCCACGCAAAATGCACCTGGGGGTTC
CGCCCCGCCTCAGAAATCCCCAAAAAAGCAAATAAAAAGTGAACCGTCTAAGTGAC
TAAGCCGGGACCCGCTCCCAGAAGCCGCAAGGGTTTACCTCGTAGGCTGGGGGAGA
GAAGCTGGCGTGCCCCAAGTAGACACTGCTCGTACACCAGACAACAACACTCCCCC
CTTCTTAACCTTCTTGGGAGACCTTGTCCGACAGGGTATGCCCGCCGCATTTACACGT
CAGTGGTGATAGGGGGATGCTTCAGCTCCCTGGCCTTACTGGACTCTGGTTCTGAAA
TCAGCCTAATGTCTTCAGATACTTTCAAGGCAATGTCCAGAACCATGTTGTCTCTTGG
GAAACCCCTACAGGTCGAACCGTGTAACATTAACATCACTAGTTACACACAAGACA
AGTCCCCCATTGTTGAACGAGCCTGGGTTACTATTGCTTTCCAAGAGGTAACTCTTGT
GCACCCTTTGTATATCTGCAATCTTGACACAGAGCCAGTGCTTATTGGCCAAGATCT
GTTGGGCCGAGTGGCACCTCTTATTGACTGTCACAGAGGTCGCATCTGGGCCCAGGT
GAAAGTCCCCAAACCATTGGATCTTGGGGACAACCGACCAGGATGTGACGGACAGG
TCACTGCAGTGGAGAGGCAAGGCCAAGAGTCCAACCCCAGCTCTGAGGTGCCAGAC
TTGCCAAACTCCTTCCAGACCCATGGTTCATTTCTCTGCGCTCTGACAAATGAGGGC
CAAATCATCTACAGCCCAAAGATCGTGGGGGGCATTCATCTTGACAACCTTCCTGTG

CCTGAGGCGCAACTGACACTCCGGTCAGAAAAGACCGCCATCAGCCAGGAATTGTA
TGACGCCTTGTGCTTGCAACGACCACCTCCCACCTTCGTGCAGCATAGTCATAGGCT
CCTCCCAGCAGATCAGTCAAGACCCCTCCTTAAGGCAATCGGGGTCTGCGCCGCCAC
CATTCGAATTGGGGGTCGGGTCATCAAACATTTCCTAAGTGTAGTTCCCCACTTGCA
GCCGCCTGCTCTTGTGGGCGCCGATCTCCTTGTGCGGCTAGGTACGCACCTTGACAC
CATCAACCAAGTCTTATGGTCGCAGGTCAGCCCAAGAGACACCTCACAAAGTCCTG
AAATAGAGAACATGAAATCGGGACAGACCATTCCCCACGCGTGCCAAGTTGCTGCG
GAGAGCGACAAGACCATTCCCCCCAGAACGGCAGGTAACCCGGTCCGTCTAGTAGT
CATGAAGGGGCAGAAAATCCATGGCTCACAAGCATTCTTCCAACCCCTTCCTAGTTT
CGCCGAACTTAACCTGACGGTCTGTGGCACCCCTCTCCTAGAACTGGACAATCGTGC
AACCTACCTCTTGGTTCAGAATCCAACCAATGCACCGATCCTTGTTAAGGCCCAACA
ACCGCTGGGCGTCTTGGTAGAAAGTTCGTTTCATGACTTTGAGCTGACGGTACCCGT
CATTGGGGAATTACCAGCGTCTCTAATCAGGGAGGGAAGTGTGGGAGACGTCCTCC
TGTCTCACCCAACGGGGATGATCTCTGTATTGCGCTATGAGCCACTGAGAGACCAAC
CAATCTGTGGGGCAACACTTGAACCAGGGGGTGACATGCTGGTCTATACCTTGGCTG
CCCAGCCAGAGTGCACTGCCCAAGCGGGAACCAACACCTCTCCGGCGGGAGAGGAA
CCTTACCCCGGCTTTGAAGCAGAAATCAACCAGCAACTTGAAAAAGCGGACGCATT
GACAAATGAGGGGCAAAAAGAAGCCCTTCGAGATCTCTTCCGCGAATTCAAACCCA
TCTTCTCACGAGATTCGCATGACTGTGGGGTCACTGACTTACACACCGTGCGCATTC
CCACTGACCCCCATGCACCACCAACGTTTGTGCGACAGTATAGGATTCCTCTGGCAG
CATATGAGTCAATCCAAGAGATTCTGGACACACTGCTGAAGAAGGAGATCATAAGG
GAATGTAACTCGACATATAACTCGCCGATCTGGCCGGTTCTGAAACCGACAGGGAA
ATGGCGTCTTACCATCGATTACCGACCCTTGAACAAAAACGTTCCCTTGTCTCGATG
GCCCATGATTCACTTAGACCAAGAACTGGCTAAGATAAAAGACGCCCATTTCTTCTC
CACTGTAGATGTTGCCAATGGCTTTTGGACCATGAGAGTGGATCCAGCTGACCAGTA
TAAACTGGCCTTTTCCTTCGGCAACCGCCAGTACACCTGGAACCGTTGCCCCTTCGG
GTACTCAAACTCCCCTGCAGAGTTCAATATCTTCCTCCATAAAGCAATGAGTGACGC
TGCCGCTCGCGGTAACCTCATCTACGTAGACGATATCCTCATGAGGAGTCAGACATT
TGAAGGACACTTGGCAGAAATCCGACATGTGCTCTCACAACTCCTGGCTGCAGGAG
CCAAGTTAGCATTAACCAAGGGTCAGTGGTGTCGAACCAAGGTGAAGTATGTCGGG
CTGACAGTGGGCCAACATGGTATTGAGCCTCAGGCAGACAGAATCCAGGCTATTAG
AACCATTAAAGCCCCAAACACACTGTCAGGACTCAGGAGCTTCCTAGGGGTCTGTA
ACTACTCACGCCAGTTTGTAGAAGACTATGCGGAGATAGCCAGACCCCTCACAGAA
CTCCTTCGTAATGATACTCCCTATATATGGACGGAGCGCCAACAACAGTCATTTGAG
AAAATGAAAGAAGCGCTCAGCACAGCCCCGTGCCTAGCTTATCCTGACAAGGATAA
AGAATTCTATCTGGAAGCTAGCTTCTCCCAACAAAGCCTCAGCGCTGCTCTGGCCCA
AAAACATGACACAGACAAACGTGTTGTCGCCTATGCAAGCCGCCCACTGAGTAGCG
TGGAGATTAAGTTCTCAGACTGTGAGAAAGCCTTGCTAGCCACCATCTGGGCTGTGG
AACATTTCCGCAGTTACATTGGTGGTCAAAAGATCATCATTGAAACTTGTCACCAAC
CTGTGTCGTTCTTGAACAGCCAGAGGCTCAGAGAAGGCAGAGTTTCCAACAGTCGC
ATTGCCGCGTGGATGATGGCCCTACAGGGCTATGATGTTGAAGTGAAGTACTCAAA
GAGTCACAAGGTAACACTCACCCAAGGGCTAGCAGACTGCCAACACTGCGACTGTG
AGGGACAACCAGATCCGTCACCGCTGGTCACAATCTCACCGACGCTTCCCTCTAACC
ACCACTACTTCAATGAAAATGCCTGCCAGGGCCTCCCTGAAGTTTATGTGGACGGCT
GCTCCTTCCGCTGTGACAGCCAACTCAGGGCTGGAGTAGGTATCGTGTGGGCCAATC
GCCAGTTTGAGGAGCCAAACCACTACCTGTTAGGCCCTAAAACAAGCCAGTATGCA

GAGATTGCCGCCGCTTTGATTGCCATTCAACAGGCAGCCAGACAAGCTCAATCGCA
ACTTGTGATATGCTCAGACTCAAACTATGCCCGCCACAGTTTCATCTCTCACTTCCCC
ACATGGAAAGCCAGTGGGATGAAGAACGCCAGGAACAAACCTGTAAAACACGCAG
AGTTATTCCTGGCATGCGATCAACTGGTAACAGACAAGGGAATGACGGTGTACTGG
CGGAAAGTCAAAGGTCATTCCCAAACCTCTGGCCCAGACAAAGAGGGGAACGATGA
GGCAGACCGCCTGGCGAGACTGGGAGCAGAGGAGGGCCCCCTGTGGGAACTCCAAC
CTGACGGGCTCCCAACTCCCCACAGCTGTACGGTTAATGCAGTCACCCGTAGGCAGT
CCAGAGAACAGCAACATAGCCAGCCTGACAAGTACGAGATTCTGCACCTGGGTAGG
AAACCAGGAGATGCGGATCTCATCACCATGCAGGAACAGGACCCCACCATCTGCGC
CGCCAGGAAGTGGGTAACCAACCCACCAGCAGCGGAGTCACTCCCTGACAGAACAG
AGGAGCTGAAAGACCTGGGAGCAATGAGGCATGCCCTACCGCACCTGAGGCTGGAG
AAAGGCCTATTGGTGTATGTCCATCACAACCAAGGACCAATCCGCTGGGTTGTCCCC
ACAGACCATAGGGGGGTGATGCTCAATCATGCCCATGATTCACCAGTAGGGGGTCA
CCGAGGCTACAAGGCCACTCTCCAGACCTTGCAGCAAGTGGCATACTGGCCCTCTAT
GAGTAAAGATGTGAAAATCCATGTACAAAGTTGCCTGGTTTGCTGTCAATTCAACCC
ATCACGACCACTAAACAGGGCCCCACTTCAGAAAAGGGGTATGACATTCCCCTGGT
CCCACATCCAAGTAGATTGGATCGGACCCGTCCCAAAATCCTCCAGAGGTAATAAA
TATCTCCTTACAGTGACCTGTGCATTTACCAAGTGGGTCGAGTGCCTGCCAGCACCA
AATGACACAGCCGTCACCACCGCAGTCCTTCTACTAAACCATGTGTTCAGCCGTTGG
GGGCTGCCATTGTCAGTCGACACAGACAGAGGGACACATTTCACCTCCCATGTCATG
ACCGAACTATACCATATCCTAGGGGTGAACGTGAGATTCCACATATCTCACCACCCT
CAATCATCAGGACAAGTAGAGCGTGCAAACCGCACCATTGTCCAGATGCTTAAGAA
ATACGTGAGCAGCCACGGCAAGGACTGGGACATGAAGTTGCCCCTGGTGCTAATGG
CTATCAGGGCCACCCCACACAGAGCTACAGGTGTCACACCTTTCGAGATGATGACTG
GAAGACAAATGACTTTACCCCTCCACCTCCTCTACCAGGCAGAAGACGTCAGTGTGG
TGACAGCCTACACCGCACACCAGTACGTGTCTGACCTGCGGGAGCACCTTCAGACG
ACCTTTGCTTGGGCCCAAAGCAACTTGGGATCCAGTGCAGAGGGAGCCAAAGCCTA
TTACGATAGAAAAGCATCCCACCGAGAGTACGAAGTAGGCGATAAGGTCCTGTACT
TTCAGTTCGCAAAACCTGTCGGACGCTCTGCAAAGTTCCTACCCAGCTGGTCAGGCC
CCTTTGAGATTGTCGGAAAACTCTCCCCGGTAGCCTATAGAATACAGGTTTCAAAGC
CCAGAAAGTCACCGGCATACAGATGGGTCCATGCAAACCAGATAAAGCCCTTCAAG
CTTAACGTTTCCCAGCCAGGGGAAGCTGCTGTGTCAACCCAAAAGGACGAGCAGCC
AACCCCTCGTAGCAACAACTAACTCACAACCAGTCAGGGCAGCATGCTTAACTCCA
CCTGCATGCACCGTAAGGCATAAATCCAACCCTATAATAACCAGAGCTTCATCTGTT
GCAGATGAAACTCCCTTGGGCCATCATCATCCTCAACATACAGTGGGCCCTCTCCAT
TGAGATAACCGAACCTGGCCCGCCAGCCGGCATCGTCCTACAAAACGTACCTGGCC
TTCTGATCACACATTGCCGTCTTTACACCCAGAGGATCTATGTCCGACTAGACCCCT
GGGACGCCTATCGCAATCACATCCACCTACCCCCTCGGTTAACTGAAGGCAGGCCA
AGCGGGACCCAAACCCACGAGACTGTGGAACACACCCGCTCCACAGTGATCCACGT
CCTAGAACAGCTGCAGAAGTTCCTTGTCACAGAGGCAGACCTCAGCGGCAAAGGAA
GACCAAAACGGTTCCTAGGAGGTCTAATCGCAGCAGCATCAGCTATCGGGTCCCTCT
TCTCTATCGGCCTATCTGCCGCAAACTCAGTTAGCCTTAGCACACTCCAGCGACACG
TTACCGAGCTTGATGAAGAGATGCCAGAGGTTCACCAAAAGCTTAGCTTTCAACAGC
AGGAGTTGCAGAGTTTAGGAAAGACAGTACAGGGTACTATAACGACAGTAAACCTC
CACTCAGCTCTCCTCAACAACACCTTGCATGCCCTCGAAACCCTGTCAGGGATTGTC
AAAGAAGACCTTACCTACGTACGAGTGGTGAGAGACTTGATGCAAGAACTCATCAG

AGAAGTCAGCTCATCAATCAACAGCCTCAGTAGTGGTAGAATCCCCACTTACTTAGT
TAACTTAGATTTGGTCAGAGATATCCTAGGATCCGCCACCACCACAGTGGTGCAACC
TTCGCAGGTCCATCTAGCCTACAACCTAGGCAGTGCTATTCCCATAGCTGTAGACCC
CAAGAACCTGGAACTAGGCTTTATTCTGAATTTGCCAATCATAGAAACTAGGAATAT
ATATAAGTTAAAGTCTGTGTTGAATGTAGGATTTTGGCAAGACAACGCTCACATTCT
ACTTAAGACACCTCCAGTGCTAGCATACCATAGTGAAGACCCATCTCTCTACCTCGT
TCCAAATCTCAACCTCTGCACCAACACCAGGGACATCCATTGGGTCTGCCCAAGTAA
CCCTTTCATCAGGGACACGACCAACTCCCTATGCGGCCTTAGGCCCGACGCCCCCGA
GCAAAAGTGCCAGGGCAGCCTGTCAGACAGGGACAGAGCAGCAGAACCTAGGGTG
GAACGAGCTGGCAACCGATGGCTACTCAGCACACCCGCCACCGAAGCCTTGATTGT
CTATGACCACCATGACACAACCACTACGTTGGCCCTACCAAACCAAACCCTGTTCTT
AACGGTACCTCAGGGTGCTACCGTCCATATTGAGGACATCATCCTCCATTACCTCAG
TTCAGAGAGACATGACACTGAGATTGAGATCATGGACGTGTTCAAGGGTTACAATA
TCACGATTGATGACACCCTTCAGCGGCAACTCATGCTAGAAGGGACCAAATTGGTA
AAGTTTAGCCTCAGGCAGACGGGGCTGACCACAACCTACGCGGATCAGGTGACTCG
CCCGGTAACCTACAGCGAACACCCAATCAGCATGGTCGCGTTGGGACTCCTTCTTAG
TGGTTGGGTCATTACAGCAATGGTAGCATACGGAATGTACAAATATATCAAGGAATT
ACAAGCTAGGTTAGATTCTGCCCTGGTAGTGCCAGTACGCTTTCCCCCACAGCTTAA
CGCACCTACATTCAACTTAGCCGCCAGGACCTAAGACACCCCAGGCTCCTGTCTTCC
TCTCTCACACACATCTGTCCTTATCTGTCCTATCTTCCCTTTGCTCGTTTATTATTTTC
CTTCATGGTTAACAAGAACCTTATGATGGTAATATCTAGAAATATTAGACTGTTGAC
TCTGCTCAGACATTGCTTTTGCCTCTGCTCAGACATTGCTTTTGCCTCTGCTCAGACA
TTGCTTTTGCCTCTGCTCAGACATGCCTCTAAAATGCTCAGACTTGCCTCCAACTCTA
GCAAACAAGGACTGTTTGGAATACATGGACTGTCAATGAACATTGATAATGCCCAC
AGGAGATGACAAGTACACTGAGGACCTTAACGAACCTGTCAGAGCACATGGACCCT
TTCAACGTCTTGTTCCGAAGCGACGGCTGACACGCCATCATTCAGAACAAAAAGGG
GGGATGTTGGGGGACCTCATATGTATAATCTGTTATTAACAGAATTATTATTAAGGT
CCCACATGCTCAGACAAAGGAAATACTGCGGCCTGCATGTAAAGTCCTCATCCTGAT
AGTAGTGAATGAGAACAAAGAGATTTCAAATCTCGGCGCCAAAACCAGGCGCCAGG
CTGGGGAACCACACCTGGCTCCCCGCTGTGAAGGCCGAAACCCCGCTGACACAAGC
GACAAACGGAGGAGAAACTTGGACGTCAGCCATCCTGATTGGACCAAACAATAGGA
TGGGCGTGACCAAAAACACGGTTATAAAAGTCGCCGTGACCCAAGACTCGTCCCTT
GTCCTCTTCGATCACCCGTTACCCCTTGTTCTCCCCCTTGTTCTCTTCGATCTCCCGTT
TCCCCTTCCTCTCTTCGACCCTCTCGCTGCCCCACCTGGAGCTTCCCTCGGCACCAGC
ACGCCTGAAGCCTCCACCGCAGCCAGCAGCTGAACCAACGGAGACGTCGTCACAGA
GAACGGGCCTGATCACCCATCTACTTCCAGCCATACGGCAGGAAACCGCCAGCTTA
ACCGCCTCACTCGCCGCCGCATCAGCTGATCACGACGCGATCACCGCGACGCACAC
CTGGACCGGACCGGACCGCAACAGAGCAAGCACGCACGCACGCACGCACGCAACG
CCGGAGCTTCTCACCAGGATTGAGCAGTAAGGCAGAGGCATTCTTAAGTCTGAGCA
GAGTAGAACTCTTAGGGTATCAGTAATTGTTTTATTGTTGTGTTTCTTTTCCTGCACA
CACGATCTGACCGCTGTGTATTTCCGCGATCACGTGACTGTAACGCTGTTACAGATC
TACGTGAATCTAAGAGGTGCACATTGAAGTGTTGTAACCTGTGATTTCGCTAGCTTA
AGTGCTTTATTGTGTTACTGTGTAGTGCTCTGTTTCTGCTGCGACCACGCCAAGCGCA
CGTTTCAACCACTGCACGCTTCCGCGATCACGTGACTATATCGCAGAATATAGCCCC
ACGTGTTTGTGTGACCACGAGCCTTATCCACGCTAGTCCCACTTCACACTCCTTTACA
CGCTTCTTTCCCTGTGAGCGCCAGACACGCGCGCTCCATTCAAAGGGCTCACCGGCA

```
CACGCGCGCAGCTGTGCCACCTCACCACGCCCCTAAAGGGGGGCGCTCTCCTCCTCC
TCCTCTTCTCTCTTTCTCTCTCTCTCTCTCACCCTCACACACATATACACACACACAC
ACACACACACACACCCCTCTTATCTCAGGTAACGCGCATACACATACAGAACGCTA
GACGAACACACCATATAATTAAGTTGTGATTTGCTGAGTTATTCAAGATAGTGTTGA
CACTAGAGTTAATATTAACTGTTTAATAAAGCGTTCATAATTTAACTAGTCGTTGTTT
GTGATTATTTGTATATGGTTTTGCAGCACTAGTGAATTGAGTCGGCTATGGTACGGA
GTTCACCCTTCAACTAACTTAAATACAAATGAGACTGTATTGGCTATCCCAAATCGG
TTATTGGTCCCTGGAAACAGGGTGGTGCCCCGTAAATTAAGTAATTATTAATTCAAT
TAATAATTATTAAATCAATAATAGTAAGTGTTTCACTTGCCAAACTATAGACATTTG
TCATATAGTTATATATATTTAAGCTAATAACCAAGATTTGTGAATCTAAATTATAAA
CATAATTGGTATCTCCACTCAGGAGCTATAAATCCAATTATTATAATTTGTGCTCCTC
GCATATCCCCAACAGAAGTGAACAAATTT
> Loki-Bsp consensus sequence 5’LTR
TGTTGGGGGACCTCATATGTATAATCTGTTATTAACAGAATTATTATTAAGGTCCCA
CATGCTCAGACAAAGGAAATACTGCGGCCTGCATGTAAAGTCCTCATCCTGATAGTA
GTGAATGAGAACAAAGAGATTTCAAATCTCGGCGCCAAAACCAGGCGCCAGGCTGG
GGAACCACACCTGGCTCCCCGCTGTGAAGGCCGAAACCCCGCTGACACAAGCGACA
AACGGAGGAGAAACTTGGACGTCAGCCATCCTGATTGGACCAAACAATAGGATGGG
CGTGACCAAAAACACGGTTATAAAAGTCGCCGTGACCCAAGACTCGTCCCTTGTCCT
CTTCGATCACCCGTTACCCCTTGTTCTCCCCCTTGTTCTCTTCGATCTCCCGTTTCCCC
TTCCTCTCTTCGACCCTCTCGCTGCCCCACCTGGAGCTTCCCTCGGCACCAGCACGCC
TGAAGCCTCCACCGCAGCCAGCAGCTGAACCAACGGAGACGTCGTCACAGAGAACG
GGCCTGATCACCCATCTACTTCCAGCCATACGGCAGGAAACCGCCAGCTTAACCGCC
TCACTCGCCGCCGCATCAGCTGATCACGACGCGATCACCGCGACGCACACCTGGAC
CGGACCGGACCGCAACAGAGCAAGCACGCACGCACGCACGCACGCACGCAACGCC
GGAGCTTCTCACCAGGATTGAGCAGTAAGGCAGAGGCATTCTTAAGTCTGAGCAGA
GTAGAACTCTTAGGGTATCAGTAATTGTTTTATTGTTGTGTTTTCTTTTCCTGCACAC
ACGATCTGACCGCTGTGTATTTCCGCGATCACGTGACTGTAACGCTGTTACAGATCT
ACGTGAATCTAAGAGGTGCACATTGAAGTGTTGTAACCTGTGATTTCGCTAGCTTAA
GTGCTTTATTGTGTTACTGTGTAGTGCTCTGTTTCTGCTGCGACCACGCCAAGCGCAC
GTTTCAACCACTGCACGCTTCCGCGATCACGTGACTATATCGCAGAATATAGCCCCA
CGTGTTTGTGTGACCACGAGCCTTATCCACGCTAGTCCCACTTCACACTCCTTTACAC
GCTTCTTTCCCTGTGAGCGCCAGACACGCGCGCTCCATTCAAAGGGCTCACCGGCAC
ACGCGCGCAGCTGTGCCACCTCACCACGCCCCTAAAGGGGGGCGCTCTCCTCCTCCT
CCTCTTCTCTCTTTCTCTCTCTCTCTCTCACCCTCACACACATATACACACACACACA
CACACACACACACACCCCTCTTATCTCAGGTAACGCGCATACACATACAGAACGCTA
GACGAACACACCATATAATTAAGTTGTGATTTGCTGAGTTATTCAAGATAGTGTTGA
CACTAGAGTTAATATTAACTGTTTAATAAAGCGTTCATAATTTAACTAGTCGTTGTTT
GTGATTATTTGTATATGGTTTTGCAGCACTAGTGAATTGAGTCGGCTATGGTACGGA
ATTCACCCTTCAACTAACTTAAATACAAATGAGACTGTATTGGCTATCCCAAATCGG
TTATTGGTCCCTGGAAACAGGGTGGTGCCCCGTAAATTAAGTAATTATTAATTCAAT
TAATAATTATTAAATCAATAATAGTAAGTGTTTCACTTGCCAAACTATAGACATTTG
TCATATAGTTATATATATTTAAGCTAATAACCAAGATTTGTGAATCTAAATTATAAA
CATAATTGGTATCTCCACTCAGGAGCTATAAATCCAATTATTATAATTTGTGCTCCTC
GCATATCCCCAAC
> Loki-Bsp consensus sequence 3’LTR
```

GTTGGGGGACCTCATATGTATAATCTGTTATTAACAGAATTATTATTAAGGTCCCAC
ATGCTCAGACAAAGGAAATACTGCGGCCTGCATGTAAAGTCCTCATCCTGATAGTAG
TGAATGAGAACAAAGAGATTTCAAATCTCGGCGCCAAAACCAGGCGCCAGGCTGGG
GAACCACACCTGGCTCCCCGCTGTGAAGGCCGAAACCCCGCTGACACAAGCGACAA
ACGGAGGAGAAACTTGGACGTCAGCCATCCTGATTGGACCAAACAATAGGATGGGC
GTGACCAAAAACACGGTTATAAAAGTCGCCGTGACCCAAGACTCGTCCCTTGTCCTC
TTCGATCACCCGTTACCCCTTGTTCTCCCCCTTGTTCTCTTCGATCTCCCGTTTCCCCT
TCCTCTCTTCGACCCTCTCGCTGCCCCACCTGGAGCTTCCCTCGGCACCAGCACGCCT
GAAGCCTCCACCGCAGCCAGCAGCTGAACCAACGGAGACGTCGTCACAGAGAACGG
GCCTGATCACCCATCTACTTCCAGCCATACGGCAGGAAACCGCCAGCTTAACCGCCT
CACTCGCCGCCGCATCAGCTGATCACGACGCGATCACCGCGACGCACACCTGGACC
GGACCGGACCGCAACAGAGCAAGCACGCACGCACGCACGCACGCAACGCCGGAGC
TTCTCACCAGGATTGAGCAGTAAGGCAGAGGCATTCTTAAGTCTGAGCAGAGTAGA
ACTCTTAGGGTATCAGTAATTGTTTTATTGTTGTGTTTCTTTTCCTGCACACACGATCT
GACCGCTGTGTATTTCCGCGATCACGTGACTGTAACGCTGTTACAGATCTACGTGAA
TCTAAGAGGTGCACATTGAAGTGTTGTAACCTGTGATTTCGCTAGCTTAAGTGCTTT
ATTGTGTTACTGTGTAGTGCTCTGTTTCTGCTGCGACCACGCCAAGCGCACGTTTCAA
CCACTGCACGCTTCCGCGATCACGTGACTATATCGCAGAATATAGCCCCACGTGTTT
GTGTGACCACGAGCCTTATCCACGCTAGTCCCACTTCACACTCCTTTACACGCTTCTT
TCCCTGTGAGCGCCAGACACGCGCGCTCCATTCAAAGGGCTCACCGGCACACGCGC
GCAGCTGTGCCACCTCACCACGCCCCTAAAGGGGGGCGCTCTCCTCCTCCTCCTCTT
CTCTCTTTCTCTCTCTCTCTCTCACCCTCACACACATATACACACACACACACACACA
CACACACCCCTCTTATCTCAGGTAACGCGCATACACATACAGAACGCTAGACGAAC
ACACCATATAATTAAGTTGTGATTTGCTGAGTTATTCAAGATAGTGTTGACACTAGA
GTTAATATTAACTGTTTAATAAAGCGTTCATAATTTAACTAGTCGTTGTTTGTGATTA
TTTGTATATGGTTTTGCAGCACTAGTGAATTGAGTCGGCTATGGTACGGAGTTCACC
CTTCAACTAACTTAAATACAAATGAGACTGTATTGGCTATCCCAAATCGGTTATTGG
TCCCTGGAAACAGGGTGGTGCCCCGTAAATTAAGTAATTATTAATTCAATTAATAAT
TATTAAATCAATAATAGTAAGTGTTTCACTTGCCAAACTATAGACATTTGTCATATA
GTTATATATATTTAAGCTAATAACCAAGATTTGTGAATCTAAATTATAAACATAATT
GGTATCTCCACTCAGGAGCTATAAATCCAATTATTATAATTTGTGCTCCTCGCATATC
CCCAACAG
> Loki-Bsp consensus sequence *gag* gene
TTGTGCGATACTATTCTAGCTGTTGTGAAAACCTTACAAATAAGGATGAGTAACAAG
GGCTCCCCTGGGTCAGGAAGTCTGGACTTCCAAGGTGTTGTTGCAGGACTTGTCAGA
ATATCCACGAAAGAGCTAGAGCGCCTTAACAAGTATACTGTCGCAGAGTGTGAGGG
AGAAATGCAGAGGCTGATGTCGGTCGTGAGTCAACCCACGGGAGAAACCCGCCTGA
CTCAAGAGGTAGTGGGCCAGCTGGCATTGCTATGCCATCAGAAAGCCCAACTGGAG
TCCGCAGAATCTGCTGAGAAGATCGCGGCCCTACACAGTAGCTACTTAGCTGAGCA
AAAGACCGTCACCCATCTAGAGCATAGCTTAGAGGTAATGAAAAGGGAACTGGAAC
AACTCGAGGAGAGCTATAAGAGGTTGCAGCGAGACCATGACATACCCCCTAGGTCC
CGAACAGTTCACCAGGAGGAGGTAGGAGATAGGAATGAAAATCAGTGCAGACAAA
TAGGTATGGCGGCTCAGTCAAGTGACCAAGAGCTGGACAAACCTGGTCCTGGGAGT
GAGACTTTCCACAGCCTAGAGGCCCCAAAGTCAGTCCCCACAAACCGGACCAATCT
AGCACAAAGAGAGCCCCGTGGGTCCTTTGGAGAAGCGACCCCAGCTCACGCCCACA
AGGGCACTCATTCAAACTCTTACCTCACTACTGACACCCCAGTCTACTACGCATCGA

CCATAGCCCATGGTAGCGTAGTCAGGGGAACAGGAAGACTGAACAGGGACAGGTAT
GCACCCATGGATACGCCAATCCACAGAAATGTACTCCCGAGTTACCCCCAGCCAGA
GCCACACCTCGAAGACCAGAGATGGCAAGGTCCTAGTAGAGAAACCTACTTTAAGC
CTGAGCCCACTAATGGTAGATATGAGCCACTGCCACCCAGTGGGGGATACCTTCCTC
GCAACGCCCCCCCCCTACGCCCTCACCCAGTCCCTTACTACGGGGGCGACTCAAGCT
CAGAGGACTCTGACATTGTGAACCCTGTCCCAAGGCGGGAGATACGCACACGCCAA
CTCGAATCTATGGCTAAGGACATTGAGCGCTTTGATCCCAGTGACAGGGAGTCCACC
ATAGATGACTACCTGCGGGAAGTTGAGCGGTGCCTCCTGGACTTGCCTGAAGCAAC
AGCTCGCGAAAAGCTGAAGCTGATCTGGAAGACCACGGCTAGAAGTGTGCATAGTT
TTATAGAAACCCTTCCCTGCAGCATACGTGACAGTTACCCAAGGCTCTGTGATGCCC
TTAGGGAAGAGTTCTCGCCCTTTACTGACTCTGCCGCTGCCACCCTTGGAGCATTTG
AGGTCCTGCACAAAAGGGACGAATCTCCAAAGGAATATTATAGGCGCCTCCGAGCA
GCCTACTTCCAGGGTCACAATGCCCCAGGTCTAGAGGAAGACAGCGCATTCCGATC
CCTCTTCCTGCATAACCTCCACGAGAGCATTCGGTATGAGGTGACAATGCATAGCCA
ATCAGGTCGACTCACTATGCAGGAACTCCGAAAATACGCGCAAATGGCCTGGCAAA
CCCGCGCACGACCAAGGCGTCGCCCTGACCGGGACGCCAGAGTCCTGGGCATCCAA
ACTACGGACGACTTGGCGTTAGAAGGCCCAGAATTACAGCGGGGAAGGGGAGAAA
ATAGAGTACAACCCCATGCCACAGAGTTTTCCCATAGGCCCGGGAGTCGACCCAGA
CAAGGTGGGGAACACACCCACCGGGCCTCAGAGAACAACCAACAGTATGACACTGC
ACCACAGAGATATAAGGGACGGAGTGAACAAAAGTCAGCTCCGAAAGGGAAGTAC
AGTAGAAAGGCTGATAACCATCTCAGTGCAGAACTGAGAGCAGCCATTGCGCAAGG
CATAGCTGATGCTCTGGCCACGCAAAATGCACCTGGGGGTTCCGCCCCGCCTCAGAA
ATCCCCAAAAAAGCAAATAAAAAGTGAACCGTCTAAGTGA
> Loki-Bsp consensus sequence Gag protein
LCDTILAVVKTLQIRMSNKGSPGSGSLDFQGVVAGLVRISTKELERLNKYTVAECEGEM
QRLMSVVSQPTGETRLTQEVVGQLALLCHQKAQLESAESAEKIAALHSSYLAEQKTVTH
LEHSLEVMKRELEQLEESYKRLQRDHDIPPRSRTVHQEEVGDRNENQCRQIGMAAQSSD
QELDKPGPGSETFHSLEAPKSVPTNRTNLAQREPRGSFGEATPAHAHKGTHSNSYLTTDT
PVYYASTIAHGSVVRGTGRLNRDRYAPMDTPIHRNVLPSYPQPEPHLEDQRWQGPSRET
YFKPEPTNGRYEPLPPSGGYLPRNAPPLRPHPVPYYGGDSSSEDSDIVNPVPRREIRTRQL
ESMAKDIERFDPSDRESTIDDYLREVERCLLDLPEATAREKLKLIWKTTARSVHSFIETLP
CSIRDSYPRLCDALREEFSPFTDSAAATLGAFEVLHKRDESPKEYYRRLRAAYFQGHNAP
GLEEDSAFRSLFLHNLHESIRYEVTMHSQSGRLTMQELRKYAQMAWQTRARPRRRPDR
DARVLGIQTTDDLALEGPELQRGRGENRVQPHATEFSHRPGSRPRQGGEHTHRASENNQ
QYDTAPQRYKGRSEQKSAPKGKYSRKADNHLSAELRAAIAQGIADALATQNAPGGSAP
PQKSPKKQIKSEPSK*
> Loki-Bsp consensus sequence *pol* gene
CTGGGGGAGAGAAGCTGGCGTGCCCCAAGTAGACACTGCTCGTACACCAGACAACA
ACACTCCCCCCTTCTTAACCTTCTTGGGAGACCTTGTCCGACAGGGTATGCCCGCCG
CATTTACACGTCAGTGGTGATAGGGGGATGCTTCAGCTCCCTGGCCTTACTGGACTC
TGGTTCTGAAATCAGCCTAATGTCTTCAGATACTTTCAAGGCAATGTCCAGAACCAT
GTTGTCTCTTGGGAAACCCCTACAGGTCGAACCGTGTAACATTAACATCACTAGTTA
CACACAAGACAAGTCCCCCATTGTTGAACGAGCCTGGGTTACTATTGCTTTCCAAGA
GGTAACTCTTGTGCACCCTTTGTATATCTGCAATCTTGACACAGAGCCAGTGCTTATT
GGCCAAGATCTGTTGGGCCGAGTGGCACCTCTTATTGACTGTCACAGAGGTCGCATC
TGGGCCCAGGTGAAAGTCCCCAAACCATTGGATCTTGGGGACAACCGACCAGGATG

TGACGGACAGGTCACTGCAGTGGAGAGGCAAGGCCAAGAGTCCAACCCCAGCTCTG
AGGTGCCAGACTTGCCAAACTCCTTCCAGACCCATGGTTCATTTCTCTGCGCTCTGA
CAAATGAGGGCCAAATCATCTACAGCCCAAAGATCGTGGGGGGCATTCATCTTGAC
AACCTTCCTGTGCCTGAGGCGCAACTGACACTCCGGTCAGAAAAGACCGCCATCAG
CCAGGAATTGTATGACGCCTTGTGCTTGCAACGACCACCTCCCACCTTCGTGCAGCA
TAGTCATAGGCTCCTCCCAGCAGATCAGTCAAGACCCCTCCTTAAGGCAATCGGGGT
CTGCGCCGCCACCATTCGAATTGGGGGTCGGGTCATCAAACATTTCCTAAGTGTAGT
TCCCCACTTGCAGCCGCCTGCTCTTGTGGGCGCCGATCTCCTTGTGCGGCTAGGTAC
GCACCTTGACACCATCAACCAAGTCTTATGGTCGCAGGTCAGCCCAAGAGACACCTC
ACAAAGTCCTGAAATAGAGAACATGAAATCGGGACAGACCATTCCCCACGCGTGCC
AAGTTGCTGCGGAGAGCGACAAGACCATTCCCCCCAGAACGGCAGGTAACCCGGTC
CGTCTAGTAGTCATGAAGGGGCAGAAAATCCATGGCTCACAAGCATTCTTCCAACCC
CTTCCTAGTTTCGCCGAACTTAACCTGACGGTCTGTGGCACCCCTCTCCTAGAACTG
GACAATCGTGCAACCTACCTCTTGGTTCAGAATCCAACCAATGCACCGATCCTTGTT
AAGGCCCAACAACCGCTGGGCGTCTTGGTAGAAAGTTCGTTTCATGACTTTGAGCTG
ACGGTACCCGTCATTGGGGAATTACCAGCGTCTCTAATCAGGGAGGGAAGTGTGGG
AGACGTCCTCCTGTCTCACCCAACGGGGATGATCTCTGTATTGCGCTATGAGCCACT
GAGAGACCAACCAATCTGTGGGGCAACACTTGAACCAGGGGGTGACATGCTGGTCT
ATACCTTGGCTGCCCAGCCAGAGTGCACTGCCCAAGCGGGAACCAACACCTCTCCG
GCGGGAGAGGAACCTTACCCCGGCTTTGAAGCAGAAATCAACCAGCAACTTGAAAA
AGCGGACGCATTGACAAATGAGGGGCAAAAAGAAGCCCTTCGAGATCTCTTCCGCG
AATTCAAACCCATCTTCTCACGAGATTCGCATGACTGTGGGGTCACTGACTTACACA
CCGTGCGCATTCCCACTGACCCCCATGCACCACCAACGTTTGTGCGACAGTATAGGA
TTCCTCTGGCAGCATATGAGTCAATCCAAGAGATTCTGGACACACTGCTGAAGAAGG
AGATCATAAGGGAATGTAACTCGACATATAACTCGCCGATCTGGCCGGTTCTGAAAC
CGACAGGGAAATGGCGTCTTACCATCGATTACCGACCCTTGAACAAAAACGTTCCCT
TGTCTCGATGGCCCATGATTCACTTAGACCAAGAACTGGCTAAGATAAAAGACGCCC
ATTTCTTCTCCACTGTAGATGTTGCCAATGGCTTTTGGACCATGAGAGTGGATCCAG
CTGACCAGTATAAACTGGCCTTTTCCTTCGGCAACCGCCAGTACACCTGGAACCGTT
GCCCCTTCGGGTACTCAAACTCCCCTGCAGAGTTCAATATCTTCCTCCATAAAGCAA
TGAGTGACGCTGCCGCTCGCGGTAACCTCATCTACGTAGACGATATCCTCATGAGGA
GTCAGACATTTGAAGGACACTTGGCAGAAATCCGACATGTGCTCTCACAACTCCTGG
CTGCAGGAGCCAAGTTAGCATTAACCAAGGGTCAGTGGTGTCGAACCAAGGTGAAG
TATGTCGGGCTGACAGTGGGCCAACATGGTATTGAGCCTCAGGCAGACAGAATCCA
GGCTATTAGAACCATTAAAGCCCCAAACACACTGTCAGGACTCAGGAGCTTCCTAG
GGGTCTGTAACTACTCACGCCAGTTTGTAGAAGACTATGCGGAGATAGCCAGACCCC
TCACAGAACTCCTTCGTAATGATACTCCCTATATATGGACGGAGCGCCAACAACAGT
CATTTGAGAAAATGAAAGAAGCGCTCAGCACAGCCCCGTGCCTAGCTTATCCTGAC
AAGGATAAAGAATTCTATCTGGAAGCTAGCTTCTCCCAACAAAGCCTCAGCGCTGCT
CTGGCCCAAAAACATGACACAGACAAACGTGTTGTCGCCTATGCAAGCCGCCCACT
GAGTAGCGTGGAGATTAAGTTCTCAGACTGTGAGAAAGCCTTGCTAGCCACCATCTG
GGCTGTGGAACATTTCCGCAGTTACATTGGTGGTCAAAAGATCATCATTGAAACTTG
TCACCAACCTGTGTCGTTCTTGAACAGCCAGAGGCTCAGAGAAGGCAGAGTTTCCAA
CAGTCGCATTGCCGCGTGGATGATGGCCCTACAGGGCTATGATGTTGAAGTGAAGTA
CTCAAAGAGTCACAAGGTAACACTCACCCAAGGGCTAGCAGACTGCCAACACTGCG
ACTGTGAGGGACAACCAGATCCGTCACCGCTGGTCACAATCTCACCGACGCTTCCCT

CTAACCACCACTACTTCAATGAAAATGCCTGCCAGGGCCTCCCTGAAGTTTATGTGG
ACGGCTGCTCCTTCCGCTGTGACAGCCAACTCAGGGCTGGAGTAGGTATCGTGTGGG
CCAATCGCCAGTTTGAGGAGCCAAACCACTACCTGTTAGGCCCTAAAACAAGCCAG
TATGCAGAGATTGCCGCCGCTTTGATTGCCATTCAACAGGCAGCCAGACAAGCTCAA
TCGCAACTTGTGATATGCTCAGACTCAAACTATGCCCGCCACAGTTTCATCTCTCACT
TCCCCACATGGAAAGCCAGTGGGATGAAGAACGCCAGGAACAAACCTGTAAAACAC
GCAGAGTTATTCCTGGCATGCGATCAACTGGTAACAGACAAGGGAATGACGGTGTA
CTGGCGGAAAGTCAAAGGTCATTCCCAAACCTCTGGCCCAGACAAAGAGGGGAACG
ATGAGGCAGACCGCCTGGCGAGACTGGGAGCAGAGGAGGGCCCCCTGTGGGAACTC
CAACCTGACGGGCTCCCAACTCCCCACAGCTGTACGGTTAATGCAGTCACCCGTAGG
CAGTCCAGAGAACAGCAACATAGCCAGCCTGACAAGTACGAGATTCTGCACCTGGG
TAGGAAACCAGGAGATGCGGATCTCATCACCATGCAGGAACAGGACCCCACCATCT
GCGCCGCCAGGAAGTGGGTAACCAACCCACCAGCAGCGGAGTCACTCCCTGACAGA
ACAGAGGAGCTGAAAGACCTGGGAGCAATGAGGCATGCCCTACCGCACCTGAGGCT
GGAGAAAGGCCTATTGGTGTATGTCCATCACAACCAAGGACCAATCCGCTGGGTTGT
CCCCACAGACCATAGGGGGGTGATGCTCAATCATGCCCATGATTCACCAGTAGGGG
GTCACCGAGGCTACAAGGCCACTCTCCAGACCTTGCAGCAAGTGGCATACTGGCCCT
CTATGAGTAAAGATGTGAAAATCCATGTACAAAGTTGCCTGGTTTGCTGTCAATTCA
ACCCATCACGACCACTAAACAGGGCCCCACTTCAGAAAAGGGGTATGACATTCCCC
TGGTCCCACATCCAAGTAGATTGGATCGGACCCGTCCCAAAATCCTCCAGAGGTAAT
AAATATCTCCTTACAGTGACCTGTGCATTTACCAAGTGGGTCGAGTGCCTGCCAGCA
CCAAATGACACAGCCGTCACCACCGCAGTCCTTCTACTAAACCATGTGTTCAGCCGT
TGGGGGCTGCCATTGTCAGTCGACACAGACAGAGGGACACATTTCACCTCCCATGTC
ATGACCGAACTATACCATATCCTAGGGGTGAACGTGAGATTCCACATATCTCACCAC
CCTCAATCATCAGGACAAGTAGAGCGTGCAAACCGCACCATTGTCCAGATGCTTAA
GAAATACGTGAGCAGCCACGGCAAGGACTGGGACATGAAGTTGCCCCTGGTGCTAA
TGGCTATCAGGGCCACCCCACACAGAGCTACAGGTGTCACACCTTTCGAGATGATG
ACTGGAAGACAAATGACTTTACCCCTCCACCTCCTCTACCAGGCAGAAGACGTCAGT
GTGGTGACAGCCTACACCGCACACCAGTACGTGTCTGACCTGCGGGAGCACCTTCA
GACGACCTTTGCTTGGGCCCAAAGCAACTTGGGATCCAGTGCAGAGGGAGCCAAAG
CCTATTACGATAGAAAAGCATCCCACCGAGAGTACGAAGTAGGCGATAAGGTCCTG
TACTTTCAGTTCGCAAAACCTGTCGGACGCTCTGCAAAGTTCCTACCCAGCTGGTCA
GGCCCCTTTGAGATTGTCGGAAAACTCTCCCCGGTAGCCTATAGAATACAGGTTTCA
AAGCCCAGAAAGTCACCGGCATACAGATGGGTCCATGCAAACCAGATAAAGCCCTT
CAAGCTTAACGTTTCCCAGCCAGGGGAAGCTGCTGTGTCAACCCAAAAGGACGAGC
AGCCAACCCCTCGTAGCAACAACTAA
> Loki-Bsp consensus sequence Pol protein
LGERSWRAPSRHCSYTRQQHSPLLNLLGRPCPTGYARRIYTSVVIGGCFSSLALLDSGSEI
SLMSSDTFKAMSRTMLSLGKPLQVEPCNINITSYTQDKSPIVERAWVTIAFQEVTLVHPL
YICNLDTEPVLIGQDLLGRVAPLIDCHRGRIWAQVKVPKPLDLGDNRPGCDGQVTAVER
QGQESNPSSEVPDLPNSFQTHGSFLCALTNEGQIIYSPKIVGGIHLDNLPVPEAQLTLRSEK
TAISQELYDALCLQRPPPTFVQHSHRLLPADQSRPLLKAIGVCAATIRIGGRVIKHFLSVV
PHLQPPALVGADLLVRLGTHLDTINQVLWSQVSPRDTSQSPEIENMKSGQTIPHACQVA
AESDKTIPPRTAGNPVRLVVMKGQKIHGSQAFFQPLPSFAELNLTVCGTPLLELDNRATY
LLVQNPTNAPILVKAQQPLGVLVESSFHDFELTVPVIGELPASLIREGSVGDVLLSHPTGM
ISVLRYEPLRDQPICGATLEPGGDMLVYTLAAQPECTAQAGTNTSPAGEEPYPGFEAEIN

QQLEKADALTNEGQKEALRDLFREFKPIFSRDSHDCGVTDLHTVRIPTDPHAPPTFVRQY
RIPLAAYESIQEILDTLLKKEIIRECNSTYNSPIWPVLKPTGKWRLTIDYRPLNKNVPLSRW
PMIHLDQELAKIKDAHFFSTVDVANGFWTMRVDPADQYKLAFSFGNRQYTWNRCPFG
YSNSPAEFNIFLHKAMSDAAARGNLIYVDDILMRSQTFEGHLAEIRHVLSQLLAAGAKL
ALTKGQWCRTKVKYVGLTVGQHGIEPQADRIQAIRTIKAPNTLSGLRSFLGVCNYSRQF
VEDYAEIARPLTELLRNDTPYIWTERQQQSFEKMKEALSTAPCLAYPDKDKEFYLEASFS
QQSLSAALAQKHDTDKRVVAYASRPLSSVEIKFSDCEKALLATIWAVEHFRSYIGGQKIII
ETCHQPVSFLNSQRLREGRVSNSRIAAWMMALQGYDVEVKYSKSHKVTLTQGLADCQ
HCDCEGQPDPSPLVTISPTLPSNHHYFNENACQGLPEVYVDGCSFRCDSQLRAGVGIVW
ANRQFEEPNHYLLGPKTSQYAEIAAALIAIQQAARQAQSQLVICSDSNYARHSFISHFPT
WKASGMKNARNKPVKHAELFLACDQLVTDKGMTVYWRKVKGHSQTSGPDKEGNDE
ADRLARLGAEEGPLWELQPDGLPTPHSCTVNAVTRRQSREQQHSQPDKYEILHLGRKPG
DADLITMQEQDPTICAARKWVTNPPAAESLPDRTEELKDLGAMRHALPHLRLEKGLLV
YVHHNQGPIRWVVPTDHRGVMLNHAHDSPVGGHRGYKATLQTLQQVAYWPSMSKDV
KIHVQSCLVCCQFNPSRPLNRAPLQKRGMTFPWSHIQVDWIGPVPKSSRGNKYLLTVTC
AFTKWVECLPAPNDTAVTTAVLLLNHVFSRWGLPLSVDTDRGTHFTSHVMTELYHILG
VNVRFHISHHPQSSGQVERANRTIVQMLKKYVSSHGKDWDMKLPLVLMAIRATPHRAT
GVTPFEMMTGRQMTLPLHLLYQAEDVSVVTAYTAHQYVSDLREHLQTTFAWAQSNLG
SSAEGAKAYYDRKASHREYEVGDKVLYFQFAKPVGRSAKFLPSWSGPFEIVGKLSPVA
YRIQVSKPRKSPAYRWVHANQIKPFKLNVSQPGEAAVSTQKDEQPTPRSNN*
> Loki-Bsp consensus sequence *env* gene
CTGTTGCAGATGAAACTCCCTTGGGCCATCATCATCCTCAACATACAGTGGGCCCTC
TCCATTGAGATAACCGAACCTGGCCCGCCAGCCGGCATCGTCCTACAAAACGTACCT
GGCCTTCTGATCACACATTGCCGTCTTTACACCCAGAGGATCTATGTCCGACTAGAC
CCCTGGGACGCCTATCGCAATCACATCCACCTACCCCCTCGGTTAACTGAAGGCAGG
CCAAGCGGGACCCAAACCCACGAGACTGTGGAACACACCCGCTCCACAGTGATCCA
CGTCCTAGAACAGCTGCAGAAGTTCCTTGTCACAGAGGCAGACCTCAGCGGCAAAG
GAAGACCAAAACGGTTCCTAGGAGGTCTAATCGCAGCAGCATCAGCTATCGGGTCC
CTCTTCTCTATCGGCCTATCTGCCGCAAACTCAGTTAGCCTTAGCACACTCCAGCGA
CACGTTACCGAGCTTGATGAAGAGATGCCAGAGGTTCACCAAAAGCTTAGCTTTCAA
CAGCAGGAGTTGCAGAGTTTAGGAAAGACAGTACAGGGTACTATAACGACAGTAAA
CCTCCACTCAGCTCTCCTCAACAACACCTTGCATGCCCTCGAAACCCTGTCAGGGAT
TGTCAAAGAAGACCTTACCTACGTACGAGTGGTGAGAGACTTGATGCAAGAACTCA
TCAGAGAAGTCAGCTCATCAATCAACAGCCTCAGTAGTGGTAGAATCCCCACTTACT
TAGTTAACTTAGATTTGGTCAGAGATATCCTAGGATCCGCCACCACCACAGTGGTGC
AACCTTCGCAGGTCCATCTAGCCTACAACCTAGGCAGTGCTATTCCCATAGCTGTAG
ACCCCAAGAACCTGGAACTAGGCTTTATTCTGAATTTGCCAATCATAGAAACTAGGA
ATATATATAAGTTAAAGTCTGTGTTGAATGTAGGATTTTGGCAAGACAACGCTCACA
TTCTACTTAAGACACCTCCAGTGCTAGCATACCATAGTGAAGACCCATCTCTCTACC
TCGTTCCAAATCTCAACCTCTGCACCAACACCAGGGACATCCATTGGGTCTGCCCAA
GTAACCCTTTCATCAGGGACACGACCAACTCCCTATGCGGCCTTAGGCCCGACGCCC
CCGAGCAAAAGTGCCAGGGCAGCCTGTCAGACAGGGACAGAGCAGCAGAACCTAG
GGTGGAACGAGCTGGCAACCGATGGCTACTCAGCACACCCGCCACCGAAGCCTTGA
TTGTCTATGACCACCATGACACAACCACTACGTTGGCCCTACCAAACCAAACCCTGT
TCTTAACGGTACCTCAGGGTGCTACCGTCCATATTGAGGACATCATCCTCCATTACCT
CAGTTCAGAGAGACATGACACTGAGATTGAGATCATGGACGTGTTCAAGGGTTACA

ATATCACGATTGATGACACCCTTCAGCGGCAACTCATGCTAGAAGGACCAAATTG
GTAAAGTTTAGCCTCAGGCAGACGGGGCTGACCACAACCTACGCGGATCAGGTGAC
TCGCCCGGTAACCTACAGCGAACACCCAATCAGCATGGTCGCGTTGGGACTCCTTCT
TAGTGGTTGGGTCATTACAGCAATGGTAGCATACGGAATGTACAAATATATCAAGG
AATTACAAGCTAGGTTAGATTCTGCCCTGGTAGTGCCAGTACGCTTTCCCCCACAGC
TTAACGCACCTACATTCAACTTAGCCGCCAGGACCTAA
> Loki-Bsp consensus sequence Env protein
LLQMKLPWAIIILNIQWALSIEITEPGPPAGIVLQNVPGLLITHCRLYTQRIYVRLDPWDA
YRNHIHLPPRLTEGRPSGTQTHETVEHTRSTVIHVLEQLQKFLVTEADLSGKGRPKRFLG
GLIAAASAIGSLFSIGLSAANSVSLSTLQRHVTELDEEMPEVHQKLSFQQQELQSLGKTV
QGTITTVNLHSALLNNTLHALETLSGIVKEDLTYVRVVRDLMQELIREVSSSINSLSSGRI
PTYLVNLDLVRDILGSATTTVVQPSQVHLAYNLGSAIPIAVDPKNLELGFILNLPIIETRNI
YKLKSVLNVGFWQDNAHILLKTPPVLAYHSEDPSLYLVPNLNLCTNTRDIHWVCPSNPF
IRDTTNSLCGLRPDAPEQKCQGSLSDRDRAAEPRVERAGNRWLLSTPATEALIVYDHHD
TTTTLALPNQTLFLTVPQGATVHIEDIILHYLSSERHDTEIEIMDVFKGYNITIDDTLQRQL
MLEGTKLVKFSLRQTGLTTTYADQVTRPVTYSEHPISMVALGLLLSGWVITAMVAYGM
YKYIKELQARLDSALVVPVRFPPQLNAPTFNLAART*
>Loki-Aro consensus sequence
TTGGCCTCTGCGACGACAACGAACACAAACGCTTTCGACATGACAAAGAAAACAGG
AATTACCCTTTCTACTCCAGACTCAATGATCAAAATATGTTCCTTAACCCGCAGAGG
CCCTGCACACAGACCCCATGCCATGCTGCAGTTCGACCAGACTTTCAAGGTAGAGGC
CCTTCTGGACACTGGCGCAGAAATCTCCCTCATGACTTACTCACTCTTTGAAGCATTA
CGAGAAACAAGGGAATGCACAGGTCACACCCTAAGCCTCAGTGAAAGCAACTGGA
ACTTTACAGCCTATGGCAGTACTAAGCCACAAAAACTAAAGGGAGTAATTGAGCTG
AAACTAATTTTTGAAGGATTGTCCATAACACACCCCCTCTATGTGGCATCACATGAC
CGTCCCATGCTTTTGATAGGCATGGACCTACTCAACCGCCTCAGGCCTACCATTGAC
TGTGATAACAACCTACTCTGGGGTCACAAACCAAAGGCTGACCACTATGTTCCTGGG
CCCACAAAGGCTATTTGTGCTTGCACACTAGGTTGGGTCTCCTCCGACAACCATGAA
ACGGGACGTGCACGCACCCCTACCCCACCAAACACGTTCCAACTACCCCAGGCTCC
ACTTGTCCCTAATAGCAAGACCACTCCGACAATACAGGAAAACCTGTATATATTCAG
GCCAGAGATTAACCCCTCCGTCCTTCCAGACGCAGGATCTGACCTGCCAAAGATGCA
TCACAATGCACAAATGGAAGAACTCCCTAACCCACCTGCTCGTACCAACCCCTATGC
CAAAGCTCAAAGCACTGACCACCCTCCCTATGGCCGACCCCCACTAATTCCACAGAC
AAACCACAGATCTCTCCAAATAGCGGCTAGCACAAACACTCCTAAAGCCGTTATGC
AAATAAACGCGAACCAGCCCACCACAAGTGTCTCACCACTAACCACTAATGGTACC
ACCCCTTCTCCTGAACCACACCCTAACTTGGACCCAAACCTCCTGCAACAAATCCAA
AAACAAGTGGATACTGCCGACACCTTAAACACAGATGAAGANATGACTCGCCCTGA
AAACCCTCATGCTCAGCTGTACTCCCAATCTGCGGCAAAACTGNACTGCAGCAAAA
CTGCTCTCCATGAAACCATAATTCCCACATTACCCAATGCTAACCCAGTGTATATCA
GAAAATATCGTATTCCAATAGCCGCACATGACCCTGTACAGGACACCATTACGGAA
ATCCTACATGCAGGAATTATCAGGAAATGCAATTCCACATACAACAGCCCTATTTGG
CCTGTACTCAAACCCTCTGGGAAACAGAGCCTCACAGTTGACTACAGAAAACTTAAT
AAACAAGTTCCACTCTCGCGCTGCCTATGGCCTACCTGGATGAAACCCTTGCAAAAG
TCCAGAATGCCACATACTTCACAAGCCTNAGACATATCAAATGGCTTCTGTACAATT
CCAGTCCGTGAACAAGATCAACATAAACTGGCTTTTTCATTCAACAATGTCCAATAC
ACTTGGACAGTCACCCCTTTTGGCTATGCAAACTCCCCAGCCGAATTCAACATCTTC

TTCCACAAAGCAATCCCTGACGTCAACTCCAAGAACATCGCCATTTATGTGGATGAC
ATCCTCATGTGGAGTGACTCTTGGGGGGAACACCAACAGCTGCTCTCCTATGTCCTT
ACAAGACTCAGTCAAGCAGGCGCCAAAATATCATTGGACAAGGGCCAGTGGTGTAA
ACGTGCTGTCAATTACCTGGGCTTCCAAGTCGGAGGTGATGGCATCAGACCCCGCAC
CGAGGCACCCAAAGCAATTCCTGCACCCACTACTCTGAAGCAGCTGTGCAGCTTCTT
AGGCATTTGCAACTTCTCCCAACATTTTGTGGCGGACTATGCCAAAATTTCCCAGCC
CCTCACACGGCTCCTAAAGAAAGACGTTCCATGGACATGGGGCCACAAACAAGATG
CACCAATGAAAATACTCAAGGACTCCATCTCCTCGGCACCAGCCTTAGCCTTCCCAG
ACAAAACCAAGCCTTTCATACTAGAAACCGCCTTCACACCCCTCAACATCACCGCAG
ACCTCAAACAACAACAGGACCAAAAAAACAAGGTTGTCTCTTATGCCAGCAAATCT
CTCTCCCCAGTAGAACAAAGGTTTACTGACTGTGAAAAGGCCCTGTTTTCAGCCATA
TGGGCCATCCAGCGATACCGCAACTTAGTCTGTGGAGCACAAATCATCATTGAAACC
TGTCATCAAATACTCACCTACCTCACATGTGGGAAACTGAGAAAGGGAAAGGTTAC
CAATGACCGCATAGCTGGTTGGATCCTAGCCACAGAAGGGCTTTCCTTACAGGTCAA
ATATTCCAAAAGCAAAAACAATGTGCTGGCACATGGCCTTGCCACATTACACGATTG
TGCATCATCACCGCAACAGACTGGTAAAGAATGCTTCTATACAGACTCACCCACACT
CCTTAACCACAAAAGTTTCTCTGATGCCCCTTGTACCACTTTCCCCATGGTCTACATG
GACGGCTCTGCCCACCGCAAAACGGGCACACCCGGTGCTGGCGTTGGTTTACTCTGG
ACAAATTGTGATTTGCCTCCACAAAAATTCAAGCTGGGACCTAAATCCAGCCAGTAT
GCTGAACTAGCGGGCACACTAATTGCCCTCNACGCAGCACAAAACAACTACCATGA
TCTTGTCATTTGCACTGACTCAGACTACACGGTACGTAGTTTCCTGGACTTTTTGCCA
AACTGGCTCGCAAACAATATGATTAAGTCTGATGGTAAACTAGTGGCCCACACTGA
ACTGTTCACGGCCGTTCAACACCTACTGGACACACACAACCTCACAGTATACTGGAA
AAAGGTCCCGGGTCATTCCAAACTAGACAGCCCAAACAAACACGGCAATGATCAGG
CTGACACCTTGGTAAAAGAGGGAGTATCTACAGGCCAACCCTGGGAATTCTGCCCA
CCAGTCATGCAGTTCCCGCAAACCCCCTGCTCAGGTGGACATGGTGACCCGCAGCCG
TGCCAAGGCGCCAGACCCTCACAGTCCCTTCAGCAGTACCACAGCTACCTGATCACC
CCAGACACACACAATGACCTGGTACGAGCACAGGAAGAGGATCAGATTCTGATACA
GTTCAAAAAAATCCTTCAAAACACAGCCTCAGTCGCCGACAGTGGGAACCTCAAAG
CGGACCCATACTTCCAGATACTCTGGAGAGAAAAGGAAAGGCTGTCCCTGCGCGAC
GACCTACTGATCTTCAAAGGAACAAATGTGACAGGTGCTGGTGTTTTGGTCCTCAAG
TANGGGCAAGGTCATTTTTCAGTTTTTTTACAGATTATGAAAGTGTATTATAAGTGG
ATTATAAGATTTCAAATGATGTGTCATACATTGAAATCTGAAAATAATTGNAGATTG
CGGACCTTTTCTCCAGCACCCGTCACAAATGATCAACACCATCGCTGGGTAGTTCCA
CGACAATGGAGAGCCACCATGATCCGACATGCCCATGATGAGCCATGTGGGGGGCA
CAGGGGAGAACAACCCACTATCCACATGCTATGGCAGGTGGCATACTGGTTCAAAA
TATGGACTGATGTGAAGGACTATGTGAACGGCTGTACCATTTGTCATAAGTTCCAAC
CCGTGTTGCCGCGACACAGGGCCCCCCTCCAATCAGCCCCAAGCGCGCAACCGTGG
TCACACATTCAAATTGACTGGATAGGTCCGGTCACTCGCTCTACTTGGGGAAACTAN
TATGCCCTGACAATAACAGATAGATTCACCAAATGGACCAAGTGCCTGACTGCGCTG
GCAGACACTGCAGAAACAACAGCCGTCCTCCTCATGAAGCACGTCTTCACTCACTGG
GGCCTGCCACAGAGCGTGGATTCTGATGGAGGCACACACTTCACTGCCACCATCATG
GAAATCGAGGTGTGGAAAACACTGGGAGTAAAAAGACAATTGCACATAGCGTACCA
ACCCCAGTCATCCGGGCAAGTTGAGCGAGCTAACCAGACGGTCATCAACATGCTAA
AAAAACAACTGGATGCAAATCAGAAAGACTGGGACATCAAGCTGCCACTCGTCCTT
ATGACCATANGGGCTATGCCTTCAAAAGCTACGCAGGTGTCCCTATTTGAAATGATG

ACTGATCGGCAGATAGTCCTACCCCTCCACCTCCTCTACACTACCCCAGTGGAAGAC
ATCCATGAAGCAATGCACACAGCCAACTACCTAGAGGAACTAGGCAGACACCTGCA
GGACATTTTCAAGTTTTCCAGACAGAACCTGAAAACAGAGGCAACTGGGGTCAAAG
CCTATTATGACAAAAAGGCCTGCGACCGCGAGTATGAAGTGGGGGACAAAATTTTC
TACTACCACTTCATGATCAAAAACGATCCAAAAAGCTTCTTCAGAGTTGGCAAGGAC
CGTACTTCATAACAGGCAAACTCTCCCCTGTTGTGTATCGTATCAAAGTTCAGACCC
CCCGTCGGACTGCCTTTANAGTGGGTCCACATTAATCAACTCCGACCCTACACCCCA
TCAGCACTCCAAACAGACACCGAAGAATGATTAAGCTCATGATGACAAGTCTCCCC
CTTCTGTACACCACAACCCTCCTGCTCGCCCTCACAATTCCCAGCCTTGGTCTCAAAA
TAACAGAGATTGTCCAACCTGGTCCCAAGACAAGAATCATCCTCAGAGACAACTCA
AGCTTCCTGGTAGCCGATGCCCAGACTGTTACTCATCAAATCTATGTCTCCCTCGAC
CCCCTTAATGTAATTTGGACGTACAAGAAAGAACCTCAACATGCCTACCGAGAAGA
GACAAAGACGTGGTACCTGTATCACCTAGCACATACCCGGGCTGCAGTCAAAACCA
CCCTGAGACAGTTAGAGCGAACGGCTGTATACATGGACACCCCGGACAGACCTCTG
CGGTCCAAAAGAGCAGTTTTCGGGGTCCTCCTATCCCTCTCCGCTGTTGGCAGCCTC
TTCTCATCCTCCATGTCTGTAGCTAACACTGTCTCTTACAACGCCTTGAAGAAAAAC
ATGGAAACTTTACAATGTGACCTGAAACACATAAAACTACTGATACAACAGCAACA
GTCCGCCCTTGAGGAAATGGGTACCACCCTAGGGGATACCATCCTGCTGCTCAACCT
ACACAACAAGCTCCTCAACTCGACTATCACCCAGTTACAGATCCTGACAAATAGTCT
TGTGGGAGCAAGTTACATTTATATATGTTTCAGTTTGATAAATAAGCGCAGTCAGTA
TTTTGTTTAA
> Loki-Aro consensus sequence *pol* gene
TTGGCCTCTGCGACGACAACGAACACAAACGCTTTCGACATGACAAAGAAAACAGG
AATTACCCTTTCTACTCCAGACTCAATGATCAAAATATGTTCCTTAACCCGCAGAGG
CCCTGCACACAGACCCCATGCCATGCTGCAGTTCGACCAGACTTTCAAGGTAGAGGC
CCTTCTGGACACTGGCGCAGAAATCTCCCTCATGACTTACTCACTCTTTGAAGCATTA
CGAGAAACAAGGGAATGCACAGGTCACACCCTAAGCCTCAGTGAAAGCAACTGGA
ACTTTACAGCCTATGGCAGTACTAAGCCACAAAAACTAAAGGGAGTAATTGAGCTG
AAACTAATTTTTGAAGGATTGTCCATAACACACCCCCTCTATGTGGCATCACATGAC
CGTCCCATGCTTTTGATAGGCATGGACCTACTCAACCGCCTCAGGCCTACCATTGAC
TGTGATAACAACCTACTCTGGGGTCACAAACCAAAGGCTGACCACTATGTTCCTGGG
CCCACAAAGGCTATTTGTGCTTGCACACTAGGTTGGGTCTCCTCCGACAACCATGAA
ACGGGACGTGCACGCACCCCTACCCCACCAAACACGTTCCAACTACCCCAGGCTCC
ACTTGTCCCTAATAGCAAGACCACTCCGACAATACAGGAAAACCTGTATATATTCAG
GCCAGAGATTAACCCCTCCGTCCTTCCAGACGCAGGATCTGACCTGCCAAAGATGCA
TCACAATGCACAAATGGAAGAACTCCCTAACCCACCTGCTCGTACCAACCCCTATGC
CAAAGCTCAAAGCACTGACCACCCTCCCTATGGCCGACCCCCACTAATTCCACAGAC
AAACCACAGATCTCTCCAAATAGCGGCTAGCACAAACACTCCTAAAGCCGTTATGC
AAATAAACGCGAACCAGCCCACCACAAGTGTCTCACCACTAACCACTAATGGTACC
ACCCCTTCTCCTGAACCACACCCTAACTTGGACCCAAACCTCCTGCAACAAATCCAA
AAACAAGTGGATACTGCCGACACCTTAAACACAGATGAAGANATGACTCGCCCTGA
AAACCCTCATGCTCAGCTGTACTCCCAATCTGCGGCAAAACTGNACTGCAGCAAAA
CTGCTCTCCATGAAACCATAATTCCCACATTACCCAATGCTAACCCAGTGTATATCA
GAAAATATCGTATTCCAATAGCCGCACATGACCCTGTACAGGACACCATTACGGAA
ATCCTACATGCAGGAATTATCAGGAAATGCAATTCCACATACAACAGCCCTATTTGG
CCTGTACTCAAACCCTCTGGGAAACAGAGCCTCACAGTTGACTACAGAAAACTTAAT

AAACAAGTTCCACTCTCGCGCTGCCTATGGCCTACCTGGATGAAACCCTTGCAAAAG
TCCAGAATGCCACATACTTCACAAGCCTNAGACATATCAAATGGCTTCTGTACAATT
CCAGTCCGTGAACAAGATCAACATAAACTGGCTTTTTCATTCAACAATGTCCAATAC
ACTTGGACAGTCACCCCTTTTGGCTATGCAAACTCCCCAGCCGAATTCAACATCTTC
TTCCACAAAGCAATCCCTGACGTCAACTCCAAGAACATCGCCATTTATGTGGATGAC
ATCCTCATGTGGAGTGACTCTTGGGGGGAACACCAACAGCTGCTCTCCTATGTCCTT
ACAAGACTCAGTCAAGCAGGCGCCAAAATATCATTGGACAAGGGCCAGTGGTGTAA
ACGTGCTGTCAATTACCTGGGCTTCCAAGTCGGAGGTGATGGCATCAGACCCCGCAC
CGAGGCACCCAAAGCAATTCCTGCACCCACTACTCTGAAGCAGCTGTGCAGCTTCTT
AGGCATTTGCAACTTCTCCCAACATTTTGTGGCGGACTATGCCAAAATTTCCCAGCC
CCTCACACGGCTCCTAAAGAAAGACGTTCCATGGACATGGGGCCACAAACAAGATG
CACCAATGAAAATACTCAAGGACTCCATCTCCTCGGCACCAGCCTTAGCCTTCCCAG
ACAAAACCAAGCCTTTCATACTAGAAACCGCCTTCACACCCCTCAACATCACCGCAG
ACCTCAAACAACAACAGGACCAAAAAAACAAGGTTGTCTCTTATGCCAGCAAATCT
CTCTCCCCAGTAGAACAAAGGTTTACTGACTGTGAAAAGGCCCTGTTTTCAGCCATA
TGGGCCATCCAGCGATACCGCAACTTAGTCTGTGGAGCACAAATCATCATTGAAACC
TGTCATCAAATACTCACCTACCTCACATGTGGGAAACTGAGAAAGGGAAAGGTTAC
CAATGACCGCATAGCTGGTTGGATCCTAGCCACAGAAGGGCTTTCCTTACAGGTCAA
ATATTCCAAAAGCAAAAACAATGTGCTGGCACATGGCCTTGCCACATTACACGATTG
TGCATCATCACCGCAACAGACTGGTAAAGAATGCTTCTATACAGACTCACCCACACT
CCTTAACCACAAAAGTTTCTCTGATGCCCCTTGTACCACTTTCCCCATGGTCTACATG
GACGGCTCTGCCCACCGCAAAACGGGCACACCCGGTGCTGGCGTTGGTTTACTCTGG
ACAAATTGTGATTTGCCTCCACAAAAATTCAAGCTGGGACCTAAATCCAGCCAGTAT
GCTGAACTAGCGGGCACACTAATTGCCCTCNACGCAGCACAAAACAACTACCATGA
TCTTGTCATTTGCACTGACTCAGACTACACGGTACGTAGTTTCCTGGACTTTTTGCCA
AACTGGCTCGCAAACAATATGATTAAGTCTGATGGTAAACTAGTGGCCCACACTGA
ACTGTTCACGGCCGTTCAACACCTACTGGACACACACAACCTCACAGTATACTGGAA
AAAGGTCCCGGGTCATTCCAAACTAGACAGCCCAAACAAACACGGCAATGATCAGG
CTGACACCTTGGTAAAAGAGGGAGTATCTACAGGCCAACCCTGGGAATTCTGCCCA
CCAGTCATGCAGTTCCCGCAAACCCCCTGCTCAGGTGGACATGGTGACCCGCAGCCG
TGCCAAGGCGCCAGACCCTCACAGTCCCTTCAGCAGTACCACAGCTACCTGATCACC
CCAGACACACACAATGACCTGGTACGAGCACAGGAAGAGGATCAGATTCTGATACA
GTTCAAAAAAATCCTTCAAAACACAGCCTCAGTCGCCGACAGTGGGAACCTCAAAG
CGGACCCATACTTCCAGATACTCTGGAGAGAAAAGGAAAGGCTGTCCCTGCGCGAC
GACCTACTGATCTTCAAAGGAACAAATGTGACAGGTGCTGGTGTTTTGGTCCTCAAG
TANGGGCAAGGTCATTTTTCAGTTTTTTTACAGATTATGAAAGTGTATTATAAGTGG
ATTATAAGATTTCAAATGATGTGTCATACATTGAAATCTGAAAATAATTGNAGATTG
CGGACCTTTTCTCCAGCACCCGTCACAAATGATCAACACCATCGCTGGGTAGTTCCA
CGACAATGGAGAGCCACCATGATCCGACATGCCCATGATGAGCCATGTGGGGGGCA
CAGGGGAGAACAACCCACTATCCACATGCTATGGCAGGTGGCATACTGGTTCAAAA
TATGGACTGATGTGAAGGACTATGTGAACGGCTGTACCATTTGTCATAAGTTCCAAC
CCGTGTTGCCGCGACACAGGGCCCCCCTCCAATCAGCCCCAAGCGCGCAACCGTGG
TCACACATTCAAATTGACTGGATAGGTCCGGTCACTCGCTCTACTTGGGGAAACTAN
TATGCCCTGACAATAACAGATAGATTCACCAAATGGACCAAGTGCCTGACTGCGCTG
GCAGACACTGCAGAAACAACAGCCGTCCTCCTCATGAAGCACGTCTTCACTCACTGG
GGCCTGCCACAGAGCGTGGATTCTGATGGAGGCACACACTTCACTGCCACCATCATG

GAAATCGAGGTGTGGAAAACACTGGGAGTAAAAAGACAATTGCACATAGCGTACCA
ACCCCAGTCATCCGGGCAAGTTGAGCGAGCTAACCAGACGGTCATCAACATGCTAA
AAAAACAACTGGATGCAAATCAGAAAGACTGGGACATCAAGCTGCCACTCGTCCTT
ATGACCATANGGGCTATGCCTTCAAAAGCTACGCAGGTGTCCCTATTTGAAATGATG
ACTGATCGGCAGATAGTCCTACCCCTCCACCTCCTCTACACTACCCCAGTGGAAGAC
ATCCATGAAGCAATGCACACAGCCAACTACCTAGAGGAACTAGGCAGACACCTGCA
GGACATTTTCAAGTTTTCCAGACAGAACCTGAAAACAGAGGCAACTGGGGTCAAAG
CCTATTATGACAAAAAGGCCTGCGACCGCGAGTATGAAGTGGGGGACAAAATTTTC
TACTACCACTTCATGATCAAAAACGATCCAAAAAGCTTCTTCAGAGTTGGCAAGGAC
CGTACTTCATAA
> Loki-Aro consensus sequence Pol protein
LASATTTNTNAFDMTKKTGITLSTPDSMIKICSLTRRGPAHRPHAMLQFDQTFKVEALLD
TGAEISLMTYSLFEALRETRECTGHTLSLSESNWNFTAYGSTKPQKLKGVIELKLIFEGLS
ITHPLYVASHDRPMLLIGMDLLNRLRPTIDCDNNLLWGHKPKADHYVPGPTKAICACTL
GWVSSDNHETGRARTPTPPNTFQLPQAPLVPNSKTTPTIQENLYIFRPEINPSVLPDAGSD
LPKMHHNAQMEELPNPPARTNPYAKAQSTDHPPYGRPPLIPQTNHRSLQIAASTNTPKA
VMQINANQPTTSVSPLTTNGTTPSPEPHPNLDPNLLQQIQKQVDTADTLNTDEXMTRPE
NPHAQLYSQSAAKLXCSKTALHETIIPTLPNANPVYIRKYRIPIAAHDPVQDTITEILHAGII
RKCNSTYNSPIWPVLKPSGKQSLTVDYRKLNKQVPLSRCLWPTWMKPLQKSRMPHTSQ
AXDISNGFCTIPVREQDQHKLAFSFNNVQYTWTVTPFGYANSPAEFNIFFHKAIPDVNSK
NIAIYVDDILMWSDSWGEHQQLLSYVLTRLSQAGAKISLDKGQWCKRAVNYLGFQVGG
DGIRPRTEAPKAIPAPTTLKQLCSFLGICNFSQHFVADYAKISQPLTRLLKKDVPWTWGH
KQDAPMKILKDSISSAPALAFPDKTKPFILETAFTPLNITADLKQQQDQKNKVVSYASKS
LSPVEQRFTDCEKALFSAIWAIQRYRNLVCGAQIIIETCHQILTYLTCGKLRKGKVTNDRI
AGWILATEGLSLQVKYSKSKNNVLAHGLATLHDCASSPQQTGKECFYTDSPTLLNHKSF
SDAPCTTFPMVYMDGSAHRKTGTPGAGVGLLWTNCDLPPQKFKLGPKSSQYAELAGTL
IALXAAQNNYHDLVICTDSDYTVRSFLDFLPNWLANNMIKSDGKLVAHTELFTAVQHLL
DTHNLTVVYWKKVPGHSKLDSPNKHGNDQADTLVKEGVSTGQPWEFCPPVMQFPQTPC
SGGHGDPQPCQGARPSQSLQQYHSYLITPDTHNDLVRAQEEDQILIQFKKILQNTASVAD
SGNLKADPYFQILWREKERLSLRDDLLIFKGTNVTGAGVLVLKXGQGHFSVFLQIMKVY
YKWIIRFQMMCHTLKSENNXRLRTFSPAPVTNDQHHRWVVPRQWRATMIRHAHDEPC
GGHRGEQPTIHMLWQVAYWFKIWTDVKDYVNGCTICHKFQPVLPRHRAPLQSAPSAQP
WSHIQIDWIGPVTRSTWGNXYALTITDRFTKWTKCLTALADTAETTAVLLMKHVFTHW
GLPQSVDSDGGTHFTATIMEIEVWKTLGVKRQLHIAYQPQSSGQVERANQTVINMLKKQ
LDANQKDWDIKLPLVLMTIXAMPSKATQVSLFEMMTDRQIVLPLHLLYTTPVEDIHEAM
HTANYLEELGRHLQDIFKFSRQNLKTEATGVKAYYDKKACDREYEVGDKIFYYHFMIK
NDPKSFFRVGKDRTS*
> Loki-Aro consensus sequence *env* gene
ATGATTAAGCTCATGATGACAAGTCTCCCCCTTCTGTACACCACAACCCTCCTGCTC
GCCCTCACAATTCCCAGCCTTGGTCTCAAAATAACAGAGATTGTCCAACCTGGTCCC
AAGACAAGAATCATCCTCAGAGACAACTCAAGCTTCCTGGTAGCCGATGCCCAGAC
TGTTACTCATCAAATCTATGTCTCCCTCGACCCCCTTAATGTAATTTGGACGTACAAG
AAAGAACCTCAACATGCCTACCGAGAAGAGACAAAGACGTGGTACCTGTATCACCT
AGCACATACCCGGGCTGCAGTCAAAACCACCCTGAGACAGTTAGAGCGAACGGCTG
TATACATGGACACCCCGGACAGACCTCTGCGGTCCAAAAGAGCAGTTTTCGGGGTC
CTCCTATCCCTCTCCGCTGTTGGCAGCCTCTTCTCATCCTCCATGTCTGTAGCTAACA

CTGTCTCTTACAACGCCTTGAAGAAAAACATGGAAACTTTACAATGTGACCTGAAAC
ACATAAAACTACTGATACAACAGCAACAGTCCGCCCTTGAGGAAATGGGTACCACC
CTAGGGGATACCATCCTGCTGCTCAACCTACACAACAAGCTCCTCAACTCGACTATC
ACCCAGTTACAGATCCTGACAAATAGTCTTGTGGGAGCAAGTTACATTTATATATGT
TTCAGTTTGATAAATAAGCGCAGTCAGTATTTTGTTTAA
> Loki-Aro consensus sequence Env protein
MIKLMMTSLPLLYTTTLLLALTIPSLGLKITEIVQPGPKTRIILRDNSSFLVADAQTVTHQI
YVSLDPLNVIWTYKKEPQHAYREETKTWYLYHLAHTRAAVKTTLRQLERTAVYMDTP
DRPLRSKRAVFGVLLSLSAVGSLFSSSMSVANTVSYNALKKNMETLQCDLKHIKLLIQQ
QQSALEEMGTTLGDTILLLNLHNKLLNSTITQLQILTNSLVGASYIYICFSLINKRSQYFV*
>Loki-Oja consensus sequence
ATGACTGACGCAGTTTCTTCTGAGGCCTCTAGTCATGGGGACCTGTTTGATGATCTA
GAAGCTGCTTTGACAGAAATCACTGAAAATCAAATTCCAGGTTTTGAAGGTCGAGTT
CGGAACTTAACCGCACCTGAACTGGAGACTAAGCTTTCTGAGATAGTGGCTGAAATT
ACTAATGATCCACCGAAGGATAAGGAGTTGACCAAAACACTCGGTAGCCTAGCCAT
ACTGCTCACAGCTCGACTCAAAAGCGCTCACACCCTATTGAATCAAGCACTAGGGA
AACTTGACACACGCTCGCAGAAACTTTCTCAAACCCAGGGAGAAATTAAGCTTTTGC
AGCATCAGCGCGAGGTTGCCCAAAGACGAACAGATGAGTTGGAGTGTCAGAACGAA
GAACTAAAAACTGAAAATGATCAACTGAAAAGAGAAATAAAAGATTGTGAATCGCA
GACCCAGGCCTGTCAGCATGAACTAACCGCAATGCATGTGCAGCTCATTGAAGTTAA
AGAAACCTACAAAGCCTCTCAAATAGGACTTTACTCAGCTGAAGATGATATAATTAA
GTTGCAGCAAGAGAATGCTAATTTGACCAGACAGTTAGGTGACCAAAAAGAACAAG
TTGAGCTAATGCGAGAGCAACTGGCCCACCTTACCTCAAAGCTTGAGGCCTCTGACC
AACGGTATGAACGCTTGATGGTGGAACGAAAGTCTGTGACAACAGACATCAAACCA
CCCATCGAATCTGACATCAGTCAGCAGTTTTATCCAAGTTTCAGAGACGTAAAGGGG
GGTGCACAGGCTGGAAGGGGTGCGTCTGTTGGATTACCTTTGGTCTCACCTGTACCT
GCCCCCCTTCAGGCTGAAAGGGGTGATTCTGGTGGATCACCTCGGTTGTCACCTGTA
CCTGCTCCTCTTTCTGCATTGACTTCAAAGCCGACACAGGCAAATCTATTGCATGACT
TTAGGGCAGAAGCCACTCGTCATCGTGCTCCTTCTGATCGAGATTTAGATAAAATTT
CTCGAAATCTTACCAAGTTCGAACCAACCCCTGGTGGCTCTAATGATCCTCGAGCAT
ACCTGGATGACATTGAGTTCTATCTCAGGAAATACTCGGAGGCCACCATTCAAGATA
AGATTTATCTTATCAAACTCACTTCGAGCCGTGAGGTAATCCGGTTCATCGAAAGGC
AGCCAGAGTTTACCAGACGCGAATGGTACTTTTTGCGACAAGCCATTGAGAGGGAA
TTCTCTGACTACACTATACAATCAGGTTTAACGGTTGCCATGTCAGTAAAGCAAGGT
AGGCAGGAGTCTCCACAGCAATACTATTACCGACTTTTGCGAGCATACTTTGGTACT
CGTAATGAGCCAGGAATGGAGGAGGACTTGAACTTCAAATCTCTTTTCCTTCAGAAC
CTCCATCCAGGAACCAGTCATCACTTGGGAGTCTTTGCAAACCCCCACACAGCAACA
ATAAGTCACCTTCGTGAACTTGCTGCGTTGGGTTTCCAGAAGCAGAAACAGGTACAC
CAAAAGTCAGGCGACCCCATTGTTCTGTCTACCACAGCAGATCAATCAACTTTGGAA
CTTGAAGGTGCACCAAACCACAAGTTGTATGGACCTGGTTCCGTTGTATCTCATCCA
CCCAGAGCCAAAGGCGACCCAAACGGGGTTCCGTCAAGGACACCCAACCAAAGTGA
GTCACGGAAACATTGGTCTGGGAAGAAAACACAACAGGAGGGGCGAAAACCAGAT
GACACATCTGGGCCTAGGTCTAGCTGGCATAGAGGGAAACCTCAGCCAAAATGGTT
TGACAATACCCCTACCTACGAGAGACACAGGGATCATGGGAGACGACCACAAAACA
GGTTTGGAGGCCAGTATGCTGGACCCCCACGTGATTTCCATGAGGGAAAAGAACCC
ACAGAAAATTTTTACTTAAAGAAATCTTGGCCCAAATCAGAAAAGGAAAGGGAATT

TTCAGGGTTAAGTAAAGAAGAAATACAGGAAATCAAACGTCTACTAAAAGACTTAG
CACAGAAGCCTAAGCCAGATAAACCAGATGTGTTGGACCTCTCCCTGAGTCAATCC
AACTTTGAACTAAATGGTCCGGTCCCAACCTCCGGTGACGGACCCCAGACGAGCAC
AGATCGTCACACTGCCAATCTGGTCCCTGTGACGGCCTCAGAGGGGGGGTGCCTTGC
ATCCCCACCTATAAGCGATGAGGAAGAAAATGAGACCCTTCAATGCCCCAGTGCTG
TATTGGTGGTTAGACTGGACCCACAACAACAGTCACTTGAGCAGGGTGACTCCATGT
TTCCGGTCTACCCCCCAGTACAACAATTTCTGTGCGACCTCACAGAGAAGGACTCG
CACGAAAATTTTACATGTCACTAACCCTAGAAGATGTGTTAACACATGAGGCCTTAC
TTGACACTGCTGCAGACATCACATTGATCTCATCCACCTTGTTTAACCAGCTGCGCG
CTGCCGTTAACGATGACAATAAAGACCTGAAACTGCAACCATGCTCTGTTAAGGTGC
GTCCTTATGCCAACACCAACACCATCTTGTCCGCAGTGTCGCTTATCCAACTGTCTAT
TGGTCCAATGACAGTGGTGCATCCAGTCTATGTGTCACCCATCAATTCGATTCCTTTC
CTCGTGGGTAAGGATCTTCTTAATCGCTTTGAACCTTTGATCGATTTCCAACAGTTGA
AACTGTGGGCACAAGTTCGACAACCCCTAAGCATTTCTTGGCCGAAGGACGCTTCAT
CGGCTCACTGCTATGCTGTGAGAACTGAAGAGCTTCACCTCTCAGATGAACAGGACA
GTTCAACCAGGGAAGCAGAAATGTCAGAGCCAACAGAGAAGAAACCCCCTGCACA
AAACGAACAGGTGATGTTCCAGCATGACACCTTCTTGTGTGCAATTAGCCCGCCAGA
CTCGGAGGGAGAATACATTCCCCACATCCTTGGTGGCATCCAGATGCTGGGCCACAT
GGCTAGCGACGCCACCTTGTACCTCTGGGCTGACCAGTCCGCCATCAGTCACGAGTT
TTATCAAACTCTGAAAAAAGAAAACCCTGAACTTCCATTCGCAGGAAAATCATGGC
GTTTTCCAATCCATTCATGCTCGAACTTGACTTTGAAAGCCATAGGCGTGTGCTCACT
CGCGCTAAAATGGAACCACCAAACACTCACACATTCCTTCTTGGTGGTACCAGACCT
GCCACATACTGCCTGCATAGGCAGTGACCTCCTAGTCCGGCTAGGTGTACAGGTAGA
CACCATTAACAACGTCTTATGGTCCCCTGTACCGGCTCGAGAAACTCCAATGACCCT
CGACGCACTGAGGTCCGGACAGACCATTCCAGAAGCTTGTCAGGTGGTCACTTCGG
ACCATGTATCCTTACCTGCTGGTATTCGGGAAGCCATCATCAGAATGCACATTCAAC
CAGGCCAAAGTCTTGGTCATTCTCAAGCACTCTTCCAGCCGGCCCCTTTGTTTTTCAA
ACTTGGACTTTCCTTGCCAGCCACACCTCTTTTGAGTCTGAATTCCAGAACCATATCT
TTGAAGGTGAGTAATGACAACCACAATGAAGAAACTCTTGTGCCAAGAGGGACCAC
GTTGGGCTGGCTCATAAGTACAGATTTCCATGACTTTGAACTAACGGTCCCTGTTCT
GGGAGGCATTCCCCACTCCCTTCTTGGTGAAGAGGTGTCAGATGTCATCTACACAAA
ACCTTCCAAGGCTATCGCCATCTACTCTGTCATGGACCTGACATCCGATTCGGTGGG
ACGTGTGGACTTGTCAGACAAAAATCACTTGGTTTTCCAGACCATGAATGACGTATC
AATTAACACCATTGATGCACATCCCAACACTGAGACCCAAACTCAGCCTCCACCTTC
CATCAGTGAACAAATTGATGAGGTTGTGAAATTGGCTGATGCTTTAACTACTGACAA
TGAACGTCAACAGTTACGTCAGGTTCTGTTAAAGTATCAATCATCGTTGTCCACAGA
TTCCCTTGATTGTGGATTAACTTCTATCCACTCAGTGCGCATCCCAACACCAGAAAA
CGCACCGCCTACCTTTGTTAGACAATACAAGATTCCTCTTGCATCGTATGAACCGGT
TCAGGAGATCATTGATGATCTGACCAACAAAGGGATCATTAGGCCCTGTAATAGCA
CCTACTCAGCCCCAATTTGGCCTGTACGTAAGCCAAATGGCAAATGGAGATTGACTA
TCGACTACAGACGGTTGAACCAGCAGGTTCCCCTATCTAGATGGCCAATGACGCAG
CTGGAACAGGAACTACCCAGGGTCAAAGAAGCTAAATACTTCTCTACCCTGGATGT
AGCCTCAGGATTTTGGACAATCCCTGTCCATCCTGACGACCAGCACAAGCTAGCATT
CACCTTTGCGAGGCAACAATACACTTTTACCCGCTGTCCGTTTGGATATGCAAATTCT
CCAGCAGAGTTTAACATCTTTCTGAACAAAGCATGTCCCGATGCCACACAACGTGGT
ACACTTATTTACGTAGATGACATTCTCATTCGTAATGCAACATTGGAAGATCACATT

GAAGAGATTGAACATGTACTTAAACAACTCACCTCTGCTGGAGCCAAAATTTCTCTC
TCTAAATGTCAGTGGTGTCGAACCAAAGTCAACTATGTGGGCCTCCTGGTGGGACCC
ACCGGAGTCCAACCACAAGTTGGCAGAGTTCAAGGTCTCAAGGACATTGCAGCTCC
AACAAACCTCAAAGAATTAAGAAGCTTTCTAGGAGTATGCAATTACTCACGCCAGTT
TATTGAGGACTATGCTGACATAGCTAAACCGCTCACTGAACTACTAAAAAAGGATGT
GCCCTTCAGGTGGGCTCAACCACAACAACAAGCCATGGAAACTCTGATAACCAAAC
TGAGTTCAGCCCCATGCTTGGCCTACCCTAACTGTGAAAAAGAATTCCACCTTGAAG
TTGGATTCTCCAATAACTGTTTGAGCGCTGGGTTGTATCAAATATATGACCAAGATA
AACGCGTTGTTGCTTATGCTAGTAAGACACTCGCTACACCTGAACTGAAGTACTCGG
ACTGTGAAAAAGCGTTGCTTGCCACTGTTTGGGCTGTGAAGCATTTCTCTAACTATG
TTGGAGGACAAAAGGTAATTGTTGAAACCCAACATCAGCCGGTCACATTCCTGAAC
AGTCAGCGAATCAGAGACGGTGTGGTCACAAACGCACGTGTAGCTTCATGGCTAAT
GGCACTGCAGAGCTTTGACCTCGAAATCAGGTACGCACAAAACAAAAAGAATGCTT
TGGGTACTGACCTAGCTGCATGCCAGCGTTGTGTTGACGACGCTCCAACAGAAATCA
GAGTACAAAGTTCAGGGATCATCACTGATCCACACCACCACTACTTTGATCAAAATA
TCTGCCAAGATATGGTTACTGCATACGTGGATGGGTGTTCTTTCCGCAAGGATAAAG
AACTCCTGGCTGGTGTGGGTATTGCATGGCTTCACGACGACCCCTGTGAACCCAAGA
GTTTCCATCTTGGGTCTCAAACCTCACAATACGCAGAAATTGCTGCAGTTTTAATTGT
TTTACAACTTGCAGTTTCTTTGAACATCCATACACTTGTGATCTGTACCGACTCAAAC
TACGCACGACTGAGTTTTTCATGTCATCTGGACCTCTGGAAACAGAACAATTTCCAG
ACGTCCAACAGAAAACCAGTGAAAAACAAAGAGCTTTTTGTGGCATGCGATCACTT
GATCACCTCTCACAACATCCAAATCTACTGGAAAAAGGTGAGAGGACATTCTCGAA
TCCCTGGTCAGGACAAAGTGATGAATGACTTGAGTGACTCAATGGCAAAACAGGGT
GCATTGACGGGCGAACACTGGGACTTCGCTGAAACATGTCGGCTTTTTCCCCACCCA
CCCGCCCCTGAAGCTCAGATTTGCGCGGTCACCAGATCCACACCCTTGGCCTCCGTC
TCGGCCAAACCCATTTATTCTGACGATGACCTCATCAAACTGCAGGCAGACGATCCT
GCCATTGCCAAGGTCCTGGATTTTCTGGTCGATCCGACCAACAACGCCCCCTCGTCT
GGTGACTTTGAAACGATTCCAGAACTTCGGGCTCTTTATGCCCTTCATTCTTCCTTGA
AGGTGGTCAAGGGCATTCTTCTTCGTGTTTCCGGTGACCACGCTCCGCCGGCTCTCGT
GGTCCCTCGTAACCAAAGGGGGGTGATGTTGATGTATGCCCATGACCTCCCCTCTTC
AGGTCACAAAGGCATCAACGCAACATACTCTGCACTTAAACAGATCGCATTTTGGCC
ACACATGAAGAAAGATGTGACAGATTATGTCAATGGCTGCTTGGTTTGCTGTCAGTT
CCAACCAACTAACCCAACACACAGAGCACCACTGCAAAAGAAAGGTATTTCATTTC
CATGGTCAGACATCCAGATCGACTGGGTGGGCCCCTTGACAAAGTCATCCAGAGGA
AATAAATACTTCCTGACAGTGACATGTGCCTTCACAAAATGGGTGGAGTGTCTACCA
GCTCCAAATGACACTGCTCAAATGACCGCTCTGTTGCTTATGAACCACGTGTTCTCA
CGTTTTGGTTTACCAATGCGAGTGAATTCGGATCGTGGTACTCACTTCACAGCCGAA
GTGATGCAACAGGTCTGGAACACTCTGGGGGTAAAAGCGAACTTTCACGTGAGCCA
CCACCCTCAATCCTCAGGGCAGGTGGAACGAGCAAACAGAACTGTAGTAAACATAC
TGAAAAAGTATGTAGCGGCCAGTCACAAAGACTGGGACATCAAACTACCTCTGGTA
ATGATGGCCATCAGGGCCACCCCTCATGAATCCACTGGTACATCACCTTTCGAGATG
ATGACGGGGAGACAAATGACTTTGCCTCTACACCTTCTTTACCAGCCAGGAGAGGCT
AGTGTAGCAACAGCCTACACGGCGCATCAATACGTGGCTGACCTGCATGAACACTT
GAAGGCTACATTTGCTTTTGCGCAAAAGCATATGGAGAAAGCAGCTGAGGGTCGTA
AAACTTATTACGACCAGAAAGCGTCGCACCAAACGCTTGAGGTAGGAGACAAAGTC
TGGTACTACCTTTTTGTCCAGCCAACGGGACAGGCCAACCAGAAGGTTGGCAAACTG

GCCCGCAAACTTCTTCCACGCTGGACAGGCCCACATGTCATCACTGACAAGCTGTCG
CCAGTGGTATACCAGATCAAAATCAGCCAAGGCCGAAAGGAGCCCATTTTGAAATG
GGTCCATCGCAACCAGCTCAAACTGCACAAACACCCTGTGGGACTTGTTGGGTTCAA
CCAAACCCAATCAAATTAAACCACAGAGATTCCTTATTGTGGTGGTGTACCAGTTCA
TCCATAACTGAGCACCCACCCCCTCCAATAAGTGAAAGTAAGATTGTACAAATTGTG
AGCTCAAACAATAGCACGTCAATAGTACAAATCTTGCTCCACTGGAAATCGTTGTTT
GTAAAACTATTTGGCTTTGATCTCTTGGTGTGTTTTCTCTAAGATCCATAGTGTTACT
AGTTACTGGCTGACCAATCGAGTTACGTTGTGTTTCTTAACCACAAAGGGTAAGTTT
GCTAATACCAATCTACGTTGCCTCTCTGTCTTATCAGGAAGATGATACTGCTGGTGA
TCGCAGTGCTCCTCCTTGTCTCAAGCCAGGCACGTTTGGAGATTGTTGAACCGGGAC
CGGAGTCAGGAATTGTGCTGAAGGAACACCCCGGACTGTTAATCACGAACTGTCGC
TGGCACACCCAAAAGATCTTTGTGAGGCTCCGTCCTGATGCAGTTTACAGAAAACAT
GTGCAGGCGACATCCAGCTTTCTCTCTTGGGCTGCTGTTACCTGGACAATGGATGTA
ATTTCACATGCTGAGAATGATATCGTTCACATGCTAACTCAGATGCAAAAATTCACG
ATCACCCAGTCTGAGCTAAGTGGCGTTGGACACCGTGATAAACGTTTTATCGCTGGG
CTACTTACCGCTGCCTCAGCTATTGGTTCACTGTTTAGTCTGGGGGTTTCCGCTGTAA
ATGCAGTCACCCTAGCCACCCTCAAGCGTCACATGAATGAGCTACGTGATGAGATTC
CCCAAATCCGGGGTCAGCTAGATATCCAACAATCACAGGTACAGACTATGGGAAAA
ACTCTACAGGACACTGTGCTAATCGTGAACACCCACAGCGAGACTCTTAAAAACTTA
ACCAGGGCAGTAAACAAGTTGTTAAGCGTAGTTAGTATAGACTATGCACACATACA
ACTGGTAACCATGCTACTGGCTGACATGACTCGAGACATCGGGTCATCCATTGATTC
GCTAGTCCTCGGTAGAATACCCCCTTATCTAGTGTCACTAGAGTTGGTCCAACAACT
CTTATCTGAAGCAACCAGAGAACAGATAACCCCCCTACAGGCCCATCTGGCCTACA
CCTTGGGAAGTTCAATTCCCATCCATGTGGATCCGACGGCCAGGGAGCTAGCTTTCT
TGATCCACATCCCCATCGTCTCTCCAGAAAACATTTACAGACTCAAAGACGTCATAA
ATGTAGGGATCTGGCAAGGTGACGTTCATGTGAGGATGCACACACCCAAGGTCATC
GCGTACCATGACAGCGATCCAGAGTTGTACCTAGCGCCAAACTTACAGTTGTGTACG
CCCACTAGAGATATTCACTACCTGTGCCCAAGCAAACCGTTCGTGCGAGATAGCACT
GGTGGTCTGTGTGGACTTAAGCCAATCCACACAGATACTCGCTGCCCCACTACCATT
ACCCCTCGCAGGCAAATGACCACCACGCAGGCGGAGATCGTTGGAAATCGGTGGTT
GGTTCTGACACCAGTAACAACCGCGACCATCATGTATGATCAACATGACACCTCAAC
CCGAATGGAACTACCATGGAGTTCCATCTGGGTTGACGTCCCTAACGATGCTATGGT
ACACGTGGACAACCTAGCCCTCTATCACCTTCACCCAGATCAATATCAAGGTGAAGT
AGAGATTTCGACTTTCTTTAGCCAACATACGATCGAACTCAACTCTCGCACAGTGGA
GCAAGTCGACTACGAAGGGTCGCGTGTAATTGATGTAGGCCCCATTAATAAGGTCCT
CAAAGAAATGGAAGCACGTCAGAACCCTAATATTCAATCCCTTACTTACGCTTGGTC
CACCGCAGATACTGTATTAGTAACTCTGGTAGGACTCGGATACCTTCTCACTTTTGG
AACCACCTATCTTTATGTTAAGAAAACCCAACATCTTTGTAGTCAGATAGACAGGTG
TCGCAGTCGATTGGCTCGAGGGACTCGACGGAGATTTGCTGAAAGAGAGAACCATG
CTAACGACAGGACAGAGGGGTCCCTCGAGATGAATCTTAGAACAGAACCTCTCTCT
CCAACCATTCTCGACCTGGGCGATACTCTGTGA
> Loki-Oja consensus sequence *smc-gag-pol* gene
ATGACTGACGCAGTTTCTTCTGAGGCCTCTAGTCATGGGGACCTGTTTGATGATCTA
GAAGCTGCTTTGACAGAAATCACTGAAAATCAAATTCCAGGTTTTGAAGGTCGAGTT
CGGAACTTAACCGCACCTGAACTGGAGACTAAGCTTTCTGAGATAGTGGCTGAAATT
ACTAATGATCCACCGAAGGATAAGGAGTTGACCAAAACACTCGGTAGCCTAGCCAT

ACTGCTCACAGCTCGACTCAAAAGCGCTCACACCCTATTGAATCAAGCACTAGGGA
AACTTGACACACGCTCGCAGAAACTTTCTCAAACCCAGGGAGAAATTAAGCTTTTGC
AGCATCAGCGCGAGGTTGCCCAAAGACGAACAGATGAGTTGGAGTGTCAGAACGAA
GAACTAAAAACTGAAAATGATCAACTGAAAAGAGAAATAAAAGATTGTGAATCGCA
GACCCAGGCCTGTCAGCATGAACTAACCGCAATGCATGTGCAGCTCATTGAAGTTAA
AGAAACCTACAAAGCCTCTCAAATAGGACTTTACTCAGCTGAAGATGATATAATTAA
GTTGCAGCAAGAGAATGCTAATTTGACCAGACAGTTAGGTGACCAAAAAGAACAAG
TTGAGCTAATGCGAGAGCAACTGGCCCACCTTACCTCAAAGCTTGAGGCCTCTGACC
AACGGTATGAACGCTTGATGGTGGAACGAAAGTCTGTGACAACAGACATCAAACCA
CCCATCGAATCTGACATCAGTCAGCAGTTTTATCCAAGTTTCAGAGACGTAAAGGGG
GGTGCACAGGCTGGAAGGGGTGCGTCTGTTGGATTACCTTTGGTCTCACCTGTACCT
GCCCCCCTTCAGGCTGAAAGGGGTGATTCTGGTGGATCACCTCGGTTGTCACCTGTA
CCTGCTCCTCTTTCTGCATTGACTTCAAAGCCGACACAGGCAAATCTATTGCATGACT
TTAGGGCAGAAGCCACTCGTCATCGTGCTCCTTCTGATCGAGATTTAGATAAAATTT
CTCGAAATCTTACCAAGTTCGAACCAACCCCTGGTGGCTCTAATGATCCTCGAGCAT
ACCTGGATGACATTGAGTTCTATCTCAGGAAATACTCGGAGGCCACCATTCAAGATA
AGATTTATCTTATCAAACTCACTTCGAGCCGTGAGGTAATCCGGTTCATCGAAAGGC
AGCCAGAGTTTACCAGACGCGAATGGTACTTTTTGCGACAAGCCATTGAGAGGGAA
TTCTCTGACTACACTATACAATCAGGTTTAACGGTTGCCATGTCAGTAAAGCAAGGT
AGGCAGGAGTCTCCACAGCAATACTATTACCGACTTTTGCGAGCATACTTTGGTACT
CGTAATGAGCCAGGAATGGAGGAGGACTTGAACTTCAAATCTCTTTTCCTTCAGAAC
CTCCATCCAGGAACCAGTCATCACTTGGGAGTCTTTGCAAACCCCCACACAGCAACA
ATAAGTCACCTTCGTGAACTTGCTGCGTTGGGTTTCCAGAAGCAGAAACAGGTACAC
CAAAAGTCAGGCGACCCCATTGTTCTGTCTACCACAGCAGATCAATCAACTTTGGAA
CTTGAAGGTGCACCAAACCACAAGTTGTATGGACCTGGTTCCGTTGTATCTCATCCA
CCCAGAGCCAAAGGCGACCCAAACGGGGTTCCGTCAAGGACACCCAACCAAAGTGA
GTCACGGAAACATTGGTCTGGGAAGAAAACACAACAGGAGGGGCGAAAACCAGAT
GACACATCTGGGCCTAGGTCTAGCTGGCATAGAGGGAAACCTCAGCCAAAATGGTT
TGACAATACCCCTACCTACGAGAGACACAGGGATCATGGGAGACGACCACAAAACA
GGTTTGGAGGCCAGTATGCTGGACCCCCACGTGATTTCCATGAGGGAAAAGAACCC
ACAGAAAATTTTACTTAAAGAAATCTTGGCCCAAATCAGAAAAGGAAAGGGAATT
TTCAGGGTTAAGTAAAGAAGAAATACAGGAAATCAAACGTCTACTAAAAGACTTAG
CACAGAAGCCTAAGCCAGATAAACCAGATGTGTTGGACCTCTCCCTGAGTCAATCC
AACTTTGAACTAAATGGTCCGGTCCCAACCTCCGGTGACGGACCCCAGACGAGCAC
AGATCGTCACACTGCCAATCTGGTCCCTGTGACGGCCTCAGAGGGGGGTGCCTTGC
ATCCCCACCTATAAGCGATGAGGAAGAAAATGAGACCCTTCAATGCCCCAGTGCTG
TATTGGTGGTTAGACTGGACCCACAACAACAGTCACTTGAGCAGGGTGACTCCATGT
TTCCGGTCTACCCCCCAGTACAACAATTTCTGTGCGACCTCACAGAGAAGGGACTCG
CACGAAAATTTTACATGTCACTAACCCTAGAAGATGTGTTAACACATGAGGCCTTAC
TTGACACTGCTGCAGACATCACATTGATCTCATCCACCTTGTTTAACCAGCTGCGCG
CTGCCGTTAACGATGACAATAAAGACCTGAAACTGCAACCATGCTCTGTTAAGGTGC
GTCCTTATGCCAACACCAACACCATCTTGTCCGCAGTGTCGCTTATCCAACTGTCTAT
TGGTCCAATGACAGTGGTGCATCCAGTCTATGTGTCACCCATCAATTCGATTCCTTTC
CTCGTGGGTAAGGATCTTCTTAATCGCTTTGAACCTTTGATCGATTTCCAACAGTTGA
AACTGTGGGCACAAGTTCGACAACCCCTAAGCATTTCTTGGCCGAAGGACGCTTCAT
CGGCTCACTGCTATGCTGTGAGAACTGAAGAGCTTCACCTCTCAGATGAACAGGACA

GTTCAACCAGGGAAGCAGAAATGTCAGAGCCAACAGAGAAGAAACCCCCTGCACA
AAACGAACAGGTGATGTTCCAGCATGACACCTTCTTGTGTGCAATTAGCCCGCCAGA
CTCGGAGGGAGAATACATTCCCCACATCCTTGGTGGCATCCAGATGCTGGGCCACAT
GGCTAGCGACGCCACCTTGTACCTCTGGGCTGACCAGTCCGCCATCAGTCACGAGTT
TTATCAAACTCTGAAAAAAGAAAACCCTGAACTTCCATTCGCAGGAAAATCATGGC
GTTTTCCAATCCATTCATGCTCGAACTTGACTTTGAAAGCCATAGGCGTGTGCTCACT
CGCGCTAAAATGGAACCACCAAACACTCACACATTCCTTCTTGGTGGTACCAGACCT
GCCACATACTGCCTGCATAGGCAGTGACCTCCTAGTCCGGCTAGGTGTACAGGTAGA
CACCATTAACAACGTCTTATGGTCCCCTGTACCGGCTCGAGAAACTCCAATGACCCT
CGACGCACTGAGGTCCGGACAGACCATTCCAGAAGCTTGTCAGGTGGTCACTTCGG
ACCATGTATCCTTACCTGCTGGTATTCGGGAAGCCATCATCAGAATGCACATTCAAC
CAGGCCAAAGTCTTGGTCATTCTCAAGCACTCTTCCAGCCGGCCCCTTTGTTTTTCAA
ACTTGGACTTTCCTTGCCAGCCACACCTCTTTTGAGTCTGAATTCCAGAACCATATCT
TTGAAGGTGAGTAATGACAACCACAATGAAGAAACTCTTGTGCCAAGAGGGACCAC
GTTGGGCTGGCTCATAAGTACAGATTTCCATGACTTTGAACTAACGGTCCCTGTTCT
GGGAGGCATTCCCCACTCCCTTCTTGGTGAAGAGGTGTCAGATGTCATCTACACAAA
ACCTTCCAAGGCTATCGCCATCTACTCTGTCATGGACCTGACATCCGATTCGGTGGG
ACGTGTGGACTTGTCAGACAAAAATCACTTGGTTTTCCAGACCATGAATGACGTATC
AATTAACACCATTGATGCACATCCCAACACTGAGACCCAAACTCAGCCTCCACCTTC
CATCAGTGAACAAATTGATGAGGTTGTGAAATTGGCTGATGCTTTAACTACTGACAA
TGAACGTCAACAGTTACGTCAGGTTCTGTTAAAGTATCAATCATCGTTGTCCACAGA
TTCCCTTGATTGTGGATTAACTTCTATCCACTCAGTGCGCATCCCAACACCAGAAAA
CGCACCGCCTACCTTTGTTAGACAATACAAGATTCCTCTTGCATCGTATGAACCGGT
TCAGGAGATCATTGATGATCTGACCAACAAAGGGATCATTAGGCCCTGTAATAGCA
CCTACTCAGCCCCAATTTGGCCTGTACGTAAGCCAAATGGCAAATGGAGATTGACTA
TCGACTACAGACGGTTGAACCAGCAGGTTCCCCTATCTAGATGGCCAATGACGCAG
CTGGAACAGGAACTACCCAGGGTCAAAGAAGCTAAATACTTCTCTACCCTGGATGT
AGCCTCAGGATTTTGGACAATCCCTGTCCATCCTGACGACCAGCACAAGCTAGCATT
CACCTTTGCGAGGCAACAATACACTTTTACCCGCTGTCCGTTTGGATATGCAAATTCT
CCAGCAGAGTTTAACATCTTTCTGAACAAAGCATGTCCCGATGCCACACAACGTGGT
ACACTTATTTACGTAGATGACATTCTCATTCGTAATGCAACATTGGAAGATCACATT
GAAGAGATTGAACATGTACTTAAACAACTCACCTCTGCTGGAGCCAAAATTTCTCTC
TCTAAATGTCAGTGGTGTCGAACCAAAGTCAACTATGTGGGCCTCCTGGTGGGACCC
ACCGGAGTCCAACCACAAGTTGGCAGAGTTCAAGGTCTCAAGGACATTGCAGCTCC
AACAAACCTCAAAGAATTAAGAAGCTTTCTAGGAGTATGCAATTACTCACGCCAGTT
TATTGAGGACTATGCTGACATAGCTAAACCGCTCACTGAACTACTAAAAAAGGATGT
GCCCTTCAGGTGGGCTCAACCACAACAACAAGCCATGGAAACTCTGATAACCAAAC
TGAGTTCAGCCCCATGCTTGGCCTACCCTAACTGTGAAAAAGAATTCCACCTTGAAG
TTGGATTCTCCAATAACTGTTTGAGCGCTGGGTTGTATCAAATATATGACCAAGATA
AACGCGTTGTTGCTTATGCTAGTAAGACACTCGCTACACCTGAACTGAAGTACTCGG
ACTGTGAAAAAGCGTTGCTTGCCACTGTTTGGGCTGTGAAGCATTTCTCTAACTATG
TTGGAGGACAAAAGGTAATTGTTGAAACCCAACATCAGCCGGTCACATTCCTGAAC
AGTCAGCGAATCAGAGACGGTGTGGTCACAAACGCACGTGTAGCTTCATGGCTAAT
GGCACTGCAGAGCTTTGACCTCGAAATCAGGTACGCACAAAACAAAAAGAATGCTT
TGGGTACTGACCTAGCTGCATGCCAGCGTTGTGTTGACGACGCTCCAACAGAAATCA
GAGTACAAAGTTCAGGGATCATCACTGATCCACACCACCACTACTTTGATCAAAATA

TCTGCCAAGATATGGTTACTGCATACGTGGATGGGTGTTCTTTCCGCAAGGATAAAG
AACTCCTGGCTGGTGTGGGTATTGCATGGCTTCACGACGACCCCTGTGAACCCAAGA
GTTTCCATCTTGGGTCTCAAACCTCACAATACGCAGAAATTGCTGCAGTTTTAATTGT
TTTACAACTTGCAGTTTCTTTGAACATCCATACACTTGTGATCTGTACCGACTCAAAC
TACGCACGACTGAGTTTTTCATGTCATCTGGACCTCTGGAAACAGAACAATTTCCAG
ACGTCCAACAGAAAACCAGTGAAAAACAAAGAGCTTTTTGTGGCATGCGATCACTT
GATCACCTCTCACAACATCCAAATCTACTGGAAAAAGGTGAGAGGACATTCTCGAA
TCCCTGGTCAGGACAAAGTGATGAATGACTTGAGTGACTCAATGGCAAAACAGGGT
GCATTGACGGGCGAACACTGGGACTTCGCTGAAACATGTCGGCTTTTCCCCACCCA
CCCGCCCCTGAAGCTCAGATTTGCGCGGTCACCAGATCCACACCCTTGGCCTCCGTC
TCGGCCAAACCCATTTATTCTGACGATGACCTCATCAAACTGCAGGCAGACGATCCT
GCCATTGCCAAGGTCCTGGATTTTCTGGTCGATCCGACCAACAACGCCCCCTCGTCT
GGTGACTTTGAAACGATTCCAGAACTTCGGGCTCTTTATGCCCTTCATTCTTCCTTGA
AGGTGGTCAAGGGCATTCTTCTTCGTGTTTCCGGTGACCACGCTCCGCCGGCTCTCGT
GGTCCCTCGTAACCAAAGGGGGGTGATGTTGATGTATGCCCATGACCTCCCCTCTTC
AGGTCACAAAGGCATCAACGCAACATACTCTGCACTTAAACAGATCGCATTTTGGCC
ACACATGAAGAAAGATGTGACAGATTATGTCAATGGCTGCTTGGTTTGCTGTCAGTT
CCAACCAACTAACCCAACACACAGAGCACCACTGCAAAAGAAAGGTATTTCATTTC
CATGGTCAGACATCCAGATCGACTGGGTGGGCCCCTTGACAAAGTCATCCAGAGGA
AATAAATACTTCCTGACAGTGACATGTGCCTTCACAAAATGGGTGGAGTGTCTACCA
GCTCCAAATGACACTGCTCAAATGACCGCTCTGTTGCTTATGAACCACGTGTTCTCA
CGTTTTGGTTTACCAATGCGAGTGAATTCGGATCGTGGTACTCACTTCACAGCCGAA
GTGATGCAACAGGTCTGGAACACTCTGGGGGTAAAAGCGAACTTTCACGTGAGCCA
CCACCCTCAATCCTCAGGGCAGGTGGAACGAGCAAACAGAACTGTAGTAAACATAC
TGAAAAAGTATGTAGCGGCCAGTCACAAAGACTGGGACATCAAACTACCTCTGGTA
ATGATGGCCATCAGGGCCACCCCTCATGAATCCACTGGTACATCACCTTTCGAGATG
ATGACGGGGAGACAAATGACTTTGCCTCTACACCTTCTTTACCAGCCAGGAGAGGCT
AGTGTAGCAACAGCCTACACGGCGCATCAATACGTGGCTGACCTGCATGAACACTT
GAAGGCTACATTTGCTTTTGCGCAAAAGCATATGGAGAAAGCAGCTGAGGGTCGTA
AAACTTATTACGACCAGAAAGCGTCGCACCAAACGCTTGAGGTAGGAGACAAAGTC
TGGTACTACCTTTTTGTCCAGCCAACGGGACAGGCCAACCAGAAGGTTGGCAAACTG
GCCCGCAAACTTCTTCCACGCTGGACAGGCCCACATGTCATCACTGACAAGCTGTCG
CCAGTGGTATACCAGATCAAAATCAGCCAAGGCCGAAAGGAGCCCATTTTGAAATG
GGTCCATCGCAACCAGCTCAAACTGCACAAACACCCTGTGGGACTTGTTGGGTTCAA
CCAAACCCAATCAAATTAA
> Loki-Oja consensus sequence Smc-Gag-Pol polyprotein
MTDAVSSEASSHGDLFDDLEAALTEITENQIPGFEGRVRNLTAPELETKLSEIVAEITNDP
PKDKELTKTLGSLAILLTARLKSAHTLLNQALGKLDTRSQKLSQTQGEIKLLQHQREVA
QRRTDELECQNEELKTENDQLKREIKDCESQTQACQHELTAMHVQLIEVKETYKASQIG
LYSAEDDIIKLQQENANLTRQLGDQKEQVELMREQLAHLTSKLEASDQRYERLMVERK
SVTTDIKPPIESDISQQFYPSFRDVKGGAQAGRGASVGLPLVSPVPAPLQAERGDSGGSPR
LSPVPAPLSALTSKPTQANLLHDFRAEATRHRAPSDRDLDKISRNLTKFEPTPGGSNDPR
AYLDDIEFYLRKYSEATIQDKIYLIKLTSSREVIRFIERQPEFTRREWYFLRQAIEREFSDYT
IQSGLTVAMSVKQGRQESPQQYYYRLLRAYFGTRNEPGMEEDLNFKSLFLQNLHPGTSH
HLGVFANPHTATISHLRELAALGFQKQKQVHQKSGDPIVLSTTADQSTLELEGAPNHKL
YGPGSVVSHPPRAKGDPNGVPSRTPNQSESRKHWSGKKTQQEGRKPDDTSGPRSSWHR

GKPQPKWFDNTPTYERHRDHGRRPQNRFGGQYAGPPRDFHEGKEPTENFYLKKSWPKS
EKEREFSGLSKEEIQEIKRLLKDLAQKPKPDKPDVLDLSLSQSNFELNGPVPTSGDGPQTS
TDRHTANLVPVTASEGGCLASPPISDEEENETLQCPSAVLVVRLDPQQQSLEQGDSMFPV
YPPVQQFLCDLTEKGLARKFYMSLTLEDVLTHEALLDTAADITLISSTLFNQLRAAVNDD
NKDLKLQPCSVKVRPYANTNTILSAVSLIQLSIGPMTVVHPVYVSPINSIPFLVGKDLLNR
FEPLIDFQQLKLWAQVRQPLSISWPKDASSAHCYAVRTEELHLSDEQDSSTREAEMSEPT
EKKPPAQNEQVMFQHDTFLCAISPPDSEGEYIPHILGGIQMLGHMASDATLYLWADQSA
ISHEFYQTLKKENPELPFAGKSWRFPIHSCSNLTLKAIGVCSLALKWNHQTLTHSFLVVP
DLPHTACIGSDLLVRLGVQVDTINNVLWSPVPARETPMTLDALRSGQTIPEACQVVTSD
HVSLPAGIREAIIRMHIQPGQSLGHSQALFQPAPLFFKLGLSLPATPLLSLNSRTISLKVSN
DNHNEETLVPRGTTLGWLISTDFHDFELTVPVLGGIPHSLLGEEVSDVIYTKPSKAIAIYS
VMDLTSDSVGRVDLSDKNHLVFQTMNDVSINTIDAHPNTETQTQPPPSISEQIDEVVKLA
DALTTDNERQQLRQVLLKYQSSLSTDSLDCGLTSIHSVRIPTPENAPPTFVRQYKIPLASY
EPVQEIIDDLTNKGIIRPCNSTYSAPIWPVRKPNGKWRLTIDYRRLNQQVPLSRWPMTQL
EQELPRVKEAKYFSTLDVASGFWTIPVHPDDQHKLAFTFARQQYTFTRCPFGYANSPAE
FNIFLNKACPDATQRGTLIYVDDILIRNATLEDHIEEIEHVLKQLTSAGAKISLSKCQWCR
TKVNYVGLLVGPTGVQPQVGRVQGLKDIAAPTNLKELRSFLGVCNYSRQFIEDYADIAK
PLTELLKKDVPFRWAQPQQQAMETLITKLSSAPCLAYPNCEKEFHLEVGFSNNCLSAGL
YQIYDQDKRVVAYASKTLATPELKYSDCEKALLATVWAVKHFSNYVGGQKVIVETQH
QPVTFLNSQRIRDGVVTNARVASWLMALQSFDLEIRYAQNKKNALGTDLAACQRCVD
DAPTEIRVQSSGIITDPHHHYFDQNICQDMVTAYVDGCSFRKDKELLAGVGIAWLHDDP
CEPKSFHLGSQTSQYAEIAAVLIVLQLAVSLNIHTLVICTDSNYARLSFSCHLDLWKQNN
FQTSNRKPVKNKELFVACDHLITSHNIQIYWKKVRGHSRIPGQDKVMNDLSDSMAKQG
ALTGEHWDFAETCRLFPHPPAPEAQICAVTRSTPLASVSAKPIYSDDDLIKLQADDPAIA
KVLDFLVDPTNNAPSSGDFETIPELRALYALHSSLKVVKGILLRVSGDHAPPALVVPRNQ
RGVMLMYAHDLPSSGHKGINATYSALKQIAFWPHMKKDVTDYVNGCLVCCQFQPTNP
THRAPLQKKGISFPWSDIQIDWVGPLTKSSRGNKYFLTVTCAFTKWVECLPAPNDTAQM
TALLLMNHVFSRFGLPMRVNSDRGTHFTAEVMQQVWNTLGVKANFHVSHHPQSSGQV
ERANRTVVNILKKYVAASHKDWDIKLPLVMMAIRATPHESTGTSPFEMMTGRQMTLPL
HLLYQPGEASVATAYTAHQYVADLHEHLKATFAFAQKHMEKAAEGRKTYYDQKASH
QTLEVGDKVWYYLFVQPTGQANQKVGKLARKLLPRWTGPHVITDKLSPVVYQIKISQG
RKEPILKWVHRNQLKLHKHPVGLVGFNQTQSN*
> Loki-Oja consensus sequence *env* gene
TTGCCTCTCTGTCTTATCAGGAAGATGATACTGCTGGTGATCGCAGTGCTCCTCCTTG
TCTCAAGCCAGGCACGTTTGGAGATTGTTGAACCGGGACCGGAGTCAGGAATTGTG
CTGAAGGAACACCCCGGACTGTTAATCACGAACTGTCGCTGGCACACCCAAAAGAT
CTTTGTGAGGCTCCGTCCTGATGCAGTTTACAGAAAACATGTGCAGGCGACATCCAG
CTTTCTCTCTTGGGCTGCTGTTACCTGGACAATGGATGTAATTTCACATGCTGAGAAT
GATATCGTTCACATGCTAACTCAGATGCAAAAATTCACGATCACCCAGTCTGAGCTA
AGTGGCGTTGGACACCGTGATAAACGTTTTATCGCTGGGCTACTTACCGCTGCCTCA
GCTATTGGTTCACTGTTTAGTCTGGGGGTTTCCGCTGTAAATGCAGTCACCCTAGCCA
CCCTCAAGCGTCACATGAATGAGCTACGTGATGAGATTCCCCAAATCCGGGGTCAG
CTAGATATCCAACAATCACAGGTACAGACTATGGGAAAAACTCTACAGGACACTGT
GCTAATCGTGAACACCCACAGCGAGACTCTTAAAAACTTAACCAGGGCAGTAAACA
AGTTGTTAAGCGTAGTTAGTATAGACTATGCACACATACAACTGGTAACCATGCTAC
TGGCTGACATGACTCGAGACATCGGGTCATCCATTGATTCGCTAGTCCTCGGTAGAA

TACCCCCTTATCTAGTGTCACTAGAGTTGGTCCAACAACTCTTATCTGAAGCAACCA
GAGAACAGATAACCCCCCTACAGGCCCATCTGGCCTACACCTTGGGAAGTTCAATTC
CCATCCATGTGGATCCGACGGCCAGGGAGCTAGCTTTCTTGATCCACATCCCCATCG
TCTCTCCAGAAAACATTTACAGACTCAAAGACGTCATAAATGTAGGGATCTGGCAA
GGTGACGTTCATGTGAGGATGCACACACCCAAGGTCATCGCGTACCATGACAGCGA
TCCAGAGTTGTACCTAGCGCCAAACTTACAGTTGTGTACGCCCACTAGAGATATTCA
CTACCTGTGCCCAAGCAAACCGTTCGTGCGAGATAGCACTGGTGGTCTGTGTGGACT
TAAGCCAATCCACACAGATACTCGCTGCCCCACTACCATTACCCCTCGCAGGCAAAT
GACCACCACGCAGGCGGAGATCGTTGGAAATCGGTGGTTGGTTCTGACACCAGTAA
CAACCGCGACCATCATGTATGATCAACATGACACCTCAACCCGAATGGAACTACCA
TGGAGTTCCATCTGGGTTGACGTCCCTAACGATGCTATGGTACACGTGGACAACCTA
GCCCTCTATCACCTTCACCCAGATCAATATCAAGGTGAAGTAGAGATTTCGACTTTC
TTTAGCCAACATACGATCGAACTCAACTCTCGCACAGTGGAGCAAGTCGACTACGA
AGGGTCGCGTGTAATTGATGTAGGCCCCATTAATAAGGTCCTCAAAGAAATGGAAG
CACGTCAGAACCCTAATATTCAATCCCTTACTTACGCTTGGTCCACCGCAGATACTG
TATTAGTAACTCTGGTAGGACTCGGATACCTTCTCACTTTTGGAACCACCTATCTTTA
TGTTAAGAAAACCCAACATCTTTGTAGTCAGATAGACAGGTGTCGCAGTCGATTGGC
TCGAGGGACTCGACGGAGATTTGCTGAAAGAGAGAACCATGCTAACGACAGGACAG
AGGGGTCCCTCGAGATGAATCTTAGAACAGAACCTCTCTCTCCAACCATTCTCGACC
TGGGCGATACTCTGTGA
> Loki-Oja consensus sequence Env protein
LPLCLIRKMILLVIAVLLLVSSQARLEIVEPGPESGIVLKEHPGLLITNCRWHTQKIFVRLR
PDAVYRKHVQATSSFLSWAAVTWTMDVISHAENDIVHMLTQMQKFTITQSELSGVGHR
DKRFIAGLLTAASAIGSLFSLGVSAVNAVTLATLKRHMNELRDEIPQIRGQLDIQQSQVQ
TMGKTLQDTVLIVNTHSETLKNLTRAVNKLLSVVSIDYAHIQLVTMLLADMTRDIGSSID
SLVLGRIPPYLVSLELVQQLLSEATREQITPLQAHLAYTLGSSIPIHVDPTARELAFLIHIPI
VSPENIYRLKDVINVGIWQGDVHVRMHTPKVIAYHDSDPELYLAPNLQLCTPTRDIHYL
CPSKPFVRDSTGGLCGLKPIHTDTRCPTTITPRRQMTTTQAEIVGNRWLVLTPVTTATIM
YDQHDTSTRMELPWSSIWVDVPNDAMVHVDNLALYHLHPDQYQGEVEISTFFSQHTIE
LNSRTVEQVDYEGSRVIDVGPINKVLKEMEARQNPNIQSLTYAWSTADTVLVTLVGLGY
LLTFGTTYLYVKKTQHLCSQIDRCRSRLARGTRRRFAERENHANDRTEGSLEMNLRTEP
LSPTILDLGDTL*
>Loki-Sri consensus sequence
GTAGCGGACACTGGTGGACGCCAGTGACCAGACACCAGATTTATGGTAAGGACACT
GGGGTGCTCTGGAGGATATGGGGAGCCTCCCATGTCCACTTTCCCATCATCAAACCA
TCACATGGATAGAAAACAGGGACGAAATGAGTTCCCCTCACCAATAGACTAGCTTTT
GAAGATACCATCGGAATTTATGGCTAGCTCATCGAACTCATTCTGAACCAACACTCC
TGCACCAAATGGGAGACAAACACAACTCTCCCCCACAGACTGACAATGAATTCATG
TTTCATCCCTGACACTCATCCGGGATTGATAGGAGGGTGCCACCTCTTATCAAATAA
CATGAAAAGGAACATTTAGCCATCAAGGTGCTAAACCACCAGGATTCACACTCCTG
GTTCATCTTCTCTCACAAGAGAAAACAACACTCTCACCTTTAACACCCTTTCTTACCC
AAAGGGGGGTCTCACGAGGAACCACAGACCCCCGCCATAAAACCATCCTACATTCC
AAGCCGGATTGATGAACTATTCCACATTTTATAAAGCATGATTGAAATGTCATTGAA
CCAAACGTCCTATCACCAGGTAATTCATGCAGTATTACTCTGGAAAGATTTCAACTA
AGTATGATCCGAAGTTCATTTCTGTAGCCCACATGGAGGCTGAGATAGGTTTTATTA
GTTCTAAGGTAAATATCCACCATGTTACAATACTGCCATCTAAGGACCAGACTGGAC

CCCTTTTTCCCTTAGGGAATTAAGAAATAGACAAACTTAATGCATCACTCTGACATT
ACTGTTTTATATTGTAACTTAGCTATATATTTGTGTGGTTACAAATTGTTATCCTCTC
ATTGATCTATGTATTCTATGTGTGTTGTTAAACATGATGATTTTATTAGTTATCTTGT
ATATTCCAAACTTTAGCTAGGTAGTTGATATAGGCCGTTGTCAAAATGAATGGGCCA
TGGTGGGAGTAGAATATGCAACACTTGCGCGTAAAACCAAATTAAACAAATTAACT
TAAGGTTTATTTCTCTCTTTCTTTATTGAAACAAAGAAAGAGAAGGAAGACCAAAG
GGGTGGCGACAACTGTCAAGCCCCCCCCAGCTACACACACCCCGAGGATAAAACAG
CAGGATCCTCCAGGACTCGTCTTCTCGTTCTTTTTCTTCTTCTCGTTCTTCTTCTTCTTC
TTTTACGTCTTCTTTTGCATAGCTTCTTTTACAATCCTTCGTCTTACAACTTTTCTTAC
AATCTTAACTCTTACGCCTCGTTTCGCGCACTAGCTCACGTTCGTTTCTAGCTTAAAC
TTTACTTCTTACACTTTGGTGCGCGATTTTTGCCTTACTATTTCAACAAGCCAAGCCT
TCCAAATTGAATCACACACAGGTGTGTATAAGTAATTTTGTCGTTTTTCTTTTTGTGA
CAAAATCCAGCTTTTGTCTAGAAGATTTTGAAGGTTTTGCGTCGTTTGGCCAACACC
AACTCCTTCCAACTTATCTGGCAAATTAAGTCAAATACTAGCGTTTATTCCGCTGTCA
ACCTCCGGCCCTCCGGCTACACGCCACACCAGGACCGGCTAACCAGATCTCGAGAG
TGCCTACGAAGAAAGTCTGCTGGCAGACTGCGCCGACGCTGATCCACCGAGACATC
GGCAGAGCCCGTCACCCCTGGCAACGGCCGTCGTCTCTGAGAGAAACCCACTTGGA
GTGTGGCGTCCCCGTCGCCTAGCAACCAAACTCCGGGGCTCCAGCCAGGATCTCAAC
AGCGCTCCGCTACTCAAAGTAGGCAACTCTATCCTTTGCTCCGAACTGGGTTGATGT
CGGAATAATACTAATCCCGATTTAATAATTCTCCACCGTTACGAATAAATGCTATAC
ACATATCTAAATCCATTCTCTTTCACCATCGCCCAAATCACTGCTAGTTCAGCTGGTA
GATATAGTTATTATTAGTTATAGATTCATGTCATTATCCTTGTTCATTTTACTGTGTTT
AGTTAGTCAATAAATATTATCAACGTAACCCAGTCGTTGCCTGTGTTTCTTTGAGGTG
AAGATCGGAGATCCCTGTAATCGAGAAATCACAACTTTGGATGTGAGATTGATCAA
AACAATCTGTACTTTAATCAAATTGTCCTCCCTGAGGACTGGTGCCCCATTCTTTTAA
CGAGAGCAAAATTCAAATTGTGGTATTCACGCTACATATATTGGTGCCCCGTGTGAG
GCGATTTCAAATTACCGTTCTTCACTACAAAGAAACCTGTTGAATGCTGAGAGTGGC
CTCACTCTCTATCTGCTGATTCAATTGCATTTTGGCTTTGTGCTTGTTACCTGTTGTGT
GTGAGCATTTGAAACCTTGAAATTTTATTTGCAAATCCTAGAATCTAAGCTCAGTGG
TAAACTGTGTGGTTGGATATAATAACGTGGACGCACGTTGATTAGGCTGAACACGTC
GAACGGAGTCACTGCGCCGACAGTGATTTCCCCGACCAGTGGTGGCGGTAACGTAC
CTAAGCGTACACCCGTCAACGTTAATGATAGCCTAGAAAAGTGTTGTAGCTAGAGCC
GGTGAGCTTTACGCTATTTATCTAAAATTGCATGCGACAGCGTGATCATACGCTTAT
TGAAAAGGCCTGTTGGTGGAAAAATAGCTATTGCGTGTCCCAATAGCCAGTTGTTAA
TCCACCCTGAGTGTGGGTGGATTTGTCCGCGAGTCCCGGTGAATTCAGCCTGCCTCA
GACCCTCGACGCTAGCCTGTTAGCTTAGCTACCTAACAAGTCAGTAACTGCGACCTG
TCGCAGCAGAGGTTTATACGTAGGGAATACCTACGCGGGAGATTTAAAGAGAAATC
CCCGCCGTCTAACGTCAAACACCGTTGTGGTTTGAATGAGAAAGGAAATAAAGTGC
TTTCTTGTCTGAGGTCCCAGAACGGACCTGTGTTGAAAATTTTGAATATTGTCTCAAT
TACGGAGCGTAACCCTAGCTCCGTGAATAATTGACAATACAAATAGGGATAGTGCTT
GAAACGATTGATTGAAAAAAACCCAAAAGGAGACTTGGGTTGGAAATTGAATACTG
CAGCAATAGTTGCGCGTGGTGTTTGCCCGTGGTGGGTGCCCGAGGCCGCCGTGCCTC
GTGATTGATTCTTTGAAAAAGTTTAAAGCTGTGGAGAAGTTGCCTGCGTGGTGCTCT
GCGCAAGACTAACCTCGTCTCACCTGAGCTGCGATAGTCTTCCAGTGGGCGGGACTA
AAGCCGCCATCGCTGACGACAGCCCCCTACTCTTCCTCTTTAATCCTAAATCGCGAA
CCAAAACAGATAAACCTCTGTCGCCCCCTGGTGGCCGATTAGGGTAACTGAGTTAAG

AGTGGGAATATTCAAACACTGACTCTGCGCATGAAACACCTCACAATCTAGAGCTTT
AAAGCGACCTCGAACTGAAAATTGGTTGTAAACTCACAACTTTATTACATTATAGCC
TGTTTTTAACCTGCCACGCTATTGATTTAACCTGCATGCTACTCTCACATTCAGACCT
TAGACTCTAACCCACTTGAAGTGAAATAGAACTAACTCTGAATCATATACTTTTTCTT
ATATCACCCGATACTTAGTGACTTGAACTGTCTGCTGGTGCATGAACACATTATCAG
CCCTTAGGTTAAGAAATACTACAGAGACACACTTGTGTTAACCTCTCAGACTCACTT
CCTCACTATAGTATTAAGCGATATTAACCACACCTTCATACTACTCCTGATCCAAAG
AGTGGACCTCACCTGGATATGCTTAAGAATTCGTGCATGCTTATGCGCTTTTCATGTC
TTTGCATCGCTTTAAGTGGATTGCTTGTTTGAAAATTCTTCTTATTGACAGTTTTATTA
ACTCCTCCCTTTTACCTGAACACCACACCACCACGATTGAACAAATATTATCTAGGA
AGATAATATCCCTAAGTTAACTAAATTAAGTTTCACTTGTGACTCCCTGTTTCGGTGT
GCTGTCCCTCTGCCTGTGTTTCTCAGCTTGCATTGCTACTAGCTTTGATACTACAAGG
TAGACTCATCAATTGGAGGTTCCAATTTGGGAGTTATCGTACCTGGACCTGGTTAAT
CTGGAACTCTACCTATAAGGTGATCACTGCCCTCCAGTGACAACTTAAACAGGTGTT
CCAACTTTAACACAGCAACGTGGGTCGTTAGCTTCCCAATTTTGAGTTCTATCCCTTT
GAGAGCCAACTGCCTTGGAGTCACTCACTGCTGATGATTGCTTGCTTGTAATAAGGC
ATTACAGTGACACCGCCGTGACTGCAGTCTCGAGAGCGATATATATGTATATACATA
TATTTTTCTGTTTTTTTTATAATTGTTGTTGTTTGGTAGTTTTCTAGTGCCTGCATATA
TCAGTCTGCACTGTTATCGGCTCTGACACTATAAGGCAACCCTTTCAATTGGAGGTT
CCAATTTTGAAGGTTACTGCACCTGGACCTGGTCAATTTGAGTTCTACCCACACCAA
GGGGTTCCAATTTGACACAGTAACGTGGGTCTTTAACTATTCAATCTTGAGTTCTATC
CCTTTGAGGACTAACTGCCCTAGAGTTACTCACTGCTGAGACGTGCTAATTGTAAGG
CAAGGCAACTGCGGGACCAACCCCCATTTGCTAGAATAATAACGCCATAACAGTTA
AGTGGAATAACACATAGCCCTCAGATATCTCAACGCTCATAACGAGTCCATAGTAA
GAGTACACAGAGGTTGTGATTAGAACTCACAGAACGACTTGATAATCACATAAGGC
CTACTAACAACTGCTTACTGAACTGACGCTTTCCCTCTCCTTTTTCTCACTTCCTGATT
TGGAAAGAACCACCCAACTCAGTAACCTAACCACCAGCATGTCTACACCCAGCAAA
GTAGACTCGGCAGCTGCTGCCACACCTATACCCAGAACCTCAAAACCTGAGGACTC
CCTTGAGAACTTAGAGAGTGCGTTAGTGGACTTAACGGAGGATGTGCTACCCGGCTA
TGCCATTCTCATCCAAAACACCGCCCCTTTAGAACTAGAACCAATAATGCTAAACTT
AGTCACTGAGGCCCTAACTAAACCCCTAAAAGATAAGGCGTTAACCAAAATGCTTG
GCAACCTGGCCATCGTGGTGACAGCCCGACTCAAATACGCAAACGATTTGTTGACA
CAAGCAGCCGATAAACTCAAAGTTCGTGCACAACAGCTTTCACAGTTACAAGAAGG
GGCCACATTCTTTCAACATCAGCTAAACCGAGCCCACCGAGCCTCGGATGAATCTGA
AACACGCATTGAAGAGCTCGCAGCTCGACTTGATGTGGCCCAAGGAAGAGAACTTG
AGTTGGAGAACAGAAATCAAGAGTTAAAACGGTTACTCCAAAACCGCGAACAAGAG
TCAACTGAACTTAAAGACCAACTTAATGAGATAAGTAAAGAACGAGATGAATCACA
TGATCTCCTTAAACGGGCAATGGAAGAGATAGATGACGTAAAAGCCAAGAATCAGA
CCTATGCAGATCACCTGCAGGATGATCAGAAACAGATTCAACGGCTACAAGAGCAT
CTACACGAAGCGGAACTCAGGGCCAAGGCATCTCGCCAGCAACAGGAAAGCAAAG
AGGCAGAGCTTGCATCAGTCAAGAAAGAGCTACAGCAGTCCTTTGAAACAAGACGC
ATGCAGCAAACGGGCCCATATGAAATATTTCGGAAAAAAGATGACGAACCACCCAT
GAGGGGAACTCAAGGTGACGTCCTACCTCTTTTCTCACCTCTTGTGTCAAATCCTTTC
GACACTGGCAAAACGCTTCCAAAGCTGGCCACCACGACTTCTTCCCCGGAACCTAG
GTTTCCCTCAAGCACCCGCACCTCGGGTGTCTCAGATCGAGACCTTGACAAGATTGC
TCGCAACTTGCCACGCTTTGAACCCACTCCTGGTGGCTCCCAAAACATCAGCAGCTA

CTTGAAGGACATCGACTTCTACCTGCGAAAGTTCCCAAATGCCACGGCTGATGACCG
GATATACTTGATCAAAGTGACGTCCGGATGTCAGGTGGGTAGGTTCCTTGAACGACA
GACAGACTACGTCAGAAGCAACTACGACCTGCTTTGCCAGGCATTAATAAGGGAAT
TCTCTGACCCTATCTCACAAACAGGTCTAACAGCTGCTATGACCACCAAACAGGGGA
GGCATGAGTCTCCTCATCAATATTACCACCGTCTTCTTTTGGCTTACTTTGGCAACCG
AAACGAACCTGGTATGGAGGAAGACTTGAATTTCAAAACTCTCTTCGTACAAAACCT
CCATCCTGGTACTAGCTGTCACTTAGGTGTCATGGCAAACCCGTACACCTCAACCAT
CCAACAACTACGTGAGCTTGCTACCCTTGGGTTTGCCAAACAAAAGCAATCCCACTC
AAAACAGTCAGAACCCAACATCGTTCTGACCCTCAACGCCAATGCTCCGCCCATGG
AGTTAGAAGGGGCCCCAAGCGGGGAGATACAGGGACCTGCCAACCAAAGCTCCAG
AAGCCATCGATTTACAGAAGAACATAAAAGGTACCTACCCCCACACAACTACCACG
AATCCTCAAGACCTCCCTGGTCTGGTAAGGGACAAAACCAACAAAGATACGAGCCA
CGCCGTAACTCAGTGGATCACCGCCGAAGCAGGTTCCCACCCAATCAATGGGAAAA
TCGTAGAGCGCGTGATGACCAGGGACAGCATCGATTCCGGCCCCAAGGTAGATACC
CACAAGACAGGCGACAAGTACAACACACTCGCCAAGACTCATCCGAAGAAGACAC
AAAGCAAAGTGGTGACTCACAAAAATCTCCACGGAAATCCGACAAGACAGAAGGTA
AGAACAAAACCTCTCGTCTAAGCGAGGAAGAGATGGAGACCATCAGGCGCCTCATC
CAAAGAATGCAGCTGAATAGCAACAAGGCTGATCTCCTATCAATAACCTCAGTAGC
CACTGAGGCTGGCCCTCAAAACAATTCCCCCGACCAACCTGATATGGTGGTTGAAG
AATTGTACCTAGCCCCTAGCTTACCTGAGTCTATGCCACCACAAGACCAGACGCTTT
CGACCCCAGACGACAACGCTAACGAATCAGCTCAGGAAAGTAACAAATCGATGCCC
CCAAGCGCAGTCTTGGTGGTGCAGCTTGACCCCGAAGACAGGTCCCCTTGTGAAGA
AGCCTCGGCCATCCAAGTCGAACCACCTGTTCAACAGTTTCTCTGCCACCTTGATGA
GAAAGGGGTTGCACGGAAATTCTACATGTCTCTGACGCTTGAAGGAACGCTCAAAC
ACGAAGCACTCCTTGACACGGCAGCAGACATCACGTTGATGTCATTAACACTTTTCA
ATAAGCTACAGGGAATGGCTCGCCAATCTAACAAAAACTTGAAACTCCAACCTTGTT
CTCTGGAGGTCCAGCCATACGCTCCCACCGGAACCACACTGACAACCATGTCAGCTA
TACAACTGACCATTGGACCCATGACTCTGGTCCATCCGGTGTATGTGTCTCCCTTGA
ACTCCATCCCCCTTCTCATTGGCAAAGACCTCCTCAATCGATTTGAACCCCTAATCG
ACTTCAAACGCCTGAAGATGTGGGCACAAGTGCGACAGCCTCTGCCTATTTCTCTTC
CCCAAGACACTCACATACAGTGTTGCGTCCTTGAGACAGAAGGGGTCAACACTTTGG
CCCAGCGTGCTCGGTCCAACCAACCAGACGTGACTCAGGTCAATGATGACCCCGAA
GTGAGGACGGACACCGCCACAATACCAACGTGGCCACCTACAGACTCTCTGATCCA
ACGAGAGACATTCTTGTGTGCATTAAATCCCTCAACTAATGCCAACGAATACTGCCC
TCAAGTGGTCAGTGGCATAAGACTTGAAAACACTGACGTTGAGGACGTGACTCTAG
CCCTCTGGGCCGACAAATCTGCTATTAGCCAAGAGCTATATAACTCCCTCTTACAAG
CAAACCCCACTCTACAGCTTATGAAGAAAACTTTTCGTTTTCCCTCAGACCCTACGC
CACAGACTATGCTCACAGCAGAAGGTGTCTGCGCCTTGAAGCTACAGTGGAATCAG
AGGGAGCTCAAACATTTCTTCCTGGTCATCCCAGACCTACCACATCCTACCTGCATA
GGCAGCGACATTCTGGTTAGGCTGGGTGCACAACTGGACACTGTCAACCATGTTTTG
TGGTCACTCACCGGAGCTCGAAACGACTCAAGCTGCCCTGACCCTGAGAACATGAA
GTCGGGACAAACCATCCCCGACGCCTGTCAGATAAGTAATGAACATCACGTAACGG
TACCACCATACACCAAAAACGTGTCAGTACGCGTACACCTGCGAACAGGTCAACGG
CTTAATCACCCACAGGCACTCTTTCAACCATCTCCCCTTTTCTTCGAAATGGGACTCT
CCTTGGAAGCCACTCCCTTACTAGAGCTATGCTCCAGGTCAACTTACTTGTTGGTTCA
AAATACCACCGCAGAGACCATACTGATTCCCAGGTCCACACCACTAGGGTGGCTTAT

CAGTACAGAGTTCCACGATTTTGAGTTAGGAATCCCAGTTATTGGGCGTATACCCCC
TCCTCTACTCCCTGACAACTCAGACGTGCAGGTGGTATACACTAAACCATCAAAAGC
GATCGCGATCTTCCCAGTCATGGCGTTTGAACACGACTCCATCGGACGTGTTGATCT
CACGGAGGATAATCAAATGTTGCTTCAGAGTATCAACGTTTTATCTCTGGATGTAAC
GTCAGAGACAGCCCCTACACCACCTCAATCAGACATCCAAAACGAACCTGGAGAGG
CGACACTGGAACAGACAACCGAGCCACGCTCGGACTTCATGGCTCAAATAGACCAG
GTGGTGACAGAAGCCGACGCACTGGACTGTGAAGAAGAACGCCAAAAACTCCGCA
ACATACTCCTGAAATACCAAACATCCTTCGCTGAGGACTCTCTTGACTGCGGACTCA
CCACCATCCACTCAGTGCGCATCCCCACTCCACCTAATGCACCACCAACCTTTGTAC
GACAATATAAGATTCCTTTAGCATCGTATGAACCTGTACAGGAAATCATAGACAACC
TGCTAGAGAAAGGAATCATCCGCCCTTGTAACAGCACCTACTCAGCACCGCTCTGGC
CAGTGCTTAAACCCAACGGCAAATGGCGGCTGACCATCGACTACCGTAAACTCAAC
CAACAGGTGCCTCTGTCTAGATGGCCTATGACACAGCTTGAACAAGAGCTCCCTAAA
ATCAGGGATGCCAAGTACTTCTCCACGCTTGACATCGCATCAGGATTCTGGACAATC
CCAGTACATGTTGCCGACCAGCACAAACTGGCATTCACTTTCGCAAATCGACAATAC
ACCTTCACCAGGTGCCCGTTTGGCTACGCCAACTCACCTGCTGAATTCAACATCTTCT
TAAACAAAGCGTGCCCTGATGCTAACGACAGGGGCACACTCATCTATGTTGATGAC
ATCCTCATGAGAAGTTCTACTCTGGACACGCATCTGGAGGAGATTGATCATGTCCTG
GAGCAGCTCACAGCCGCTGGCGCGAAGATATCACTCTCCAAATGCCAATGGTGTAA
AACCAAAGTGAACTATGTTGGCCTCCTGGTAGGACCACGGGGGGTTGAGCCACAGA
TAAATCGTGTACAAGGAGTCACAAACATCAAAGCACCAACGAATGTCTCTGAACTT
CGTAGCTTCTTGGGGGTATGCAACTACTCCCGCCAATTCATTGAAAACTACGCTGAC
TTGGCCAAGCCATTAACAGAACTCCTTCGAAAAGATGCGCTATTTGTCTGGAACGAT
TCACAACAACAGGCAATGCAGGCCCTTAAAGAAAAACTTTGCACGGCTCCTTGCTTG
GCCTATCCGAACTGCAACAAAGAGTACCACTTGGAAGTGGGATTCTCTAACCATTGC
TTGAGCGCAGGCCTTTACCAAAAGCATGACCTCGACAAACGCGTTGTTGCATATGCC
AGCAAAACTCTGTCACCCCCTGAACTTAAGTACACTGACTGTGAAAAAGCTCTCTTA
GCGACGGTGTGGGCAGTCAAACACTTCTCCAACTATCTTGGCGGACAAAAGTGATT
GTGGAAACCCACCACCAGCCAGTAACCTTCCTCAACAGTCAGCGACTCAGAGATGG
CGTAGTCACTAATGCACGCGTAGCGTCGTGGCTGATGGCCCTACAAAGCTTTAATAT
GGAGGTGCGCTACGCTCAAAACCACAAAACACCCTTGGGCACAGGACTGGCAGCCT
GCCAACGGTGTTCAGACGACTCCCCGACCAGCGTAGCACCAGTCTGCGAACTGCCA
CCTCCAACACCTCCACGCCACCACTACTTTGACCAAAACATATGCAAAGAGATGAC
AACGGCATACGTGGATGGCTGCTCTTTCCACCATGAAACTCAGATCCGTGCAGGCGT
GGGAGTTGTGTGGGTCCATGACACCCCATATAAACCCCAGATGTTCAAACTGGGCTC
CCAGTCTTCCCAGTATGCTGAAATTGCTGGCATCCTGATCACGCTGCAACTGGCTGT
AGAAAAAGAAATCAAAACGCTGGTAATCTGCACCGACTCCAACTATGCGCGACTCA
GCTTCTCTTGCCACCTTCCCCTCTGGCAACGCAATGGCTTCATAACTTCCAACAGGA
AACCAGTCAAGCACAAAGAGTTGTTCGTGGCCTGCAACCACCTCGTGGACACACAA
GACATGCAGGTCTACTGGAAAAAGGTCAGGGGCCACTCCCGTGTTCCCGGGCAAGA
CAAAGAACTGAACGACCTAACTGACACACTTGCTAAACGTGGCGCTGTTGAAGGCA
CGCCGTGGTCACTGGAACCAGAACCTGTGCCTGAGCCGGATTCCCTTGCCCTCCCAC
CCACCCCTAGCATCTGTGTCATCACTCGGTCTCATGCGGCCACTTCAGCGCTGCCAC
GCCCAACTGACATAACCACTGCCTCTGTCCGTCCGGCTTTCTCTGATTCGGACCTCAT
CTCACTGCAAGCGCTTGATCCAGCAATCAATACAATGATTGGCCACCTAACCGACCC
AACTGCTCGGTTGATCACAAGCGATGAACTTGCCCAACTACCAGACCTGAAGATCCT

CTTCAACGGACGTGCGTCCCTGCGCCTGCTCAAAGGCATCCTGGTTTATGCTTCCGA
TGCGCTCACTTCACCTGTGTTCGTGGTACCTCGTAACCACAGGGGGGTGATGCTAAT
GCATGCCCACGACTCCCCATGCGCAGGACACAAAGGCTCCAAAGCGACATACAATG
CACTCAGGCAGGTCGCCTACTGGCCACAAATGCGAAAAGATGTGGCTGACTACATC
AAAGGCTGCTTAGTCTGCTGCCAGTTTCAACCAGCCAATCCTGTGCACAGAGCCCCC
TTACAGCGAAGAGGCATTTCCTTCCCCTGGTCCGACCTCCAAATCGACTGGGTTGGA
CCTCTGACCAAATCAACCAGAGGCAACAAATACTTCCTCACGGTCACATGCGCATTC
ACTAAATGGCTCGAGTGCTTACCAGCCCCAAACGACACAGCTCAAACAACCGCATA
CCTACTGATGAACCACATCTTCTCCAGGTTCGGTCTGCCGAGCCGCGTCAACTCGGA
CCGGGGCACACATTTCACTGCGGATGTAATGCAAGAACTCTGGCAGATACTCGGTGT
GAAGGCCAACTTCCACATCAGCCATCATCCCCAAGCCTCTGGTCAGGTAGAACGAG
CTAATCGAACCGTGGTTAGCGTTCTAAAGAAGTACGTGACTTCAAACCACAGGGACT
GGGATGTGAAACTGCCGCTGGTCTTAATGGCCATCAGAGCAACACCGCATGAAGCA
ACGGGGTTTTCGCCATTTGAAATGATGACAGGCCGACAAATGACACTCCCACTGCAC
CTGCTTTACCAACCAGGTGAAGCTAGCGTTGCAACAGCATACACAACCCACCAGTA
CATGACGGACTTGAACCAACACCTTCAAGCCACCTTCGCCTTTGCCCAAAAACACCT
GGACAAAAGCGCAGAAGGACGTAAGGCCTACTATGACCAAAAGGCTTCACATGAAG
AACTCCAAGTGGGGGACAAGGTGTGGTACTACAATTTCTTCCAACCAACAGGTCAA
ACCAACCCAAGCACGGGAAGGCTCGCCCGAAAGCTCCTCCCTCGGTGGGCCGGTCC
CTACCTGATCACAGACAAGCTTTCCCCCGTTGTGTACCAAATCAAGGTTAGCAAGGG
CCAGAAAGAACCGACCCTCAAATGGGTCCACCGCAACCAGATCAAACCACATAAGA
CCCCCATGGGATATACAGGGGACGACAGCACCCCAGTCAGGTCATAAACTGATTAT
AAAACCCTTAGCACACCTAGATATCGCTCACCAACATCGAACAATCGCGATAGTGC
CTCGATCGTTGTGCTAACCTAAGAACGAGAAATCGAATTAAATACACTACGAACCC
AAATTTTAAACTAAAACCCCAATCTACCTTCAAGCCCGTTCAACACTTCCAACAAGT
GATTGACCTCGCAAAGTGAACTAGTGGGGGTGACTTAACTAATCGCCAGCCCATTTC
AAACGAAACACTAACTAAATTGGTGGCGCTAATAACCGGTAACCCTGTGTGTTTTTC
CCATCAGGATGACGTGGCTGCTGCTCGCCATCGTGACTCTCGGGATGATAACTACCC
CGGCGAACCCCGAGACCGTCGAACCTGGGCCTAACTCGGGGATTGTTCTAAAGGAA
CAACCAGGACTTCTAATCACAAATTGTCGTCTACACACCCAGAAAGTGTTCGTTAGA
TTCAACCCTGAGATGGTTTGTCGCCGAAATGTCCCCACCACAGCCAAACTAACCTCT
TGGGCTGGCATTCGCTGGAACAAGGAGGTTATAACCCACGCAGAGGTAGACATAAC
CCACATGCTGCAACAGCTTCAGAAATTCACCATCACTCGGTCTGAATTAAGCGGTGC
CAACAAAAGAGAGAAGCGGTTCATTGGAGGCTTGCTAACCGCAGCTGCTGCTCTCG
GCTCTCTCTTTAGTCTAGGGGTCTCCTCAGTTAATGCTGTAAACATTGCCGCTGTCAA
ACGACACATGGGGGAGCTGCAGGCTGAAATCCCCGGGATTCAGGAACAAATGGTAA
AACAACAAGCACAGCTAAAAACTATTGGGAAGACGTTGCAGGGAACAGTCCTAGTA
CTTAACACTCACAGTGAAACCTTGCTCAAGACTTTAGGAGCGGTAGACCTGTTACTG
TCAGTCCTGCAGGCAGACTATGCTCACACACAACTGGTGACTATGCTAATGTCCGAC
ATGCTTAGGGATATTGGGTCATCAATTGATAGTCTAGCTATAGGTAGGATCCCTCCC
TACCTGGTTCCACTCTCCCTGGTACAGGACATCCTTGCCACTGCCACTCGAGACATA
GTGACACCACTACAGGCTCATTTAGCTTACACCCTAGGTAGCTCAGTACCAATCTAT
GTAGACCCAGAGGCCAGAGAAGTAGCCTTTATTATAAATCTACCCATAGTGACACCT
GAGAACATCTACAGACTGAAGGACGTAGTTAATGTTGGCTCTTGGCAGGGTGAAAC
GCACGTCAAAGTGCACACACCACCCGTTGTGGCATACCACGACAGCAACCCTGACT
TATACCTTGCTCCCAACCTCCGCATGTGTACACTGACAAAAGACATCCACTACCTCT

GCCCCAGCAAACCATTCGTACGTGATAACACAGGTGGTATCTGCGGCATTCAAACC
ATGTCTAAAGACTCTCAATGCCCAACAACAGTCACGCCCCTCTACCAAATGACTACG
ACTCAGGCAGAAATAGTGGGAAGTCGCTGGTTAGTTAACACCCCAGCGAAAATAGC
AACTCTCACTTATGACCGACACGACACCTCCACCCGTATCGAACTACCTCACCGCAC
CCTGTGGGTTACCGTGCCTGAGGGTGCCATCTTTCACATAGACGACTTAGCCCTCTA
CCATCTAAACCCTGACCAGTATGAGAGTAACATTGAGGTGTCTGACTTCTTCAACAA
ACATGCCCTCGAGTTGGATTCCGAGACCCTAACCCAAGTCCAATACGAGGGAACTC
AGACCATAGACGTAGGCCCTGTTGACAACGTTCTAAAAGAAATTGCCTCCCAGACC
CATACTCCTAGCCAGCCTTTAACCTATGCCTGGTCTACTCCAGACACTGCCATAGCC
GCCCTTGTTGGCTTAGGATACATCCTGACCTTTGGTTTGACGTGCCTATACATCAGGC
GCATGAAAACTTGGCAGTGTCACGTAGATCAGGGTCTGGACAGACTGAACAGAGGA
ATAAGAATGAGGAAAAGAGAAACGAAACCAACCACTAATGACCTAGTGACTGTTGA
CGATGAGACAACCTTCACCCGAAACCCTTCCAGTGGCCCTGCCATCCTGGAACTATT
ATGCGACTAGTTAGATAAACCCATGTACCCAGTAGTATCGATATAGAGCCTAGTAAG
CACTACTGGTAAGAAGTAGTTTAGATTTAGTGTCTTAATATAGGAAGTAACTATGCT
AGTTTTTCTTTGTTTTCTCAAGTTGACACCCTACATACATAACTGATACACACTTAGT
TAAAGACATCAACCCAGATTGCAATGGTAGTGTACTGTCACGTCACTGCCACCCAGA
TGTCCTCAACTCAGCTCAACGGTCCAAAACTCCCTGGACAGCGACCATTGGACCGAG
TGGGGGAATGTAGCGGACACTGGTGGACGCCAGTGACCAGACACCAGATTTATGGT
AAGGACACTGGGGTGCTCTGGAGGATATGGGGAGCCTCCCATGTCCACTTTCCCATC
ATCAAACCATCACATGGATAGAAAACAGGGACGAAATGAGTTCCCCTCACCAATAG
ACTAGCTTTTGAAGATACCATCGGAATTTATGGCTAGCTCATCGAACTCATTCTGAA
CCAACACTCCTGCACCAAATGGGAGACAAACACAACTCTCCCCCACAGACTGACAA
TGAATTCATGTTTCATCCCTGACACTCATCCGGGATTGATAGGAGGGTGCCACCTCT
TATCAAATAACATGAAAAGGAACATTTAGCCATCAAGGTGCTAAACCACCAGGATT
CACACTCCTGGTTCATCTTCTCTCACAAGAGAAAACAACACTCTCACCTTTAACACC
CTTTCTTACCCAAAGGGGGGTCTCACGAGGAACCACAGACCCCCGCCATAAAACCA
TCCTACATTCCAAGCCGGATTGATGAACTATTCCACATTTTATAAAGCATGATTGAA
ATGTCATTGAACCAAACGTCCTATCACCAGGTAATTCATGCAGTATTACTCTGGAAA
GATTTCAACTAAGTATGATCCGAAGTTCATTTCTGTAGCCCACATGGAGGCTGAGAT
AGGTTTTATTAGTTCTAAGGTAAATATCCACCATGTTACAATACTGCCATCTAAGGA
CCAGACTGGACCCCTTTTTCCCTTAGGGAATTAAGAAATAGACAAACTTAATGCATC
ACTCTGACATTACTGTTTTATATTGTAACTTAGCTATATATTTGTGTGGTTACAAATT
GTTATCCTTTCATTGATCTATGTATTCTATGTGTGTTGTTAAACATGATGATTTTATTA
GTTATCTTGTATATTCCAAACTTTAGCTAGGTAGTTGATATAGGCCGTTGTCAAAATG
AATGGGCCATGGTGGGAGTAGAATATGCAACACTTGCGCGTAAAACCAAATTAAAC
AAATTAACTTAAGGTTTTATTTCTCTCTTTCTTTATTGAAACAAAGAAAGAGAAGGA
AGACCAAAGGGGTGGCGACAACTGTCAAGCCCCCCCCAGCTACACACACCCCGAGG
ATAAAACAGCAGGATCCTCCAGGACTCGCTCTTCTCGTTCTTCTTCTTCTCGTTCTTC
TTCTTCTTCTTTTACGTCTTCTTTTGCATAGCTTCTTTTACAATCCTTCGTCTTACAACT
TTTCTTACAATCTTAACTCTTACGCCTCGTTTCGCGCACTAGCTCACGTTCGTTTCTA
GCTTAAACTTTACTTCTTACACTTTGGTGCGCGATTTTTGCCTTACTATTTCAACAAG
CCAAGCCTTCCAAATTGAATCACACACAGGTGTGTATAAGTAATTTTGTCGTTTTTCT
TTTTGTGACAAAATCCAGCTTTTGTCTAGAAGATTTTGAAGGTTTTGCGTCGTTTGGC
CAACACCAACTCCTTCCAACTTATCTGGCAAATTAAGTCAAATACTAGCGTTTATTC
CGCTGTCAACCTCCGGCCCTCCGGCTACACGCCACACCAGGACCGGCTAACCAGATC

TCGAGAGTGCCTACGAAGAAAGTCTGCTGGCAGACTGCGCCGACGCTGATCCACCG
AGACATCGGCAGAGCCCGTCACCCCTGGCAACGGCCGTCGTCTCTGAGAGAAACCC
ACTTGGAGTGTGGCGTCCCCGTCGCCTAGCAACCAAACTCCGGGGCTCCAGCCAGG
ATCTCAACAGCGCTCCGCTACTCAAAGTAGGCAACTCTATCCTTTGCTCCGAACTGG
GTTGATGTCGGAATAATACTAATCCCGATTTAATAATTCTCCACCGTTACGAATAAA
TGCTATACACATATCTAAATCCATTCTCTTTCACCATCGCCCAAATCACTGCTAGTTC
AGCTGGTAGATATAGTTATTATTAGTTATAGATTCATGTCATTATCCTTGTTCATTTT
ACTGTGTTTAGTTAGTCAATAAATATTATCAACGTAACCCAGTCGTTGCCTGTGTTTC
TTTGAGGTGAAGATCGGAGATCCCTGTAATCGAGAAATCACAACTTTGGATGTGAG
ATTGATCAAAACAATCTGTACTTTAATCAAATTGTCCTCCCTGAGGACTGGTGCCCC
ATTCTTTTAACGAGAGCAAAATTCAAATTGTGGTATTCACGCTACAC
> Loki-Sri consensus sequence 5’LTR
TGGTGGACGCCAGTGACCAGACACCAGATTTATGGTAAGGACACTGGGGTGCTCTG
GAGGATATGGGGAGCCTCCCATGTCCACTTTCCCATCATCAAACCATCACATGGATA
GAAAACAGGGACGAAATGAGTTCCCCTCACCAATAGACTAGCTTTTGAAGATACCA
TCGGAATTTATGGCTAGCTCATCGAACTCATTCTGAACCAACACTCCTGCACCAAAT
GGGAGACAAACACAACTCTCCCCCACAGACTGACAATGAATTCATGTTTCATCCCTG
ACACTCATCCGGGATTGATAGGAGGGTGCCACCTCTTATCAAATAACATGAAAAGG
AACATTTAGCCATCAAGGTGCTAAACCACCAGGATTCACACTCCTGGTTCATCTTCT
CTCACAAGAGAAAACAACACTCTCACCTTTAACACCCTTTCTTACCCAAAGGGGGT
CTCACGAGGAACCACAGACCCCCGCCATAAAACCATCCTACATTCCAAGCCGGATT
GATGAACTATTCCACATTTTATAAAGCATGATTGAAATGTCATTGAACCAAACGTCC
TATCACCAGGTAATTCATGCAGTATTACTCTGGAAAGATTTCAACTAAGTATGATCC
GAAGTTCATTTCTGTAGCCCACATGGAGGCTGAGATAGGTTTTATTAGTTCTAAGGT
AAATATCCACCATGTTACAATACTGCCATCTAAGGACCAGACTGGACCCCTTTTTCC
CTTAGGGAATTAAGAAATAGACAAACTTAATGCATCACTCTGACATTACTGTTTTAT
ATTGTAACTTAGCTATATATTTGTGTGGTTACAAATTGTTATCCTCTCATTGATCTAT
GTATTCTATGTGTGTTGTTAAACATGATGATTTTATTAGTTATCTTGTATATTCCAAA
CTTTAGCTAGGTAGTTGATATAGGCCGTTGTCAAAATGAATGGGCCATGGTGGGAGT
AGAATATGCAACACTTGCGCGTAAAACCAAATTAAACAAATTAACTTAAGGTTTTAT
TTCTCTCTTTCTTTATTGAAACAAAGAAAGAGAAGGAAGACCAAAGGGGTGGCGAC
AACTGTCAAGCCCCCCCCAGCTACACACACCCCGAGGATAAAACAGCAGGATCCTC
CAGGACTCGTCTTCTCGTTCTTTTTCTTCTTCTCGTTCTTCTTCTTCTTCTTTTACGTCT
TCTTTTGCATAGCTTCTTTTACAATCCTTCGTCTTACAACTTTTCTTACAATCTTAACT
CTTACGCCTCGTTTCGCGCACTAGCTCACGTTCGTTTCTAGCTTAAACTTTACTTCTT
ACACTTTGGTGCGCGATTTTTGCCTTACTATTTCAACAAGCCAAGCCTTCCAAATTGA
ATCACACACAGGTGTGTATAAGTAATTTTGTCGTTTTTCTTTTTGTGACAAAATCCAG
CTTTTGTCTAGAAGATTTTGAAGGTTTTGCGTCGTTTGGCCAACACCAACTCCTTCCA
ACTTATCTGGCAAATTAAGTCAAATACTAGCGTTTATTCCGCTGTCAACCTCCGGCC
CTCCGGCTACACGCCACACCAGGACCGGCTAACCAGATCTCGAGAGTGCCTACGAA
GAAAGTCTGCTGGCAGACTGCGCCGACGCTGATCCACCGAGACATCGGCAGAGCCC
GTCACCCCTGGCAACGGCCGTCGTCTCTGAGAGAAACCCACTTGGAGTGTGGCGTCC
CCGTCGCCTAGCAACCAAACTCCGGGGCTCCAGCCAGGATCTCAACAGCGCTCCGCT
ACTCAAAGTAGGCAACTCTATCCTTTGCTCCGAACTGGGTTGATGTCGGAATAATAC
TAATCCCGATTTAATAATTCTCCACCGTTACGAATAAATGCTATACACATATCTAAA
TCCATTCTCTTTCACCATCGCCCAAATCACTGCTAGTTCAGCTGGTAGATATAGTTAT

TATTAGTTATAGATTCATGTCATTATCCTTGTTCATTTTACTGTGTTTAGTTAGTCAAT
AAATATTATCAACGTAACCCAGTCGTTGCCTGTGTTTCTTTGAGGTGAAGATCGGAG
ATCCCTGTAATCGAGAAATCACAACTTTGGATGTGAGATTGATCAAAACAATCTGTA
CTTTAATCAAATTGTCCTCCCTGAGGACTGGTGCCCCATTCTTTTAACGAGAGCAAA
ATTCAAATTGTGGTATTCACGCTACATATATTGGTGCCCCGTGTGAGGCGATTTCAA
ATTACCGTTCTTCACTACAAAGAAACCTGTTGAATGCTGAGAGTGGCCTC
> Loki-Sri consensus sequence 3’LTR
TGGTGGACGCCAGTGACCAGACACCAGATTTATGGTAAGGACACTGGGGTGCTCTG
GAGGATATGGGGAGCCTCCCATGTCCACTTTCCCATCATCAAACCATCACATGGATA
GAAAACAGGGACGAAATGAGTTCCCCTCACCAATAGACTAGCTTTTGAAGATACCA
TCGGAATTTATGGCTAGCTCATCGAACTCATTCTGAACCAACACTCCTGCACCAAAT
GGGAGACAAACACAACTCTCCCCCACAGACTGACAATGAATTCATGTTTCATCCCTG
ACACTCATCCGGGATTGATAGGAGGGTGCCACCTCTTATCAAATAACATGAAAAGG
AACATTTAGCCATCAAGGTGCTAAACCACCAGGATTCACACTCCTGGTTCATCTTCT
CTCACAAGAGAAAACAACACTCTCACCTTTAACACCCTTTCTTACCCAAAGGGGGGT
CTCACGAGGAACCACAGACCCCCGCCATAAAACCATCCTACATTCCAAGCCGGATT
GATGAACTATTCCACATTTTATAAAGCATGATTGAAATGTCATTGAACCAAACGTCC
TATCACCAGGTAATTCATGCAGTATTACTCTGGAAAGATTTCAACTAAGTATGATCC
GAAGTTCATTTCTGTAGCCCACATGGAGGCTGAGATAGGTTTTATTAGTTCTAAGGT
AAATATCCACCATGTTACAATACTGCCATCTAAGGACCAGACTGGACCCCTTTTTCC
CTTAGGGAATTAAGAAATAGACAAACTTAATGCATCACTCTGACATTACTGTTTTAT
ATTGTAACTTAGCTATATATTTGTGTGGTTACAAATTGTTATCCTTTCATTGATCTAT
GTATTCTATGTGTGTTGTTAAACATGATGATTTTATTAGTTATCTTGTATATTCCAAA
CTTTAGCTAGGTAGTTGATATAGGCCGTTGTCAAAATGAATGGGCCATGGTGGGAGT
AGAATATGCAACACTTGCGCGTAAAACCAAATTAAACAAATTAACTTAAGGTTTTAT
TTCTCTCTTTCTTTATTGAAACAAAGAAAGAGAAGGAAGACCAAAGGGGTGGCGAC
AACTGTCAAGCCCCCCCCAGCTACACACACCCCGAGGATAAAACAGCAGGATCCTC
CAGGACTCGCTCTTCTCGTTCTTCTTCTTCTCGTTCTTCTTCTTCTTCTTTTACGTCTTC
TTTTGCATAGCTTCTTTTACAATCCTTCGTCTTACAACTTTTCTTACAATCTTAACTCT
TACGCCTCGTTTCGCGCACTAGCTCACGTTCGTTTCTAGCTTAAACTTTACTTCTTAC
ACTTTGGTGCGCGATTTTTGCCTTACTATTTCAACAAGCCAAGCCTTCCAAATTGAAT
CACACACAGGTGTGTATAAGTAATTTTGTCGTTTTTCTTTTTGTGACAAAATCCAGCT
TTTGTCTAGAAGATTTTGAAGGTTTTGCGTCGTTTGGCCAACACCAACTCCTTCCAAC
TTATCTGGCAAATTAAGTCAAATACTAGCGTTTATTCCGCTGTCAACCTCCGGCCCTC
CGGCTACACGCCACACCAGGACCGGCTAACCAGATCTCGAGAGTGCCTACGAAGAA
AGTCTGCTGGCAGACTGCGCCGACGCTGATCCACCGAGACATCGGCAGAGCCCGTC
ACCCCTGGCAACGGCCGTCGTCTCTGAGAGAAACCCACTTGGAGTGTGGCGTCCCCG
TCGCCTAGCAACCAAACTCCGGGGCTCCAGCCAGGATCTCAACAGCGCTCCGCTACT
CAAAGTAGGCAACTCTATCCTTTGCTCCGAACTGGGTTGATGTCGGAATAATACTAA
TCCCGATTTAATAATTCTCCACCGTTACGAATAAATGCTATACACATATCTAAATCC
ATTCTCTTTCACCATCGCCCAAATCACTGCTAGTTCAGCTGGTAGATATAGTTATTAT
TAGTTATAGATTCATGTCATTATCCTTGTTCATTTTACTGTGTTTAGTTAGTCAATAA
ATATTATCAACGTAACCCAGTCGTTGCCTGTGTTTCTTTGAGGTGAAGATCGGAGAT
CCCTGTAATCGAGAAATCACAACTTTGGATGTGAGATTGATCAAAACAATCTGTACT
TTAATCAAATTGTCCTCCCTGAGGACTGGTGCCCCATTCTTTTAACGAGAGCAAAAT
TCAAATTGTGGTATTCACGCTACA

> Loki-Sri consensus sequence *smc-gag-pol* gene
CTGCTTACTGAACTGACGCTTTCCCTCTCCTTTTTCTCACTTCCTGATTTGGAAAGAA
CCACCCAACTCAGTAACCTAACCACCAGCATGTCTACACCCAGCAAAGTAGACTCG
GCAGCTGCTGCCACACCTATACCCAGAACCTCAAAACCTGAGGACTCCCTTGAGAA
CTTAGAGAGTGCGTTAGTGGACTTAACGGAGGATGTGCTACCCGGCTATGCCATTCT
CATCCAAAACACCGCCCCTTTAGAACTAGAACCAATAATGCTAAACTTAGTCACTGA
GGCCCTAACTAAACCCCTAAAAGATAAGGCGTTAACCAAAATGCTTGGCAACCTGG
CCATCGTGGTGACAGCCCGACTCAAATACGCAAACGATTTGTTGACACAAGCAGCC
GATAAACTCAAAGTTCGTGCACAACAGCTTTCACAGTTACAAGAAGGGGCCACATT
CTTTCAACATCAGCTAAACCGAGCCCACCGAGCCTCGGATGAATCTGAAACACGCA
TTGAAGAGCTCGCAGCTCGACTTGATGTGGCCCAAGGAAGAGAACTTGAGTTGGAG
AACAGAAATCAAGAGTTAAAACGGTTACTCCAAAACCGCGAACAAGAGTCAACTGA
ACTTAAAGACCAACTTAATGAGATAAGTAAAGAACGAGATGAATCACATGATCTCC
TTAAACGGGCAATGGAAGAGATAGATGACGTAAAAGCCAAGAATCAGACCTATGCA
GATCACCTGCAGGATGATCAGAAACAGATTCAACGGCTACAAGAGCATCTACACGA
AGCGGAACTCAGGGCCAAGGCATCTCGCCAGCAACAGGAAAGCAAAGAGGCAGAG
CTTGCATCAGTCAAGAAAGAGCTACAGCAGTCCTTTGAAACAAGACGCATGCAGCA
AACGGGCCCATATGAAATATTTCGGAAAAAAGATGACGAACCACCCATGAGGGGAA
CTCAAGGTGACGTCCTACCTCTTTTCTCACCTCTTGTGTCAAATCCTTTCGACACTGG
CAAAACGCTTCCAAAGCTGGCCACCACGACTTCTTCCCCGGAACCTAGGTTTCCCTC
AAGCACCCGCACCTCGGGTGTCTCAGATCGAGACCTTGACAAGATTGCTCGCAACTT
GCCACGCTTTGAACCCACTCCTGGTGGCTCCCAAAACATCAGCAGCTACTTGAAGGA
CATCGACTTCTACCTGCGAAAGTTCCCAAATGCCACGGCTGATGACCGGATATACTT
GATCAAAGTGACGTCCGGATGTCAGGTGGGTAGGTTCCTTGAACGACAGACAGACT
ACGTCAGAAGCAACTACGACCTGCTTTGCCAGGCATTAATAAGGGAATTCTCTGACC
CTATCTCACAAACAGGTCTAACAGCTGCTATGACCACCAAACAGGGGAGGCATGAG
TCTCCTCATCAATATTACCACCGTCTTCTTTTGGCTTACTTTGGCAACCGAAACGAAC
CTGGTATGGAGGAAGACTTGAATTTCAAAACTCTCTTCGTACAAAACCTCCATCCTG
GTACTAGCTGTCACTTAGGTGTCATGGCAAACCCGTACACCTCAACCATCCAACAAC
TACGTGAGCTTGCTACCCTTGGGTTTGCCAAACAAAAGCAATCCCACTCAAAACAGT
CAGAACCCAACATCGTTCTGACCCTCAACGCCAATGCTCCGCCCATGGAGTTAGAAG
GGGCCCCAAGCGGGGAGATACAGGGACCTGCCAACCAAAGCTCCAGAAGCCATCG
ATTTACAGAAGAACATAAAAGGTACCTACCCCCACACAACTACCACGAATCCTCAA
GACCTCCCTGGTCTGGTAAGGGACAAAACCAACAAAGATACGAGCCACGCCGTAAC
TCAGTGGATCACCGCCGAAGCAGGTTCCCACCCAATCAATGGGAAAATCGTAGAGC
GCGTGATGACCAGGGACAGCATCGATTCCGGCCCCAAGGTAGATACCCACAAGACA
GGCGACAAGTACAACACACTCGCCAAGACTCATCCGAAGAAGACACAAAGCAAAG
TGGTGACTCACAAAAATCTCCACGGAAATCCGACAAGACAGAAGGTAAGAACAAAA
CCTCTCGTCTAAGCGAGGAAGAGATGGAGACCATCAGGCGCCTCATCCAAAGAATG
CAGCTGAATAGCAACAAGGCTGATCTCCTATCAATAACCTCAGTAGCCACTGAGGCT
GGCCCTCAAAACAATTCCCCCGACCAACCTGATATGGTGGTTGAAGAATTGTACCTA
GCCCCTAGCTTACCTGAGTCTATGCCACCACAAGACCAGACGCTTTCGACCCCAGAC
GACAACGCTAACGAATCAGCTCAGGAAAGTAACAAATCGATGCCCCCAAGCGCAGT
CTTGGTGGTGCAGCTTGACCCCGAAGACAGGTCCCCTTGTGAAGAAGCCTCGGCCAT
CCAAGTCGAACCACCTGTTCAACAGTTTCTCTGCCACCTTGATGAGAAAGGGGTTGC
ACGGAAATTCTACATGTCTCTGACGCTTGAAGGAACGCTCAAACACGAAGCACTCCT

TGACACGGCAGCAGACATCACGTTGATGTCATTAACACTTTTCAATAAGCTACAGGG
AATGGCTCGCCAATCTAACAAAAACTTGAAACTCCAACCTTGTTCTCTGGAGGTCCA
GCCATACGCTCCCACCGGAACCACACTGACAACCATGTCAGCTATACAACTGACCAT
TGGACCCATGACTCTGGTCCATCCGGTGTATGTGTCTCCCTTGAACTCCATCCCCCTT
CTCATTGGCAAAGACCTCCTCAATCGATTTGAACCCCTAATCGACTTCAAACGCCTG
AAGATGTGGGCACAAGTGCGACAGCCTCTGCCTATTTCTCTTCCCCAAGACACTCAC
ATACAGTGTTGCGTCCTTGAGACAGAAGGGGTCAACACTTTGGCCCAGCGTGCTCGG
TCCAACCAACCAGACGTGACTCAGGTCAATGATGACCCCGAAGTGAGGACGGACAC
CGCCACAATACCAACGTGGCCACCTACAGACTCTCTGATCCAACGAGAGACATTCTT
GTGTGCATTAAATCCCTCAACTAATGCCAACGAATACTGCCCTCAAGTGGTCAATGG
CATAAGACTTGAAAACACTGACGTTGAGGACGTGACTCTAGCCCTCTGGGCCGACA
AATCTGCTATTAGCCAAGAGCTATATAACTCCCTCTTACAAGCAAACCCCACTCTAC
AGCTTATGAAGAAAACTTTTCGTTTTCCCTCAGACCCTACGCCACAGACTATGCTCA
CAGCAGAAGGTGTCTGCGCCTTGAAGCTACAGTGGAATCGGAGGGAGCTCAAACAT
TTCTTCCTGGTCATCCCAGACCTACCACATCCTACCTGCATAGGCAGCGACATTCTG
GTTAGGCTGGGTGCACAACTGGACACTGTCAACCATGTTTTGTGGTCACTCACCGGA
GCTCGAAACGACTCAAGCTGCCCTGACCCTGAGAACATGAAGTCGGGACAAACCAT
CCCCGACGCCTGTCAGATAAGTAATGAACATCACGTAACGGTACCACCATACACCA
AAAACGTGTCAGTACGCGTACACCTGCGAACAGGTCAACGGCTTAATCACCCACAG
GCACTCTTTCAACCATCTCCCCTTTTCTTCGAAATGGGACTCTCCTTGGAAGCCACTC
CCTTACTAGAGCTATGCTCCAGGTCAACTTACTTGTTGGTTCAAAATACCACCGCAG
AGACCATACTGATTCCCAGGTCCACACCACTAGGGTGGCTTATCAGTACAGAGTTCC
ACGATTTTGAGTTAGGAATCCCAGTTATTGGGCGTATACCCCCTCCTCTACTCCCTGA
CAACTCAGACGTGCAGGTGGTATACACTAAACCATCAAAAGCGATCGCGATCTTCC
CAGTCATGGCGTTTGAACACGACTCCATCGGACGTGTTGATCTCACGGAGGATAATC
AAATGTTGCTTCAGAGTATCAACGTTTTATCTCTGGATGTAACGTCAGAGACAGCCC
CTACACCACCTCAATCAGACATCCAAAACGAACCTGGAGAGGCGACACTGGAACAG
ACAACCGAGCCACGCTCGGACTTCATGGCTCAAATAGACCAGGTGGTGACAGAAGC
CGACGCACTGGACTGTGAAGAAGAACGCCAAAAACTCCGCAACATACTCCTGAAAT
ACCAAACATCCTTCGCTGAGGACTCTCTTGACTGCGGACTCACCACCATCCACTCAG
TGCGCATCCCCACTCCACCTAATGCACCACCAACCTTTGTACGACAATATAAGATTC
CTTTAGCATCGTATGAACCTGTACAGGAAATCATAGACAACCTGCTAGAGAAAGGA
ATCATCCGCCCTTGTAACAGCACCTACTCAGCACCGCTCTGGCCAGTGCTTAAACCC
AACGGCAAATGGCGGCTGACCATCGACTACCGTAAACTTAACCAACAGGTGCCTCT
GTCTAGATGGCCTATGACACAGCTTGAACAAGAGCTCCCTAAAATCAGGGATGCCA
AGTACTTCTCCACGCTTGACATCGCATCAGGATTCTGGACAATCCCAGTACATGTTG
CCGACCAGCACAAACTGGCATTCACTTTCGCAAATCGACAATACACCTTCACCAGGT
GCCCGTTTGGCTACGCCAACTCACCTGCTGAATTCAACATCTTCTTAAACAAAGCGT
GCCCTGATGCTAACGACAGGGGCACACTCATCTATGTTGATGACATCCTCATGAGAA
GTTCTACTCTGGACACGCATCTGGAGGAAATTGATCATGTCCTGGAGCAGCTCACAG
CCGCTGGCGCGAAGATATCACTCTCCAAATGCCAATGGTGTAAAACCAAAGTGAAC
TATGTTGGCCTCCTGGTAGGACCACGGGGGGTTGAGCCACAGATAAATCGTGTACA
AGGAGTCACAAACATCAAAGCACCAACGAATGTCTCTGAACTTCGTAGCTTCTTGGG
GGTATGCAACTACTCCCGCCAATTCATTGAAAACTACGCTGACTTGGCCAAGCCATT
AACAGAACTCCTTCGAAAAGATGCGCTATTTGTCTGGAACGATTCACAACAACAGG
CAATGCAGGCCCTTAAAGAAAAACTTTGCACGGCTCCTTGCTTGGCCTATCCGAACT

GCAACAAAGAGTACCACTTGGAAGTGGGATTCTCTAACCATTGCTTGAGCGCAGGC
CTTTACCAAAAGCATGACCTCGACAAACGCGTTGTTGCATATGCCAGCAAAACTCTG
TCACCCCCTGAACTTAAGTACACTGACTGTGAAAAAGCTCTCTTAGCGACGGTGTGG
GCAGTCAAACACTTCTCCAACTATCTTGGCGGACAAAAAGTGATTGTGGAAACCCAC
CACCAGCCAGTAACCTTCCTCAACAGTCAGCGACTCAGAGATGGCGTAGTCACTAAT
GCACGCGTAGCGTCGTGGCTGATGGCCCTACAAAGCTTTAATATGGAGGTGCGCTAC
GCTCAAAACCACAAAACACCCTTGGGCACAGGACTGGCAGCCTGCCAACGGTGTTC
AGACGACTCCCCGACCAGCGTAGCACCAGTCTGCGAACTGCCACCTCCAACACCTCC
ACGCCACCACTACTTTGACCAAAACATATGCAAAGAGATGACAACGGCATACGTGG
ATGGCTGCTCTTTCCACCATGAAACTCAGATCCGTGCAGGCGTGGGAGTTGTGTGGG
TCCATGACACCCCATATAAACCCCAGATGTTCAAACTGGGCTCCCAGTCTTCCCAGT
ATGCTGAAATTGCTGGCATCCTGATCACGCTGCAACTGGCTGTAGAAAAAGAAATC
AAAACGCTGGTAATCTGCACCGACTCCAACTATGCGCGACTCAGCTTCTCTTGCCAC
CTTCCCCTCTGGCAACGCAATGGCTTCATAACTTCCAACAGGAAACCAGTCAAGCAC
AAAGAGTTGTTCGTGGCCTGCAACCACCTCGTGGACACACAAGACATGCAGGTCTA
CTGGAAAAAGGTCAGGGGCCACTCCCGTGTTCCCGGGCAAGACAAAGAACTGAACG
ACCTAACTGACACACTTGCTAAACGTGGCGCTGTTGAAGGCACGCCGTGGTCACTGG
AACCAGAACCTGTGCCTGAGCCGGATTCCCTTGCCCTCCCACCCACCCCTAGCATCT
GTGTCATCACTCGGTCTCATGCGGCCACTTCAGCGCTGCCACGCCCAACTGACATAA
CCACTGCCTCTGTCCGTCCGGCTTTCTCTGATTCGGACCTCATCTCACTGCAAGCGCT
TGATCCAGCAATCAATACAATGATTGGCCACCTAACCGACCCAACTGCTCGGTTGAT
CACAAGCGATGAACTTGCCCAACTACCAGACCTGAAGATCCTCTTCAACGGACGTG
CGTCCCTGCGCCTGCTCAAAGGCATCCTGGTTTATGCTTCCGATGCGCTCACTTCACC
TGTGTTCGTGGTACCTCGTAACCACAGGGGGGTGATGCTAATGCATGCCCACGACTC
CCCATGCGCAGGACACAAAGGCTCCAAAGCGACATACAATGCACTCAGGCAGGTCG
CCTACTGGCCACAAATGCGAAAAGATGTGGCTGACTACATCAAAGGCTGCTTAGTCT
GCTGCCAGTTTCAACCAGCCAATCCTGTGCACAGAGCCCCCTTACAGCGAAGAGGC
ATTTCCTTCCCCTGGTCCGACCTCCAAATCGACTGGGTTGGACCTCTGACCAAATCA
ACCAGAGGCAACAAATACTCCCTCACGGTCACATGCGCATTCACTAAATGGCTCGA
GTGCTTACCAGCCCCAAACGACACAGCTCAAACAACCGCATACCTACTGATGAACC
ACATCTTCTCCAGGTTCGGTCTGCCGAGCCGCGTCAACTCGGACCGGGGCACACATT
TCACTGCGGATGTAATGCAAGAACTCTGGCAGATACTCGGTGTGAAGGCCAACTTCC
ACATCAGCCATCATCCCCAAGCCTCTGGTCAGGTAGAACGAGCTAATCGAACCGTG
GTTAGCGTTCTAAAGAAGTACGTGACTTCAAACCACAGGGACTGGGATGTGAAACT
GCCGCTGGTCTTAATGGCCATCAGAGCAACACCGCATGAAGCAACGGGGTTTTCGC
CATTTGAAATGATGACAGGCCGACAAATGACACTCCCACTGCACCTGCTTTACCAAC
CAGGTGAAGCTAGCGTTGCAACAGCATACACAACCCACCAGTACATGACGGACTTG
AACCAACACCTTCAAGCCACCTTCGCCTTTGCCCAAAAACACCTGGACAAAAGCGC
AGAAGGACGTAAGGCCTACTATGACCAAAAGGCTTCACATGAAGAACTCCAAGTGG
GGGACAAGGTGTGGTACTACAATTTCTTCCAACCAACAGGTCAAACCAACCCAAGC
ACGGGAAGGCTCGCCCGAAAGCTCCTCCCTCGGTGGGCCGGTCCCTACCTGATCACA
GACAAGCTTTCCCCCGTTGTGTACCAAATCAAGGTTAGCAAGGGCCAGAAAGAACC
GACCCTCAAATGGGTCCACCGCAACCAGATCAAACCACATAAGACCCCCATGGGAT
ATACAGGGGACGACAGCACCCCAGTCAGGTCATAA
> Loki-Sri consensus sequence Smc-Gag-Pol polyprotein

LLTELTLSLSFFSLPDLERTTQLSNLTTSMSTPSKVDSAAAATPIPRTSKPEDSLENLESAL
VDLTEDVLPGYAILIQNTAPLELEPIMLNLVTEALTKPLKDKALTKMLGNLAIVVTARLK
YANDLLTQAADKLKVRAQQLSQLQEGATFFQHQLNRAHRASDESETRIEELAARLDVA
QGRELELENRNQELKRLLQNREQESTELKDQLNEISKERDESHDLLKRAMEEIDDVKAK
NQTYADHLQDDQKQIQRLQEHLHEAELRAKASRQQQESKEAELASVKKELQQSFETRR
MQQTGPYEIFRKKDDEPPMRGTQGDVLPLFSPLVSNPFDTGKTLPKLATTTSSPEPRFPSS
TRTSGVSDRDLDKIARNLPRFEPTPGGSQNISSYLKDIDFYLRKFPNATADDRIYLIKVTS
GCQVGRFLERQTDYVRSNYDLLCQALIREFSDPISQTGLTAAMTTKQGRHESPHQYYHR
LLLAYFGNRNEPGMEEDLNFKTLFVQNLHPGTSCHLGVMANPYTSTIQQLRELATLGFA
KQKQSHSKQSEPNIVLTLNANAPPMELEGAPSGEIQGPANQSSRSHRFTEEHKRYLPPHN
YHESSRPPWSGKGQNQQRYEPRRNSVDHRRSRFPPNQWENRRARDDQGQHRFRPQGR
YPQDRRQVQHTRQDSSEEDTKQSGDSQKSPRKSDKTEGKNKTSRLSEEEMETIRRLIQR
MQLNSNKADLLSITSVATEAGPQNNSPDQPDMVVEELYLAPSLPESMPPQDQTLSTPDD
NANESAQESNKSMPPSAVLVVQLDPEDRSPCEEASAIQVEPPVQQFLCHLDEKGVARKF
YMSLTLEGTLKHEALLDTAADITLMSLTLFNKLQGMARQSNKNLKLQPCSLEVQPYAPT
GTTLTTMSAIQLTIGPMTLVHPVYVSPLNSIPLLIGKDLLNRFEPLIDFKRLKMWAQVRQP
LPISLPQDTHIQCCVLETEGVNTLAQRARSNQPDVTQVNDDPEVRTDTATIPTWPPTDSLI
QRETFLCALNPSTNANEYCPQVVNGIRLENTDVEDVTLALWADKSAISQELYNSLLQAN
PTLQLMKKTFRFPSDPTPQTMLTAEGVCALKLQWNRRELKHFFLVIPDLPHPTCIGSDIL
VRLGAQLDTVNHVLWSLTGARNDSSCPDPENMKSGQTIPDACQISNEHHVTVPPYTKN
VSVRVHLRTGQRLNHPQALFQPSPLFFEMGLSLEATPLLELCSRSTYLLVQNTTAETILIP
RSTPLGWLISTEFHDFELGIPVIGRIPPPLLPDNSDVQVVYTKPSKAIAIFPVMAFEHDSIGR
VDLTEDNQMLLQSINVLSLDVTSETAPTPPQSDIQNEPGEATLEQTTEPRSDFMAQIDQV
VTEADALDCEEERQKLRNILLKYQTSFAEDSLDCGLTTIHSVRIPTPPNAPPTFVRQYKIP
LASYEPVQEIIDNLLEKGIIRPCNSTYSAPLWPVLKPNGKWRLTIDYRKLNQQVPLSRWP
MTQLEQELPKIRDAKYFSTLDIASGFWTIPVHVADQHKLAFTFANRQYTFTRCPFGYAN
SPAEFNIFLNKACPDANDRGTLIYVDDILMRSSTLDTHLEEIDHVLEQLTAAGAKISLSKC
QWCKTKVNYVGLLVGPRGVEPQINRVQGVTNIKAPTNVSELRSFLGVCNYSRQFIENYA
DLAKPLTELLRKDALFVWNDSQQQAMQALKEKLCTAPCLAYPNCNKEYHLEVGFSNH
CLSAGLYQKHDLDKRVVAYASKTLSPPELKYTDCEKALLATVWAVKHFSNYLGGQKVI
VETHHQPVTFLNSQRLRDGVVTNARVASWLMALQSFNMEVRYAQNHKTPLGTGLAAC
QRCSDDSPTSVAPVCELPPPTPPRHHYFDQNICKEMTTAYVDGCSFHHETQIRAGVGVV
WVHDTPYKPQMFKLGSQSSQYAEIAGILITLQLAVEKEIKTLVICTDSNYARLSFSCHLPL
WQRNGFITSNRKPVKHKELFVACNHLVDTQDMQVYWKKVRGHSRVPGQDKELNDLT
DTLAKRGAVEGTPWSLEPEPVPEPDSLALPPTPSICVITRSHAATSALPRPTDITTASVRPA
FSDSDLISLQALDPAINTMIGHLTDPTARLITSDELAQLPDLKILFNGRASLRLLKGILVYA
SDALTSPVFVVPRNHRGVMLMHAHDSPCAGHKGSKATYNALRQVAYWPQMRKDVAD
YIKGCLVCCQFQPANPVHRAPLQRRGISFPWSDLQIDWVGPLTKSTRGNKYSLTVTCAF
TKWLECLPAPNDTAQTTAYLLMNHIFSRFGLPSRVNSDRGTHFTADVMQELWQILGVK
ANFHISHHPQASGQVERANRTVVSVLKKYVTSNHRDWDVKLPLVLMAIRATPHEATGF
SPFEMMTGRQMTLPLHLLYQPGEASVATAYTTHQYMTDLNQHLQATFAFAQKHLDKS
AEGRKAYYDQKASHEELQVGDKVWYYNFFQPTGQTNPSTGRLARKLLPRWAGPYLIT
DKLSPVVYQIKVSKGQKEPTLKWVHRNQIKPHKTPMGYTGDDSTPVRS*

> Loki-Sri consensus sequence *env* gene

CTGTGTGTTTTTCCCATCAGGATGACGTGGCTGCTGCTCGCCATCGTGACTCTCGGGA
TGATAACTACCCCGGCGAACCCCGAGACCGTCGAACCTGGGCCTAACTCGGGGATT

```
GTTCTAAAGGAACAACCAGGACTTCTAATCACAAATTGTCGTCTACACACCCAGAAA
GTGTTCGTTAGATTCAACCCTGAGATGGTTTGCCGCCGAAATGTCCCCACCACAGCC
AAACTAACCTCTTGGGCTGGCATTCGCTGGAACAAGGAGGTTATAACCCACGCAGA
GGTAGACATAACCCACATGCTGCAACAGCTTCAGAAATTCACCATCACTCGGTCTGA
ATTAAGCGGTGCCAACAAAAGAGAGAAGCGGTTCATTGGAGGCTTGCTAACCGCAG
CTGCTGCTCTCGGCTCTCTCTTTAGTCTAGGGGTCTCCTCAGTTAATGCTGTAAACAT
TGCCGCTGTCAAACGACACATGGGGGAGCTGCAGGCTGAAATCCCCGGGATTCAGG
AACAAATGGTAAAACAACAAGCACAGCTAAAAACTATTGGGAAGACGTTGCAGGG
AACAGTCCTAGTACTTAACACTCACAGTGAAACCTTGCTCAAGACTTTAGGAGCGGT
AGACCTGTTACTGTCAGTCCTGCAGGCAGACTATGCTCACACACAACTGGTGACTAT
GCTAATGTCCGACATGCTTAGGGATATTGGGTCATCAATTGATAGTCTAGCTATAGG
TAGGATCCCTCCCTACCTGGTTCCACTCTCCCTGGTACAGGACATCCTTGCCACTGCC
ACTCGAGACATAGTGACACCACTACAGGCTCATTTAGCTTACACCCTAGGTAGCTCA
GTACCAATCTATGTAGACCCAGAGGCCAGAGAAGTAGCCTTTATTATAAATCTACCC
ATAGTGACACCTGAGAACATCTACAGACTGAAGGACGTAGTTAATGTTGGCTCTTGG
CAGGGTGAAACGCACGTCAAAGTGCACACACCACCCGTTGTGGCATACCACGACAG
CAACCCTGACTTATACCTTGCTCCCAACCTCCGCATGTGTACACTGACAAAGACAT
CCACTACCTCTGCCCCAGCAAACCATTCGTACGTGATAACACAGGTGGTATCTGCGG
CATTCAAACCATGTCTAAAGACTCTCAATGCCCAACAACAGTCACGCCCCTCTACCA
AATGACTACGACTCAGGCAGAAATAGTGGGAAGTCGCTGGTTAGTTAACACCCCAG
CGAAAATAGCAACTCTCACTTATGACCGACACGACACCTCCACCCGTATCGAACTAC
CTCACCGCACCCTGTGGGTTACCATGCCTGAGGGTGCCATCTTTCACATAGACGACT
TAGCCCTCTACCATCTAAACCCTGACCAGTATGAGAGTAACATTGAGGTGTCCGACT
TCTTCAACAAACATGCCCTCGAGTTGGATTCCGAGACCCTAACCCAAGTCCAATACG
AGGGAACTCAGACCATAGACGTCGGCCCTGTTGACAACGTTCTAAAAGAAATTGCC
TCCCAGACCCATACTCCTAGCCAGCCTTTAACCTATGCCTGGTCTACTCCAGACACT
GCCATAGCCGCCCTTGTTGGCTTAGGATACATCCTGACCTTTGGTTTGACGTGCCTAT
ACATCAGGCGCATGAAAACTTGGCAGTGTCACGTAGATCAGGGTCTGGACAGACTG
AACAGAGGAATAAGAATGAGGAAAAGAGAAACGAAACCAACCACTAATGACCTAG
TGACTGTTGACGATGAGACAACCTTCACCCGAAACCCTTCCAGTGGCCCTGCCATCC
TGGAACTATTATGCGACTAG
> Loki-Sri consensus sequence Env protein
LCVFPIRMTWLLLAIVTLGMITTPANPETVEPGPNSGIVLKEQPGLLITNCRLHTQKVFVR
FNPEMVCRRNVPTTAKLTSWAGIRWNKEVITHAEVDITHMLQQLQKFTITRSELSGANK
REKRFIGGLLTAAAALGSLFSLGVSSVNAVNIAAVKRHMGELQAEIPGIQEQMVKQQAQ
LKTIGKTLQGTVLVLNTHSETLLKTLGAVDLLLSVLQADYAHTQLVTMLMSDMLRDIG
SSIDSLAIGRIPPYLVPLSLVQDILATATRDIVTPLQAHLAYTLGSSVPIYVDPEAREVAFII
NLPIVTPENIYRLKDVVNVGSWQGETHVKVHTPPVVAYHDSNPDLYLAPNLRMCTLTK
DIHYLCPSKPFVRDNTGGICGIQTMSKDSQCPTTVTPLYQMTTTQAEIVGSRWLVNTPAK
IATLTYDRHDTSTRIELPHRTLWVTMPEGAIFHIDDLALYHLNPDQYESNIEVSDFFNKH
ALELDSETLTQVQYEGTQTIDVGPVDNVLKEIASQTHTPSQPLTYAWSTPDTAIAALVGL
GYILTFGLTCLYIRRMKTWQCHVDQGLDRLNRGIRMRKRETKPTTNDLVTVDDETTFTR
NPSSGPAILELLCD*
> Loki-Ofa consensus sequence
AAGTGTTCAGCGTTCCACACTTAACATTTTCACCATTATGAGTACACCATCCGGGTA
CTCATAGTGCACTGCATTTTGCCATAATTGTCAGTGTGAACACACTTATGCACTCAA
```

AACAGCGAGTGTAAGTACGGAAGAACACAATTTGAGACACAGCACAAGACTCTAGA
ACCATATTTTGCAAAGGTACAAGGATTCATTTTGCTGTACCACTTTTTGGTGTGACCA
GGCCAATGTTGCCAACTCTTTACCAGTGAAAGTAGCTAGCACTAGCTCAAAAAGTTG
CTTGTATATTTGAAGTTAAAAAGTTTAAAGAAAAAATGTATGAAGGAGGGAGAAGG
TCAAACAAACTATTTACATATTTACAAGTTTTAGGAACACCCTAACCCAGAGATCCA
TAAAGAAAAAGTATTTATATACGGAAAATATTATTCTATATCGTGGTATGAAAGCTT
AGAGCATACTATAGCACACCACTTTCTCGCGTTTTGAAAAGTTATGTTGGGACCTCA
CCACGTGGGTGAACAAAGGAGAGGCCTACATGTAAAGACAAAAGCCTGAAACAAT
GAAACGGCGCCAAATCCCATCAGGCCTTGCACTTCACACCTTGGTGTGAAGTTTGGT
TTAAATTCAAACACTGGTGTTTCATTGGATGAAACTCACGGGAACCCCAGGGAGTCC
CCGTGACGCGGCCATGACGTCACCAGGAATATAATAACCGGCCCGCACCTCTGCAC
GGGGCCGTCCGCTGTGACCCCGAGTCCGACAGGAGGTGACCCAAACGTTTGTATTTC
AACTAGCCGAGGCAACAGTCCCTAACTCGCTGGCCGAGAGCAGACTGTTTTATTTTG
TTTTTGTGTTTTTGTTTTACATCCGGACCCGGAAACAGATCGCCTGCAACGAGCCGAT
CTTGTTTCATTACCTGCAGAAGGAAGAAACATTCTTCAAGAAAGAGAAGGATAACTT
CTACGCCCCGCTCTCTTTCAATTAAGTCGACTCTCGCTGTTTCCTCCGCCTGCCATAC
GAGGACCAGCTGATTCGCCGTTCAACCCTCCACAAAGTGAGGTAACGGCTCCGTCAT
NTGTTCGAGTAGCAGTTTAGCACACTATTTGGTTTTGAGATCTGTAAGGCAGCCTAA
CACAGCTATTACTCATCAGAGGGACCGAGGTCACCTCACTAACATCCTAACAGGAT
AAGCTATACGAGAGTAATAGGCACCTAACACATTGTAAGCCACAGGTTGAAGCCTC
TCGCCTGTGCAAACGCTTTTGTATTGCAGCAACTTTTCAGTCACGACAAGTGATTAG
AACCACTGGGTGTCGCCCTTTTACTAGAAAAGTAAAATGAGTAGCCAAGACCCCCGT
AGATCAGGAGTTGAAGCTCCCCAGGAGTCAGACGATCTATCCCTGCAGGATGCGGT
CGCTGGCCTCCTCCACCTCTCTGTAGAGGACCGAAGTAACCTCGACAGCTACACTGT
TGACGCATGTAATGAGAAATTGAAGGAGCTGGTGGCTTATGTCAATCACCCGGCCG
GGAAGCCACAGCGCACCATCTCTGACGTGCTTGGTGAAATACTTCTCCTCTCTCAGC
GCAAAGCCCGACTGCAAGCTGCAGAGCATAAGGTGCAGCTTGCCCAGCTGCAGCGG
AGCAATGCCAGCCTGCGACAGGAAGTCCAGGCCGCACAAACTGAGATCTCCCAGAT
ACAGGATGATCATGATCAGCTGCAGGGGTATGCAGAAACCCTGCATGCGCAGCTCC
AGACTGCCCAACAAAAAACAGAGGTCAGGCGTGAAGAACAGATCGGCCAGGTAAG
TACCGTTGCCGACCCCTTCGCCCAAGAGTCACTTGAAGATGCCCCTGACCTTGAAGC
GGCTCCTCCTCAGGTCGCAGTCAGCCGACCCGTGCTCATAAGAGCACAGCCGACCTG
GACGAGTCTGACCGACTGCGCTCAGAACTTCATGCCGCCCGCANTGAGAGAAACAG
CCCTGCTCCAGAGACTTGAGGTTGCCCGACCGAGTGACCAAACCACTCCGGCCAAC
CCAGGAGCCATGACTGGGAACTTAGCGGCCCACAGCAGCGCGTTCTCTGAGCTGGT
CACAGTTGACAGCTATCAAGCACCTCTATCGGCCCGCTCATTCGACATGGGTGCTCC
ACCTCCCTTTAGCCGCCCTACCTTGCGTACTGACAGGAGGGGATTCCCTCCCCTCCG
GGGTGGGGATTACTCCCAAGTCCCGAGCGGCATAAGTGGTTTTAAGCCCTCATCCAC
TAGCGAGCAATATACATCGCGCAGCTCGCGCCTGCGCCTCAATAGGACTCCATCCCC
ACCGCCAAAGGGGGGAAGTCATCCTCACCGTTATATTGCATTGAGTGACCATGGGA
CGTCAGATGACTCAGACGGCCCATGGCCGCACCGAGAACAAGGCCTACGTACTCGC
CAGGTCGAGTCCCTCGCCAGGGATGTAGAGCGCTTCGACCCGGGCAATCGAGAGTC
AAACATCAATGACTACCTCAGAGAAGTAGAGCGTTGTCTCCTTGACCTACCTCACGC
TTTGTCACGTGAGAAACTGAAGCTTATCTGGAAGACCACGGCTAGGAGCGTCCATGT
CTTCATGGAAACGCTCCCACCTGGGACCAGAGACCGCTACTCAGCACTCTGCAAGG
CCCTACGGGAGGAGTACTCATTGTACACCGACCGAGCTTCCGCCACTCTCGGCGCCT

TCGCAATCATGCAAAAGAAGCATGAACCTCCAAGGGAGTATTATAGGCGTCTGAGA
GCTGCGTACTTCCTGGGACGTAATGCCCCAGGCCTGGAGGAGGATCAGGCTTTCAAG
TCCCTGTTCCTCCACAACCTCCACGAGAGTGTGCGCTATGATATAACTATGCACTGC
AGGACGCATGATCTCACCATGCCGGAGATCAGGAAATATGCGCAGATGGCATGGGA
GACACGTGCACGCCCTGTTAGGAGGTACGAAGACGACGCCAGGGTCCTTGGGATCC
AAGCCTCGGCAAACACAGACTTAGCGCTAGAAGGTGACGAAATCCCTCGCGTTAAG
ACAAACACCAGAGCAGTACCTCAGGGTCGACGGCCATTTCGCCAACAGAACTGGCA
ACAGAACCAAGGGGGTGAGAGACAGCATTACTCCCAAAACCAGGTCAGGCCTCGAC
AACCAAACACCCCTCAACCAGGAAGGCAAGGTTGGTTCGAGAGAAACCTTAAGCAA
GGAGGTAAGTATGATGGTAAGATCCCAAACAAACAGAAAGGAGATGGTTCTAAGCA
GGACATGGAGAGCCTGATTCGGAAATGCTTGGCGGAGGCAATGAAACAGCTCGAGA
CATCCCGTCATTCACCAATGACGCCGGACCCTTCAAAGGGACCTGATACCAAAACCC
CGCCCTTATGACTAGGCCCGGCGACCTTCTCTTCGACCGCCAAGGTCTACTTTGTCG
GTTGGGGAAGAGAAACAAAGGAGCCCAGGACATTCCCCGAAACAGCCACTCCTCTC
CCTTTACACACTCCTTTCTTGACCTTCTTAGGTGATCTCATCCGCAAGGGCAACGCCC
AACGGATATACACCTCAGTAGTAGTAGGGGGGTGTGTCAGCACCCTCGCCCTGCTAG
ACTCTGGTTCTGAAATAAGCTTGATGTGTTCCAGTACCTTCGCTGAGGTCGCCAGAG
CGATGCTCTCCCTCGGCAAGCCGTTACAAGTCGAGACCTGTAACGTCAACATCACCA
GCTACACACAAGATAAATCACCCATCACGAAACGTGCTTGGTTTGATATTTCCTTCC
AGGAGATGACACTAGCGCATCCCATCTACATCTGTATCCTCGATACAGAACCTTTAC
TTATCGGTCAGGACCTCTTGGACCGACTGGCCCCACTCATTGACTGCCACCGTGGAC
ATATGTGGGCTCAGGTTGACACCCCAAAGCCTCTTGGCTCCATGGGCTGTCCGCCAG
CTCTGAACGGACAAGTCGCCGTTATCCAAGAGTCCAGAGGACCCGTACCGATGCTC
ATGCCTCTCCCAGAGCCTGATCCTGACTCTGAGCCTTCTATGCCACTGGCAAGCTCA
CAACCGCTGCAGATTGGAAATGCAAACTCCTTCCACGCTCACGAGTCTTTCCTTTGC
TCGCTGAGAAACACAAACTCAGCACTATACAGCCCCCATGTGGTTGGAGGCATACA
CCTCAATGGCATTACGGCTCCAGAGGTAATACTAGCGCTCTGGTCTGAGAAGTCTGC
CGTCAGCCAAGACCTGTACAACAGCCTGAGTCAGCGCAACCCTCGTTTGGCTTTTGT
GCAAAGAAGCCACCGCCTTCTTTCTGCAGATTCACCACAGACTCTCCTCAAGGCAAT
TGGGGTATGCGTTCTCTCCATACAGATTGGAAAAAAACAAGTCACTCACTTTGTGAG
TGTTGTTCCTCGACTCCACCCCCCATTCTCTATAGGGGCGGATCTTCTTGTCAGGCTC
GGCGCCCAGCTGGACACGGTGAACCAAATCCTGTGGTCGCAGGCCAACACAAAGGA
CTACCCCTTAACCACCGAGCCAGAACACATGCTATCAGGGCAGACCATCCCTCAGG
CATGCCAGGTCGCCAGTGAAGTCGACGTGACTGTTCCTGCACGAACAGCAGGGGTT
CCCATTCGCCTTGCGATCCTTAAGGGACAGAAGATGCCTGACCTGCAGGCATTCTTC
CAACCTCTACCCCATTTCTGGGAACTCAACCTGACAGTCTGCGGTACACCCCTGCTT
GAGTTGAACAATCGTTCAACCTATTTACTGGTTCAGAACCTGACGCACGTCCCAATT
CAGGTGAGGGCACGGCAACCACTAGGGATGCTGGTCGACAGCTCTTTCCACGACTTT
GAACTGACAGTGCCTGTCATAGGAGATCTGCCCCCATCCTTAGCCAGGGATGACAA
CCTGGACAATGCCCTTTTCACTTTTCCGACCAAAATGACCACCATTGCCCGTCACGA
AATGCTACAGAACGAAGCGATCTGCGGCACCGCCCTGGGCACTGAAGGTGACATGG
TCGTGTACGCTCTCGTGACTCAGCCAGGAGACACGACCTCAGGTAGCAACGGAACC
CGTTCCTCACCAGAGGAACCGTACCCCGGCTTTGAGTCCGAAATTGAGCAGCAGCTT
GCCAAAGCAAATGCTCTGACAACAGAGTACCAGAGAGATGCACTTCGGAAGTTATT
CCGTGATTTCCAGTCTATCTTCTCGAGGGACTCTAACGACTGTGGAGTCACTGACCTT
CACACTGTTCGCATTCCCACTGACCCAGAAACACCACCGTCATTTGTACGCCAGTAC

AGAATCCCGTTGGCTGCCTATGAGTCGATCCAAGAGATACTCGATAACTTGCTGCAG
AAGCAAATTATCAGGGAATGTAACTCAACCTACAACTCGCCAATCTGGCCGGTGTTA
AAACCAACTGGTAAGTGGCGCCTAACAATTGACTACCGCCCTTTGAACAAACAAGTT
CCACTTTCACGCTGGCCCATGATCCACCTGGATCAAGAACTGGCCAAGGTAAAAGG
GGCCTGCTTCTTCTCAACGGTGGATGTGGCCAATGGTTTCTGGACCATGAGAGTCAA
TCCGGCTGACCAATACAAGCTGGCTTTCTCCTTTGGCAACCGCCAGTTCACCTGGAA
CCGTTGCCCATTTGGATACTCGAACTCCCCAGCTGAGTTCAACATTTTCCTCCACAA
AGCGATGAGTGATGCCGCTGCCCGCGGAATCTCATTTACGTGGATGACATCCTCAT
GAGGAGTCGAACCTTGNAAGAACACCTGGCAGAGATCCGTCATGTCCTCAACCAAC
TTTCTGCCGCCGGAGCCAAACTGGCCATCGCCAAGGGCCAATGGTGCCGTACCAAG
GTGGAATATGTGGGTTTGACAGTGGGTGCTGATGGCATCGAACCCCAGGCTGGGAG
GATTCGAGCCATCCAAGACATCAAAGCCCCGACGGGCGTGTCGGAACTCAGGAGCT
TCCTGGGTGTCTGCAACTACTCCCGACAGTTCATTGAGGACTACGCGGAGATAGCCA
GACCCCTCACGGAGCTCCTTCGCAAAGACAAGTCCTTTGAGTGGGGAGAACCTCAA
GAGCAGTCCTTCCGCAAAATGAAGGAGAAACTCTGCTCTGCATCTTGTTTGGCTTAC
CCAGACAAGGACAAAGAGTTCCACCTTGAGGCAAGTTTCTCTCCCCACAGCCTCAGT
GCTGCCTTGGCGCAGAAATACGACACAGACAAGCGGGTCGTCGCCTACGCGAGCCG
ACCTCTAAGCAGCGTCGAGATTAAATTCTCTGATTGCGAGAAAGCCCTCCTTGCTAC
CGTCTGGGCTGTAGAACACTTCCGCAGCTACATCGGAGGCCAGAAGGTAATCATCG
AGACTTGTCACCAACCTGTCACTTTCTTGAACAGCCAGCGGCTGAGAGAAGGACGTG
TATCTAATAGCCGCATCGCCTCCTGGATGATGGCACTGCAAGGCTACGACATCACGG
TCAAGTATGCTCAGAACCACAAGATGGCCTTGGGCCAGGGCCTGGCCGAATGTCAG
CATTGCGACTGTGAGGAGCAACCCAACTCCCAGCCCCTTCTCTTAACCACTCCATCG
CTTCCTTCCAATCATCATTATTATGATGAGAACGTTTGTCAAGACCTTCCAAGGGTCT
ATGTGGATGGCTGCTCGTTCCACCATGAAAGCCAGACACGGGCTGGAGTTGGCATTG
TTTGGGTCAACCACAAGGTTGACGAACCAAACCACTACCAGTTGGGCCCTAAGACA
AGTCAGTATGCAGAGATCGCTGCAGTGCTGATCACCCTCCAACAAGCTGCCAAGCT
AGCCATCGTGCAGCTGGTGATCTGCTCCGACTCTAACTACGCTCGTCACAGCTTTATT
TCCCACTTTCCCGTGTGGAAAGAGAACGGCATGAAAAACGCACGGAATAAAGAAGT
CAAACACTCAGAACTCTTCTTGGCTTGCGACCAACTTGTGACGGACCGAGGGATGAC
CGTGTACTGGAAGAAGGTCAAAGGTCATTCCCAAACCGCAGGTCCCGACAAGGATG
GTAACGATGAAGCCGACCGCCTAGCCAAGCTAGGTGCGGAACAGGGCACGCCCTGG
GAGTTCCAGAGTGATTGGCTCCCAATTCCCCAAACGTGTGCAGTGAATGCCATTACC
CGCCGGCAGGCAAGAGAAAAACGAGAGGACCCTCAAGCCTGCACCCAAACTCTACA
CCTGGGTCGCGCACCCAGCGATGCAGATTTAGTTACTATGCAGGAACAAGACCCTG
CCATTCAAGCCATCCGTCAGTGGATGTTGAACCCTCCACGGCAAGGGGTCTCACAAA
CCTCACTGGTCGAGTCAGAGGAGCTGCGAACTCTCCGCTCCGACCAGCTTCACCTCA
AACTGGAAAAGGGTCTGCTCGTCTACACCTACAACGGCCAGGGACCGACCCGTTGG
GTCGTCCCAACCAACCATAGGGGGGTGATGCTAGCCCACGCTCACGACTCTCCCGTC
GGAGGCCACCGGGGTTACAGAGCCACTCTCCAAACCCTCCAGCAAGTCGCCTATTG
GCCATTGATGGCTCGTGACACTAAATCCTACGTCCAAGGATGTCTGGTCTGCTGCCA
GTTCCAGCCGTCCCGACCTCTGGGTCGTGCCCCACTCCAGAAAAGGGGGGTCACATT
CCCATGGTCTCACCTCCAGGTTGACTGGGTCGGACCAGTCCCCAAGTCATCACGAGG
AAATAAATATCTTCTCACTGTCACGTGCGCCTTCACCAAATGGGTGGAATGCCTCCC
GGCACCCAACGATACTGCTATGACGACAGCAACCCTCCTGCTGAACCATGTGTTTAG
CCGCTGGGGGCTACCCCTCTCTGTCGACTCAGATCGAGGCACCCATTTCACGTCTAG

TGTCATGACCACCCTGTATGACATCCTGGGAGTGGAAGTAAAATTCCACATCTCGCA
TCATCCCCAGTCATCCGGCCAGGTCGAACGCGCCAACCGCACCATTGTTAGCATGCT
CAAGAAGTATGTCAGCAGCAATGGCAAGGATTGGGATGTGAAACTTCCTTTGGTGCT
GATGGCCATCAGGTCCACCCCGCACCGCTCCACAGGAGTCACACCCTTTGAAATGAT
GACTGGAAGAGAGATGACCCTTCCACTACATCTCTTGTATCATCCTGAAGATGTCAG
TGTTGCAACCGCCTATACGGCGCATCAATATGTGACGGACTTGCGCGACCACCTTCG
AGCCACGTTCGCGTGGGCCCAGACAAACTTGGAGGCCAGCGTTGAAGGATCAAAAG
CCTACTATGACAGGAAAGCGTCCCACCGCGAGTACCAAGTAGGCGATAAAATCTTG
TACTTCAGATTCGCCAAACCGGTAGGAATCTCCAAGAAGTTCCTACCGAGCTGGTCA
GGCCCGTTCGAGATTGTGGGAAAACTCTCACCGGTAGCTTATCGGATCAGAGTCTCG
AAACCAAAGCAACCACCAGCATACAAATGGGTCCATGCAAATCAAATTAAACCCTA
CCAGTCATCCTCCCTCTCAAGGGGAGAGGAAGCCTCTACACCCGAAAGGGGGGAGC
AGTCTTCACTCCAAGAGGGAGAAGACTCAGTCCAGTAAACCCATAGGGATAGCTGC
ATGCCTCAACTCACCCCGTCGGCATGCACTAACTCCTAATACGTTCATCCCCCTTTTT
CCCTACAATAATTCAACTCTGGTTTGTTTCAGGATGAAACTCCTTTGGATCTTCGGCA
TTCTCATCAGTCTACGAGTCACCCTGTCTATGGAAGTGGTCGAACCCGGCCCACCGA
CCGGCATCGTCCTACAAGAGTCACCCGGACTCTTAATCACCCGTTGTGGACTCTACA
CCCAGAGAGTCTATGTACGCCTCGATCCTTGGAATGTGTACCGCAAACACGTCCGCT
TACCCCCACGACTAACGGAGGGAAGGTTAAGTGGAATCCAAACTCATGACACAGTT
GAACGTGCTAAACAGACAACTGTCCATGTCCTCGAGCAATTACAGAAGTTCCTGGTC
ACTGAGGAGGACCTCAGTGGCAAGAAACGACCAAAGCGATTCCTTGGTGGTCTGCT
CGCCGCGGCGTCAGCCATCGGGTCACTATTCTCCATTGGCCTCTCAGCTGCCAACTC
AATCAGCCTTAGTGCGCTCCAAAAACACATGGGTGAACTCGACAACGAAATGCCAG
AAATCCAACAGAGACTCCTCCTTCAACAAGAACAATTACAGGGCCTGGGTAAAACC
ATCCAGGGTACCATAGTCACTGTAAACTTACACTCCACTCTCCTCAATAACACTTTG
CATGCCTTAGAAACCTTGTCAGAAGTTGTCAAAGAAGATATAACCTATGTACGAGTA
GTCAGAGATCTAATGCAGGACCTCATCAGAGAAGTTAGTTCCTCAGTCAACAGCCTG
AGTGGTGGGAAGATCCCAGCCTATCTGGTCCCCCTGGACCTTGTTGAGAAAATCCTC
AAATCTGCGACTACCACTGTGGTTCAATCCTCACAAGTACACCTGGCATACAGTCTT
GGTAGTGCAATTCCAATTTCAGTAGACCCAAAGAATCTAGAGATAGGATTTATTCTT
AATCTACCCATCGTAGAAAGGCAAAACGTATACCGACTGAAGTCTGTACTAAATGTT
GGGTTTTGGAAAGGCAACACACACGTGTACCTTAAGACACCACCAACCATTGCCTA
CCACGATGATGATCCCTCGCTCTACCTCATTCCTAACCTAAACATGTGCACTAAAAC
GAAGGACATCCATTGGGTTTGCCCAAGCAATCCCTTTGTTAGAGATGTAACCAACTA
CCTTTGTGGCTTGAGAGCTGACTCCCCTGAACAAAGGTGCCAGGGTAGCATGTCCGC
CAAAGACGAAGGAATAGAGACGAGAGTAGAACGAGCTGGTAATCGATGGCTCGTC
AGCACACCTGCCACCGAAGCCCTAATGTCCTACGACCGCCATGATACAACCACTAA
GCTAACCTTACCGAACCAGACGATGTTCTTGACGGTTCCTGAAGGAGCCACTATACA
CATCGATGACATTGTCCTCCATTATCTCAACCCCGACAGGCATGACGTGGAGATTGA
GATTATGGACGCATTCAGAGGTCACAACCTCACAATTGATGACGCCCTTCAACAACA
GCTTCTAGCTGAAGGGACTAAATTGATTCAGTTCAGCCTCAAGCCAACTGGACTTAT
CACTACACCCTTTAGTCACGTAAGCAGGTCAACACCCTACCAAGAACACCCCATCAG
TGTAGTCGCCCTTGGGCTGCTTCTTAGTGGTTGGATCATCACGGCAATTGTCGCACA
CATCATGTACAAATACATTCAAAAGCTACAGGCCAAGCTGGATTCCCTCCTGTTTGT
GCCACCACGCTTTACACACCAGTCTGCGACTATCCCACGGGCTAACTCCACCCCTTT
CCAGGAGTAAGCCTACTAATCTCTAACCACTCTTCTTTTTTTTTTTTGTGTGTTTTGTA

TCTTACGAGTTTTTAGGACATCACCAATGATCAAATGCAGCAAAGACTGTATTAGGG
TTAATGTCATGTTTACCCTTGATTTGAAATTTGTATATTGATGTATGATGTTCAATGT
CGTTAGATAGTTAGAAGTCATCTGCCCAGATGCTGCCTCAATGTCTGCCCAGACACC
ATAGTCTCTCCAACACCTATGAACTTTATAACTGCTTGGACTGTTTAAACCCCTTTCC
TGCCCCCAGGAATACATGGACGGTGTAACCCTCAGGTTACCCTTGCTACATGGACTG
TTCGGCCCTAGTTTGCCCTAAAGACTGTGGCCAGCACCCCACCCTTCAGGTCAAAGG
GGGGATGTTGGGACCTCACCACGTGGGTGAACAAAGGAGAGGCCTACATGTAAAGA
CAAAAGCCTGAAACAATGAAACGGCGCCAAATCCCATCAGGCCTTGCACTTCACAC
CTTGGTGTGAAGTTTGGTTTAAATTCAAACACTGGTGTTTCATTGGATGAAACTCAC
GGGAACCCCAGGGAGTCCCCGTGACGCGGCCATGACGTCACCAGGAATATAATAAC
CGGCCCGCACCTCTGCACGGGGCCGTCCGCTGTGACCCCGAGTCCGACAGGAGGTG
ACCCAAACGTTTGTATTTCAACTAGCCGAGGCAACAGTCCCTAACTCGCTGGCCGAG
AGCAGACTGTTTTATTTTGTTTTTGTGTTTTTGTTTTACATCCGGACCCGGAAACAGA
TCGCCTGCAACGAGCCGATCTTGTTTCATTACCTGCAGAAGGAAGAAACATTCTTCA
AGAAAGAGAAGGATAACTTCTACGCCCCGCTCTCTTTCAATTAAGTCGACTCTCGCT
GTTTCCTCCGCCTGCCATACGAGGACCAGCTGATTCGCCGTTCAACCCTCCACAAAG
TGAGGTAACGGCTCCGTCATCGTCTGAGCCTGCCGCTAAGCGCCGCCGGACGTACG
GACTAAAGCACAC

> Loki-Ofa consensus sequence *gag* gene

ATGAGTAGCCAAGACCCCCGTAGATCAGGAGTTGAAGCTCCCCAGGAGTCAGACGA
TCTATCCCTGCAGGATGCGGTCGCTGGCCTCCTCCACCTCTCTGTAGAGGACCGAAG
TAACCTCGACAGCTACACTGTTGACGCATGTAATGAGAAATTGAAGGAGCTGGTGG
CTTATGTCAATCACCCGGCCGGGAAGCCACAGCGCACCATCTCTGACGTGCTTGGTG
AAATACTTCTCCTCTCTCAGCGCAAAGCCCGACTGCAAGCTGCAGAGCATAAGGTGC
AGCTTGCCCAGCTGCAGCGGAGCAATGCCAGCCTGCGACAGGAAGTCCAGGCCGCA
CAAACTGAGATCTCCCAGATACAGGATGATCATGATCAGCTGCAGGGGTATGCAGA
AACCCTGCATGCGCAGCTCCAGACTGCCCAACAAAAAACAGAGGTCAGGCGTGAAG
AACAGATCGGCCAGGTAAGTACCGTTGCCGACCCCTTCGCCCAAGAGTCACTTGAA
GATGCCCCTGACCTTGAAGCGGCTCCTCCTCAGGTCGCAGTCAGCCGACCCGTGCTC
ATAAGAGCACAGCCGACCTGGACGAGTCTGACCGACTGCGCTCAGAACTTCATGCC
GCCCGCANTGAGAGAAACAGCCCTGCTCCAGAGACTTGAGGTTGCCCGACCGAGTG
ACCAAACCACTCCGGCCAACCCAGGAGCCATGACTGGGAACTTAGCGGCCCACAGC
AGCGCGTTCTCTGAGCTGGTCACAGTTGACAGCTATCAAGCACCTCTATCGGCCCGC
TCATTCGACATGGGTGCTCCACCTCCCTTTAGCCGCCCTACCTTGCGTACTGACAGG
AGGGGATTCCCTCCCCTCCGGGGTGGGGATTACTCCCAAGTCCCGAGCGGCATAAGT
GGTTTTAAGCCCTCATCCACTAGCGAGCAATATACATCGCGCAGCTCGCGCCTGCGC
CTCAATAGGACTCCATCCCCACCGCCAAAGGGGGGAAGTCATCCTCACCGTTATATT
GCATTGAGTGACCATGGGACGTCAGATGACTCAGACGGCCCATGGCCGCACCGAGA
ACAAGGCCTACGTACTCGCCAGGTCGAGTCCCTCGCCAGGGATGTAGAGCGCTTCG
ACCCGGGCAATCGAGAGTCAAACATCAATGACTACCTCAGAGAAGTAGAGCGTTGT
CTCCTTGACCTACCTCACGCTTTGTCACGTGAGAAACTGAAGCTTATCTGGAAGACC
ACGGCTAGGAGCGTCCATGTCTTCATGGAAACGCTCCCACCTGGGACCAGAGACCG
CTACTCAGCACTCTGCAAGGCCCTACGGGAGGAGTACTCATTGTACACCGACCGAG
CTTCCGCCACTCTCGGCGCCTTCGCAATCATGCAAAAGAAGCATGAACCTCCAAGGG
AGTATTATAGGCGTCTGAGAGCTGCGTACTTCCTGGGACGTAATGCCCCAGGCCTGG
AGGAGGATCAGGCTTTCAAGTCCCTGTTCCTCCACAACCTCCACGAGAGTGTGCGCT

ATGATATAACTATGCACTGCAGGACGCATGATCTCACCATGCCGGAGATCAGGAAA
TATGCGCAGATGGCATGGGAGACACGTGCACGCCCTGTTAGGAGGTACGAAGACGA
CGCCAGGGTCCTTGGGATCCAAGCCTCGGCAAACACAGACTTAGCGCTAGAAGGTG
ACGAAATCCCTCGCGTTAAGACAAACACCAGAGCAGTACCTCAGGGTCGACGGCCA
TTTCGCCAACAGAACTGGCAACAGAACCAAGGGGGTGAGAGACAGCATTACTCCCA
AAACCAGGTCAGGCCTCGACAACCAAACACCCCTCAACCAGGAAGGCAAGGTTGGT
TCGAGAGAAACCTTAAGCAAGGAGGTAAGTATGATGGTAAGATCCCAAACAAACAG
AAAGGAGATGGTTCTAAGCAGGACATGGAGAGCCTGATTCGGAAATGCTTGGCGGA
GGCAATGAAACAGCTCGAGACATCCCGTCATTCACCAATGACGCCGGACCCTTCAA
AGGGACCTGATACCAAAACCCCGCCCTTATGA
> Loki-Ofa consensus sequence Gag protein
MSSQDPRRSGVEAPQESDDLSLQDAVAGLLHLSVEDRSNLDSYTVDACNEKLKELVAY
VNHPAGKPQRTISDVLGEILLLSQRKARLQAAEHKVQLAQLQRSNASLRQEVQAAQTEI
SQIQDDHDQLQGYAETLHAQLQTAQQKTEVRREEQIGQVSTVADPFAQESLEDAPDLEA
APPQVAVSRPVLIRAQPTWTSLTDCAQNFMPPAXRETALLQRLEVARPSDQTTPANPGA
MTGNLAAHSSAFSELVTVDSYQAPLSARSFDMGAPPPFSRPTLRTDRRGFPPLRGGDYS
QVPSGISGFKPSSTSEQYTSRSSRLRLNRTPSPPPKGGSHPHRYIALSDHGTSDDSDGPWP
HREQGLRTRQVESLARDVERFDPGNRESNINDYLREVERCLLDLPHALSREKLKLIWKT
TARSVHVFMETLPPGTRDRYSALCKALREEYSLYTDRASATLGAFAIMQKKHEPPREYY
RRLRAAYFLGRNAPGLEEDQAFKSLFLHNLHESVRYDITMHCRTHDLTMPEIRKYAQM
AWETRARPVRRYEDDARVLGIQASANTDLALEGDEIPRVKTNTRAVPQGRRPFRQQNW
QQNQGGERQHYSQNQVRPRQPNTPQPGRQGWFERNLKQGGKYDGKIPNKQKGDGSKQ
DMESLIRKCLAEAMKQLETSRHSPMTPDPSKGPDTKTPPL*
> Loki-Ofa consensus sequence *pol* gene
TTGACCTTCTTAGGTGATCTCATCCGCAAGGGCAACGCCCAACGGATATACACCTCA
GTAGTAGTAGGGGGGTGTGTCAGCACCCTCGCCCTGCTAGACTCTGGTTCTGAAATA
AGCTTGATGTGTTCCAGTACCTTCGCTGAGGTCGCCAGAGCGATGCTCTCCCTCGGC
AAGCCGTTACAAGTCGAGACCTGTAACGTCAACATCACCAGCTACACACAAGATAA
ATCACCCATCACGAAACGTGCTTGGTTTGATATTTCCTTCCAGGAGATGACACTAGC
GCATCCCATCTACATCTGTATCCTCGATACAGAACCTTTACTTATCGGTCAGGACCTC
TTGGACCGACTGGCCCCACTCATTGACTGCCACCGTGGACATATGTGGGCTCAGGTT
GACACCCCAAAGCCTCTTGGCTCCATGGGCTGTCCGCCAGCTCTGAACGGACAAGTC
GCCGTTATCCAAGAGTCCAGAGGACCCGTACCGATGCTCATGCCTCTCCCAGAGCCT
GATCCTGACTCTGAGCCTTCTATGCCACTGGCAAGCTCACAACCGCTGCAGATTGGA
AATGCAAACTCCTTCCACGCTCACGAGTCTTTCCTTTGCTCGCTGAGAAACACAAAC
TCAGCACTATACAGCCCCCATGTGGTTGGAGGCATACACCTCAATGGCATTACGGCT
CCAGAGGTAATACTAGCGCTCTGGTCTGAGAAGTCTGCCGTCAGCCAAGACCTGTAC
AACAGCCTGAGTCAGCGCAACCCTCGTTTGGCTTTTGTGCAAAGAAGCCACCGCCTT
CTTTCTGCAGATTCACCACAGACTCTCCTCAAGGCAATTGGGGTATGCGTTCTCTCCA
TACAGATTGGAAAAAAACAAGTCACTCACTTTGTGAGTGTTGTTCCTCGACTCCACC
CCCCATTCTCTATAGGGGCGGATCTTCTTGTCAGGCTCGGCGCCCAGCTGGACACGG
TGAACCAAATCCTGTGGTCGCAGGCCAACACAAAGGACTACCCCTTAACCACCGAG
CCAGAACACATGCTATCAGGGCAGACCATCCCTCAGGCATGCCAGGTCGCCAGTGA
AGTCGACGTGACTGTTCCTGCACGAACAGCAGGGGTTCCCATTCGCCTTGCGATCCT
TAAGGGACAGAAGATGCCTGACCTGCAGGCATTCTTCCAACCTCTACCCCATTTCTG
GGAACTCAACCTGACAGTCTGCGGTACACCCCTGCTTGAGTTGAACAATCGTTCAAC

CTATTTACTGGTTCAGAACCTGACGCACGTCCCAATTCAGGTGAGGGCACGGCAACC
ACTAGGGATGCTGGTCGACAGCTCTTTCCACGACTTTGAACTGACAGTGCCTGTCAT
AGGAGATCTGCCCCCATCCTTAGCCAGGGATGACAACCTGGACAATGCCCTTTTCAC
TTTTCCGACCAAAATGACCACCATTGCCCGTCACGAAATGCTACAGAACGAAGCGA
TCTGCGGCACCGCCCTGGGCACTGAAGGTGACATGGTCGTGTACGCTCTCGTGACTC
AGCCAGGAGACACGACCTCAGGTAGCAACGGAACCCGTTCCTCACCAGAGGAACCG
TACCCCGGCTTTGAGTCCGAAATTGAGCAGCAGCTTGCCAAAGCAAATGCTCTGACA
ACAGAGTACCAGAGAGATGCACTTCGGAAGTTATTCCGTGATTTCCAGTCTATCTTC
TCGAGGGACTCTAACGACTGTGGAGTCACTGACCTTCACACTGTTCGCATTCCCACT
GACCCAGAAACACCACCGTCATTTGTACGCCAGTACAGAATCCCGTTGGCTGCCTAT
GAGTCGATCCAAGAGATACTCGATAACTTGCTGCAGAAGCAAATTATCAGGGAATG
TAACTCAACCTACAACTCGCCAATCTGGCCGGTGTTAAAACCAACTGGTAAGTGGCG
CCTAACAATTGACTACCGCCCTTTGAACAAACAAGTTCCACTTTCACGCTGGCCCAT
GATCCACCTGGATCAAGAACTGGCCAAGGTAAAAGGGGCCTGCTTCTTCTCAACGGT
GGATGTGGCCAATGGTTTCTGGACCATGAGAGTCAATCCGGCTGACCAATACAAGCT
GGCTTTCTCCTTTGGCAACCGCCAGTTCACCTGGAACCGTTGCCCATTTGGATACTCG
AACTCCCCAGCTGAGTTCAACATTTTCCTCCACAAAGCGATGAGTGATGCCGCTGCC
CGCGGGAATCTCATTTACGTGGATGACATCCTCATGAGGAGTCGAACCTTGNAAGA
ACACCTGGCAGAGATCCGTCATGTCCTCAACCAACTTTCTGCCGCCGGAGCCAAACT
GGCCATCGCCAAGGGCCAATGGTGCCGTACCAAGGTGGAATATGTGGGTTTGACAG
TGGGTGCTGATGGCATCGAACCCCAGGCTGGGAGGATTCGAGCCATCCAAGACATC
AAAGCCCCGACGGGCGTGTCGGAACTCAGGAGCTTCCTGGGTGTCTGCAACTACTCC
CGACAGTTCATTGAGGACTACGCGGAGATAGCCAGACCCCTCACGGAGCTCCTTCG
CAAAGACAAGTCCTTTGAGTGGGGAGAACCTCAAGAGCAGTCCTTCCGCAAAATGA
AGGAGAAACTCTGCTCTGCATCTTGTTTGGCTTACCCAGACAAGGACAAAGAGTTCC
ACCTTGAGGCAAGTTTCTCTCCCCACAGCCTCAGTGCTGCCTTGGCGCAGAAATACG
ACACAGACAAGCGGGTCGTCGCCTACGCGAGCCGACCTCTAAGCAGCGTCGAGATT
AAATTCTCTGATTGCGAGAAAGCCCTCCTTGCTACCGTCTGGGCTGTAGAACACTTC
CGCAGCTACATCGGAGGCCAGAAGGTAATCATCGAGACTTGTCACCAACCTGTCAC
TTTCTTGAACAGCCAGCGGCTGAGAGAAGGACGTGTATCTAATAGCCGCATCGCCTC
CTGGATGATGGCACTGCAAGGCTACGACATCACGGTCAAGTATGCTCAGAACCACA
AGATGGCCTTGGGCCAGGGCCTGGCCGAATGTCAGCATTGCGACTGTGAGGAGCAA
CCCAACTCCCAGCCCCTTCTCTTAACCACTCCATCGCTTCCTTCCAATCATCATTATT
ATGATGAGAACGTTTGTCAAGACCTTCCAAGGGTCTATGTGGATGGCTGCTCGTTCC
ACCATGAAAGCCAGACACGGGCTGGAGTTGGCATTGTTTGGGTCAACCACAAGGTT
GACGAACCAAACCACTACCAGTTGGGCCCTAAGACAAGTCAGTATGCAGAGATCGC
TGCAGTGCTGATCACCCTCCAACAAGCTGCCAAGCTAGCCATCGTGCAGCTGGTGAT
CTGCTCCGACTCTAACTACGCTCGTCACAGCTTTATTTCCCACTTTCCCGTGTGGAAA
GAGAACGGCATGAAAAACGCACGGAATAAAGAAGTCAAACACTCAGAACTCTTCTT
GGCTTGCGACCAACTTGTGACGGACCGAGGGATGACCGTGTACTGGAAGAAGGTCA
AAGGTCATTCCCAAACCGCAGGTCCCGACAAGGATGGTAACGATGAAGCCGACCGC
CTAGCCAAGCTAGGTGCGGAACAGGGCACGCCCTGGGAGTTCCAGAGTGATTGGCT
CCCAATTCCCCAAACGTGTGCAGTGAATGCCATTACCCGCCGGCAGGCAAGAGAAA
AACGAGAGGACCCTCAAGCCTGCACCCAAACTCTACACCTGGGTCGCGCACCCAGC
GATGCAGATTTAGTTACTATGCAGGAACAAGACCCTGCCATTCAAGCCATCCGTCAG
TGGATGTTGAACCCTCCACGGCAAGGGGTCTCACAAACCTCACTGGTCGAGTCAGA

GGAGCTGCGAACTCTCCGCTCCGACCAGCTTCACCTCAAACTGGAAAAGGGTCTGCT
CGTCTACACCTACAACGGCCAGGGACCGACCCGTTGGGTCGTCCCAACCAACCATA
GGGGGGTGATGCTAGCCCACGCTCACGACTCTCCCGTCGGAGGCCACCGGGGTTAC
AGAGCCACTCTCCAAACCCTCCAGCAAGTCGCCTATTGGCCATTGATGGCTCGTGAC
ACTAAATCCTACGTCCAAGGATGTCTGGTCTGCTGCCAGTTCCAGCCGTCCCGACCT
CTGGGTCGTGCCCCACTCCAGAAAAGGGGGGTCACATTCCCATGGTCTCACCTCCAG
GTTGACTGGGTCGGACCAGTCCCCAAGTCATCACGAGGAAATAAATATCTTCTCACT
GTCACGTGCGCCTTCACCAAATGGGTGGAATGCCTCCCGGCACCCAACGATACTGCT
ATGACGACAGCAACCCTCCTGCTGAACCATGTGTTTAGCCGCTGGGGGCTACCCCTC
TCTGTCGACTCAGATCGAGGCACCCATTTCACGTCTAGTGTCATGACCACCCTGTAT
GACATCCTGGGAGTGGAAGTAAAATTCCACATCTCGCATCATCCCCAGTCATCCGGC
CAGGTCGAACGCGCCAACCGCACCATTGTTAGCATGCTCAAGAAGTATGTCAGCAG
CAATGGCAAGGATTGGGATGTGAAACTTCCTTTGGTGCTGATGGCCATCAGGTCCAC
CCCGCACCGCTCCACAGGAGTCACACCCTTTGAAATGATGACTGGAAGAGAGATGA
CCCTTCCACTACATCTCTTGTATCATCCTGAAGATGTCAGTGTTGCAACCGCCTATAC
GGCGCATCAATATGTGACGGACTTGCGCGACCACCTTCGAGCCACGTTCGCGTGGGC
CCAGACAAACTTGGAGGCCAGCGTTGAAGGATCAAAAGCCTACTATGACAGGAAAG
CGTCCCACCGCGAGTACCAAGTAGGCGATAAAATCTTGTACTTCAGATTCGCCAAAC
CGGTAGGAATCTCCAAGAAGTTCCTACCGAGCTGGTCAGGCCCGTTCGAGATTGTGG
GAAAACTCTCACCGGTAGCTTATCGGATCAGAGTCTCGAAACCAAAGCAACCACCA
GCATACAAATGGGTCCATGCAAATCAAATTAAACCCTACCAGTCATCCTCCCTCTCA
AGGGGAGAGGAAGCCTCTACACCCGAAAGGGGGGAGCAGTCTTCACTCCAAGAGG
GAGAAGACTCAGTCCAGTAA
> Loki-Ofa consensus sequence Pol protein
LTFLGDLIRKGNAQRIYTSVVVGGCVSTLALLDSGSEISLMCSSTFAEVARAMLSLGKPL
QVETCNVNITSYTQDKSPITKRAWFDISFQEMTLAHPIYICILDTEPLLIGQDLLDRLAPLI
DCHRGHMWAQVDTPKPLGSMGCPPALNGQVAVIQESRGPVPMLMPLPEPDPDSEPSMP
LASSQPLQIGNANSFHAHESFLCSLRNTNSALYSPHVVGGIHLNGITAPEVILALWSEKSA
VSQDLYNSLSQRNPRLAFVQRSHRLLSADSPQTLLKAIGVCVLSIQIGKKQVTHFVSVVP
RLHPPFSIGADLLVRLGAQLDTVNQILWSQANTKDYPLTTEPEHMLSGQTIPQACQVASE
VDVTVPARTAGVPIRLAILKGQKMPDLQAFFQPLPHFWELNLTVCGTPLLELNNRSTYL
LVQNLTHVPIQVRARQPLGMLVDSSFHDFELTVPVIGDLPPSLARDDNLDNALFTFPTK
MTTIARHEMLQNEAICGTALGTEGDMVVYALVTQPGDTTSGSNGTRSSPEEPYPGFESEI
EQQLAKANALTTEYQRDALRKLFRDFQSIFSRDSNDCGVTDLHTVRIPTDPETPPSFVRQ
YRIPLAAYESIQEILDNLLQKQIIRECNSTYNSPIWPVLKPTGKWRLTIDYRPLNKQVPLSR
WPMIHLDQELAKVKGACFFSTVDVANGFWTMRVNPADQYKLAFSFGNRQFTWNRCPF
GYSNSPAEFNIFLHKAMSDAAARGNLIYVDDILMRSRTLXEHLAEIRHVLNQLSAAGAK
LAIAKGQWCRTKVEYVGLTVGADGIEPQAGRIRAIQDIKAPTGVSELRSFLGVCNYSRQF
IEDYAEIARPLTELLRKDKSFEWGEPQEQSFRKMKEKLCSASCLAYPDKDKEFHLEASFS
PHSLSAALAQKYDTDKRVVAYASRPLSSVEIKFSDCEKALLATVWAVEHFRSYIGGQKV
IIETCHQPVTFLNSQRLREGRVSNSRIASWMMALQGYDITVKYAQNHKMALGQGLAEC
QHCDCEEQPNSQPLLLTTPSLPSNHHYYDENVCQDLPRVYVDGCSFHHESQTRAGVGIV
WVNHKVDEPNHYQLGPKTSQYAEIAAVLITLQQAAKLAIVQLVICSDSNYARHSFISHFP
VWKENGMKNARNKEVKHSELFLACDQLVTDRGMTVYWKKVKGHSQTAGPDKDGND
EADRLAKLGAEQGTPWEFQSDWLPIPQTCAVNAITRRQAREKREDPQACTQTLHLGRAP
SDADLVTMQEQDPAIQAIRQWMLNPPRQGVSQTSLVESEELRTLRSDQLHLKLEKGLLV

YTYNGQGPTRWVVPTNHRGVMLAHAHDSPVGGHRGYRATLQTLQQVAYWPLMARDT
KSYVQGCLVCCQFQPSRPLGRAPLQKRGVTFPWSHLQVDWVGPVPKSSRGNKYLLTVT
CAFTKWVECLPAPNDTAMTTATLLLNHVFSRWGLPLSVDSDRGTHFTSSVMTTLYDILG
VEVKFHISHHPQSSGQVERANRTIVSMLKKYVSSNGKDWDVKLPLVLMAIRSTPHRSTG
VTPFEMMTGREMTLPLHLLYHPEDVSVATAYTAHQYVTDLRDHLRATFAWAQTNLEA
SVEGSKAYYDRKASHREYQVGDKILYFRFAKPVGISKKFLPSWSGPFEIVGKLSPVAYRI
RVSKPKQPPAYKWVHANQIKPYQSSSLSRGEEASTPERGEQSSLQEGEDSVQ*
> Loki-Ofa consensus sequence *env* gene
ATGCAAATCAAATTAAACCCTACCAGTCATCCTCCCTCTCAAGGGGAGAGGAAGCC
TCTACACCCGAAAGGGGGGAGCAGTCTTCACTCCAAGAGGGAGAAGACTCAGTCCA
GTAAACCCATAGGGATAGCTGCATGCCTCAACTCACCCCGTCGGCATGCACTAACTC
CTAATACGTTCATCCCCCTTTTTCCCTACAATAATTCAACTCTGGTTTGTTTCAGGAT
GAAACTCCTTTGGATCTTCGGCATTCTCATCAGTCTACGAGTCACCCTGTCTATGGAA
GTGGTCGAACCCGGCCCACCGACCGGCATCGTCCTACAAGAGTCACCCGGACTCTTA
ATCACCCGTTGTGGACTCTACACCCAGAGAGTCTATGTACGCCTCGATCCTTGGAAT
GTGTACCGCAAACACGTCCGCTTACCCCCACGACTAACGGAGGGAAGGTTAAGTGG
AATCCAAACTCATGACACAGTTGAACGTGCTAAACAGACAACTGTCCATGTCCTCGA
GCAATTACAGAAGTTCCTGGTCACTGAGGAGGACCTCAGTGGCAAGAAACGACCAA
AGCGATTCCTTGGTGGTCTGCTCGCCGCGGCGTCAGCCATCGGGTCACTATTCTCCA
TTGGCCTCTCAGCTGCCAACTCAATCAGCCTTAGTGCGCTCCAAAAACACATGGGTG
AACTCGACAACGAAATGCCAGAAATCCAACAGAGACTCCTCCTTCAACAAGAACAA
TTACAGGGCCTGGGTAAAACCATCCAGGGTACCATAGTCACTGTAAACTTACACTCC
ACTCTCCTCAATAACACTTTGCATGCCTTAGAAACCTTGTCAGAAGTTGTCAAAGAA
GATATAACCTATGTACGAGTAGTCAGAGATCTAATGCAGGACCTCATCAGAGAAGT
TAGTTCCTCAGTCAACAGCCTGAGTGGTGGGAAGATCCCAGCCTATCTGGTCCCCCT
GGACCTTGTTGAGAAAATCCTCAAATCTGCGACTACCACTGTGGTTCAATCCTCACA
AGTACACCTGGCATACAGTCTTGGTAGTGCAATTCCAATTTCAGTAGACCCAAAGAA
TCTAGAGATAGGATTTATTCTTAATCTACCCATCGTAGAAAGGCAAAACGTATACCG
ACTGAAGTCTGTACTAAATGTTGGGTTTTGGAAAGGCAACACACACGTGTACCTTAA
GACACCACCAACCATTGCCTACCACGATGATGATCCCTCGCTCTACCTCATTCCTAA
CCTAAACATGTGCACTAAAACGAAGGACATCCATTGGGTTTGCCCAAGCAATCCCTT
TGTTAGAGATGTAACCAACTACCTTTGTGGCTTGAGAGCTGACTCCCCTGAACAAAG
GTGCCAGGGTAGCATGTCCGCCAAAGACGAAGGAATAGAGACGAGAGTAGAACGA
GCTGGTAATCGATGGCTCGTCAGCACACCTGCCACCGAAGCCCTAATGTCCTACGAC
CGCCATGATACAACCACTAAGCTAACCTTACCGAACCAGACGATGTTCTTGACGGTT
CCTGAAGGAGCCACTATACACATCGATGACATTGTCCTCCATTATCTCAACCCCGAC
AGGCATGACGTGGAGATTGAGATTATGGACGCATTCAGAGGTCACAACCTCACAAT
TGATGACGCCCTTCAACAACAGCTTCTAGCTGAAGGGACTAAATTGATTCAGTTCAG
CCTCAAGCCAACTGGACTTATCACTACACCCTTTAGTCACGTAAGCAGGTCAACACC
CTACCAAGAACACCCCATCAGTGTAGTCGCCCTTGGGCTGCTTCTTAGTGGTTGGAT
CATCACGGCAATTGTCGCACACATCATGTACAAATACATTCAAAAGCTACAGGCCA
AGCTGGATTCCCTCCTGTTTGTGCCACCACGCTTTACACACCAGTCTGCGACTATCCC
ACGGGCTAACTCCACCCCTTTCCAGGAGTAA
> Loki-Ofa consensus sequence Env protein
MQIKLNPTSHPPSQGERKPLHPKGGSSLHSKREKTQSSKPIGIAACLNSPRRHALTPNTFIP
LFPYNNSTLVCFRMKLLWIFGILISLRVTLSMEVVEPGPPTGIVLQESPGLLITRCGLYTQR

VYVRLDPWNVYRKHVRLPPRLTEGRLSGIQTHDTVERAKQTTVHVLEQLQKFLVTEED
LSGKKRPKRFLGGLLAAASAIGSLFSIGLSAANSISLSALQKHMGELDNEMPEIQQRLLL
QQEQLQGLGKTIQGTIVTVNLHSTLLNNTLHALETLSEVVKEDITYVRVVRDLMQDLIRE
VSSSVNSLSGGKIPAYLVPLDLVEKILKSATTTVVQSSQVHLAYSLGSAIPISVDPKNLEIG
FILNLPIVERQNVYRLKSVLNVGFWKGNTHVYLKTPPTIAYHDDDPSLYLIPNLNMCTKT
KDIHWVCPSNPFVRDVTNYLCGLRADSPEQRCQGSMSAKDEGIETRVERAGNRWLVST
PATEALMSYDRHDTTTKLTLPNQTMFLTVPEGATIHIDDIVLHYLNPDRHDVEIEIMDAF
RGHNLTIDDALQQQLLAEGTKLIQFSLKPTGLITTPFSHVSRSTPYQEHPISVVALGLLLSG
WIITAIVAHIMYKYIQKLQAKLDSLLFVPPRFTHQSATIPRANSTPFQE*
>Loki-Ame consensus sequence
TTGATCATTTAAGAATGTTAAGGTTTAGAAATTATACTTGCAGTTAGGAAACCTATC
ACAAGTTAGGGCCAATACGTCACGAAGGGAACGTGTTTTCCAATAAAGGGGGATGT
TAGTTCATTCAGTAGTAGTTCGTGGTTCGCAGTTCGGAGTTCGTGCGCTTTTTAGGTT
TAGTTAAGTTCAATTTTTTTTTTAAGTAATTTTAACTTAGAGTTTTAGTTAGAATTTTG
TACTCTCTCTTTTGGTTTTACTTTCTCTTAGATTCAGCTTAGTGATTCTTAAGTTCGTG
AGTCATAGATCCTAGCTCTAGCTGCGCTTGCAGCGCACCCGGGCTGACTAACACAAG
TCCAGCGATTATCAACAGAATTAGAGAATTAGGTGCTTCCAGCTATTGAACCACCTT
AGCGCTACGGAGCTGCCTATAGTTCACCGCTCATACTTAGCGTTATACATCGCTCGC
CGTACCAGACCCAGACGCGAAGAGAATCCTGAATCAATCAGCTGCCCGGAGCGAGG
CTCCCGCTTGCCGTTGGAACCGAAGCCCTGCTTGCACCGTGAAGTCCACCTGAAGAA
CCCGCTTGGCCACCTGGATCGAGCTACCCTTCCTGGCGGTCGACACCGACGGTTCCC
TTCGCGGCAGAGACGACACCGGTTGCCGAGGCCGGTCAACGCAAAGTGAGGCGCTG
GTTAAATCTCTAATTCTCTGGTGGCGTTCTGTTGGTCTGGTTAGGTTGAGTCTTGCTA
AGTTTTCAGTTATAGGTATTCCGGTTGAGATGGTGGTGATCTCTGTCAGTTCCATAGT
TCTAATGTAATCTAGTTTAACTCAAGTAGCTTGGGTTTAGGGCGTTTTTGCTTCTTTA
GTAGTTAGTTCGTTTATAGCAGTATATAATCACTTATAATCTTGTTTAATCTCCTAAT
ATTCTTATAGTGGGTTTAGTTATCCATATTTAGGTTTATCTCCTAGTGGGTTATCATA
ATTACTTCCCCTTTGGTATATTAAACTGGGTATTGATTTGTTTATTTTTATCTACTTAT
ACATTTCTGTTGTTTGGGTCTAAAGTGTTTTGAGTTATTCATATATGCGATGAGAAGG
TTTGGTTGTAAGACTGTTATCATCAGCGTGCTTGAATTTCACTCTCAAATTGTTGTTT
GAATTATGCCCCTTATCTAATTCTGCATATATCCATGTATTACTTTTATTCTAAATAA
ACATTACTTGATTTAGCCAATTCTTGACTTAAGTGTGTTCCTCCCGTGGTTCAAAGAC
TTTCTGAAACCGACAAGAACTCACCCTGAGGTTAAATATAAATTTTGTAGCTAGCCG
CTACTGGCGCCCTATTAACCTTTAATATTAATATTAAAATTAATATTGATATCAGGTA
AATAGTTCGGTATCCATTGTAACCACTACATATTTGGCGCCCAACGTGGGGCCAACG
TCACATCTCTGATCTACGGGAGAAAGCACGGAGTTAAGAAACGTATTGGTGTGTTGT
GTTTATTAAATATCTACCTTACGCTCCGCGTTAGTAAAACGCCAGCGGGCTGGCGAA
AGCTAACCGGGTCACTCCTCTGTCTGTGTGCGCGCACATCTATATACATCTATATATA
TATCCAACTCCTTTATGTGTATGTGTTTTAAAGAGCATTGGTAATCACTTGTGAAACG
CCTCATATTACTCGCCAAAGAGCTGATCAGTGTGAGTGCGTGTGTGTATTGTAGACG
CCGCTGTACGTCTGCTTTCATTATCACTGTTATTGGTGCTTTACATTTACCCGGGGTA
TCTAAACAGGCTCAGCCGACAACTATTATCACGTGAATTATATCAGCGCTATTTGTA
AAGGCCTCAGCGAGGTTGCTGTAGCTGCGGATTTCCAGTGCAGTAAACCCGCTTCCC
AGTTTAAACGGCCCGTGTGCGCGGATTTCCGCACCCCGGTCGTACTGGGACCCTCTC
GTGTGCCCCGCGGCTTTGCAAACCGCGGCGTAACGCTATCCTTTTAGGTGGTAGAGA
GGTTTACAGTCGCGAAAACGCGACGGGCGAGGATCACGCCCTCAGCGGAAGCCCCG

CTTGACTCAGACTGTGTGAAGTAAATTCCTGCGTGCTGTGTTCCTCTCTGTGTCTCTG
TGTTAAGGTCAGCAAGACCGTTGTGTTCCCCACAAAGCTTGAAAGCAACTGAGCTCT
CAGGAGCTGAGGTTTGTCCTCCCACTCAAGGGGGCGGGACTCAGTTGTGTCACCGGT
AGGCGGAGCCTTCCCCACCAGTGGCCTCTAGTCGCCATCTCGTGGTCAACTGTATCA
GCTACATCTTGCCTCAAGGCTGTCTTACCTTTTCGTGCCTCTAGGGGTAAAGCTATAT
CTACACTGCCCTCTTCTGGTCGTTTGGGTGAATACTCAGGTTAAAGGGGGTTGCAGT
TCATTTATACGCTATTTTGGGAAATTCTTGTAGTAAACTGTGTTTACGTGTGGTTGGT
GTTGGCTTCATATGAGTTTTGGCATACCTTAATCGAGATACCCTCTTCCCCTGCTTTC
TTTGCTTGGTTACTGTTATATTTTAAATGTTGTATCGTCCTAGTTATTATTAATACTAG
TAAGTGCCATTTCCATTCTTAATTACTAGTAAGTATATATTCAAATTAGAATTCACTT
CAAATAGATACATTTAGATAACTCAATTGTAATAATTCAATATTTTGTTCTATGTAGC
CATCCTACTTCAAACATAGTGTCAACTCTCAAAGCTAACTTGCTTGTCAGCTGTTCTA
CTCACATTGCTTACATTAAATATATTCATTTGAGCTAACCATAATATCCAGTATTTTT
AAACATTCAAGTTATAATCATTGGTTGCTATAAATACCTGTTATTTAGGTAATCCATA
GTAGCCATTCATTATTTTGTTTTGTCAAACATATTATCCACTTCTTCTTGTTATTCTAA
ATACAGGCATATAATCCTTATCCTATTTGGCCATAAAAAAAAAAATAAAAAAAAAA
AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAATA
AATAAATAAAAAAAAAAAAATAAAAAAAAAAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAAAAAAAAAAAAACAAAAAAAAAAAAAAATAAATAA
ATAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAATAA
ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAAAAAAAATA
AATAAAAAAAAAAAATAAATAAATAAAAAAAAAAAATAAATAAATAAATAAATAAATAA
ATAAAAAAAAAAATAAAAAAAAAAAAAATAAATAAATAAATAAATAAATAAAAAAA
AAATAAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAA
TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAA
TAAATAAATAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAA
AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAA
AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAA
AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA
AATAAATAAATAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAA
AAAATAAAAAAAAAAAAAAATAAAATAAAAAAAAAAATAAAAAAAAAAAATAAAAA
AAAAAATAAATAAATAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAATAAATAAATAAATAAA
TAAATAAATAAATAAATAAATAAATAAAAAAAAAAATAAATAAATAAATAAATAAAT
AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAAAAA
AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA
AATAAATAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAAAAAAAAAAAAA
ATAAAAAAAAAAAATAAATAAATAAATAAATAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAACTACACTTTTGGCTTATTAGATTCACTAGGGACTACGCATCAAAGTTG
TTACCCTAGCATCAATCTCTCCAATCTCAGCAGCAGTGCTAAGTCACCCATCCGGGA
CTGGCTTCCAATACAGCTCGCATCGCTTGACCCTCGTCAAGTCAAAATGAGTTCTCC
AAGAGAGGCAGGGTCCCCGCTCCTCAGCGAGGCTGAGTCGGTGGACCAAGTAGAAG
CCACTCTTCTTGATTTAGTAAGCTGCCGCATTCCAGACTATGCTGAAGGCGTGCAAC
AGGCAGATGATGCCCACCTTCAGAGTTTGATGACTAGCTGTGTGGACGACCACCTAA
CTAAATCCCTAAAAATAAAAGAACAAATAGACACCGTGTACAAAATGGCCCTCATA
CTCTGGAGACAACTCCAACTCTCACAAGAGAGAGAGCAACATGCTACTAACCACTG
TACAGCTCTCCAGCAAACTAACCAAAAACTCAGGGACGAATTGGCGCAATGGCATA
ATACTGGTTCAGAGATACAATCCCAACAAGAAGCCCTTGAACATAAAGTAAAGTCA
CTGCAGGAAGAGGCCGCAACAGCCACAAAGCTTGCCATCCAGTCACGAGGTAAGCT
GGAGGAAGCCCAGGAGCAGCAGACCGAACTCGAGGCTGAACTAACACAGGCTATG
AGTGCTTTAAGATTAGCTCGTATAGAAAAAAGGGATGCAGAAACGCAGTTGGCAGA
GTTCAAATTAGAACGGGATGAGACATATTCCCGGGGGCCAAAGGTAAGCCCAACAA
AGGGTAGACCAGCGACCCCACATCTCTTTAAGACGGAGGTACATACTCCTACGCCC
GTAACTTCCATGGGTTACCAGGGAGGAGCTGAAGCTCCTCGATCACCCATCAGAGTC
AAACCGGAACCGGCTGGCCCCAAAAAGGTGAACCGGACCCTGTTTGGCTGGGACGA
ATCTGAACACACCCGACGACCCCAAACAAATGCAGATATGGCCAGGCCCATAACAC
CCTCAGATAGAGATTTGGATAAAGTTGCCCGCAATATCAGCCGGTTTGAACCTAAGC
CTGGCTGCCCTAATGATATCCACCTTTATCTAAATGACATTGATTACTATCTACGGCG
CTTCCCTCTGGCCACCATAGACGACAGAATCTATCTCATTAAAATAACATCCAGCCG
CGACGTTAGCAGTTTCATTGAACGCCAACCACCACACGTAAGGGGGAACTATAGGG
CTCTCTGCCAAGCATTGGAGGCAGAATTCTCTAGCTTCTCAACCCAGTTAGGCCTCT
CCGCTGCATTCACGGTGAGGCAGAACAGGCAGGAGGGTCCCGAACAGTACTACTAT
CGCCTCAGGCGCGCTTACTTTGGCTTCAAGAATGAACCAGAGATGGAGGAGGACGA
AAATTTCAAGGTCCTGTTTGTCCAGAACCTCCATCCTTCAACCAGCCACCATTTAGGT
GTCTCGGCTTGTCCCCGCACCCTGACTAGTCGGCAGCTAAGAGATCTAGCTCTTAGA
GGCTTTGCAAAAGTTCAGCAAAGTTTGGGTAAAAAAACCGACTCCGAGGTCCTTGC
AGTAGCCAGTAACCCTTGTATGGAACTAGAGGGGAAGCACAGCTTCCAGGGAACAG
ATAAGACTCACGGGTCCAGACCCTATAGACAAGAACCAGCCCAAAGTCAGCAGGTC
GATCCACAGCGACGCCGCTCACCTACTTATAAGGTGAATCGCGGCCAGGGCAGACC
CCATCGAGGTCGAGCCTCCCACCATCAACGTGGTGGCAGGCCACACCATCATGGGA
ATAACAGGCACTCTGGACCAAACCAGACTGCAAAAAGCCCTAAATCATCACTTGAC
AAAACCCTTAACGGCCTGGAGGCAGGGGATCTTGAGTTAATCCTGCAATGGGTTAA
CCAAATGCGCAAGGAGAAAAGGGGTAACCTCAACCAAGGGCGCAACATCCAGGAG
CAGTCGAGCACTGCTCCCCATGAAAAAGACCCCAAACGAGCTACTTGGGAGCATGC
CCAAGACACTGCACCTAGTGCAGTCCTAGTTGTGGAACTAGACACTGACCACTCTCT
CAAGGAAGAAGACCCATTTCTAATTCGGGGGGAGCTTCCTTTCCGACAGTTCTTGGG
CCATCTGAGCGAAAGGGGGGTGGCCAGAAAATTCTATCTTGCTACTACTTTGGAGGA
TGAATTTGTGCAAGATGCTCTTCTTGACACGGCATCAGACATCACACTGATGTCCAT
GTCCCTTTTTAACCGAGTGCGGGCATCCGCTCGGCGACGGAATAGGGAGATACAGCT
GAGTAACTGTTCTCTCAATATCCAGCCCTACTCACACACGGGCATCACCATACATTC
GATGGCCCTAATTCAACTTACGATTGGGCCCGTGACTGTTGTTCATCCAATTCATGTG
TCTCCCTTGGAGTCCATCCCATTCCTCGCAGGAAAAGATCTTCTTAATCGGTTCGAGC
CCCTTATCGACTTTCAAGGTCTTAAGATCTGGGCGCAGGTTCGGCAACCCCTTCCTGT
TCTTCCCTCTCGTGAGGAGGAGATCCAGTGCTACCGTCTTTGTGGACAGTTCCAACC

ATGCAAGGAACTGATCGACCCCGTGGTTCCTATGGAGATTGAGGAAGAGCCTACAG
CCGCCGTTTCACAGTTTACAAATCCTTCTTGGCCTCCAAGCTCAGTCTTGGAACAAC
GTGACACCTTCTTGTGTACCATGGTCCCATCTGAAAATCTTGGGGAATATTGTCCCA
GAGTGGAGAAGGGGGTCACATTGGACGACTTCCAGGTGGATGACGTGGTCATCGCC
CTGTGGTCGGACAAATCCGCAGTCAGCCAAGACTTAGTAAACTCATTATCGGCAACC
CATGCTAGCTTTCATTTCATTCAATCCCAGGGAGCTTTTCCGCTTGGGCCCGTTCCTC
AACCCACCTTGCCGGCAGTCGGCACTTGTTCTTTGACCGTCCGTTGGAACAAGAAAG
AGTTCACCCAGCTTTTTCTTGTGGTCCCAGACTTACCTCATTCCATCTATGTGGGCGG
TGACGTGTTGGTCAGGTTGGATGTTCAGGTAGACACCATCAATAACATCCTTTGGTC
CCTCACAAACACTCAGGCAGATGGCTCTTCCCCTGAGCCAGAGTATCTGCACTCGGG
TCAGACCATACCTGAAGTTTGTCAGGTGGCAAACAACCAAGCCATCACCCTACCGG
CAAATACCAAAAATATCGGCCTCTACCTGACCATCCTACCTGGACAACAACTATGTC
ACCCCCATGCACTGTTCCAACCCTCACATCGATTCTTTGATCTTGGTCTGACCGTGGA
AGCCACACCTCTTGTGGAGTTACGTGCTAGGTCAGTCTACTTGTTCGTTCAAAATTGT
ACAACCCAGGACATAGTGATTCCCCAACTCACAGAACTGGGTTGGCTGGTTAGCAC
CAAATTTCATGACTTTGAACTGCGGCTTCCAGTCATCGGTCCGATTCCTCAGCAGTTG
CTCCCAAAGCAAGAGTCCAAGACATTATACACCCGACCGTCTGGGGCAATTGCTATT
TTTCCTGTCTTGGAGACCAATGAAGATGCGTTGTACCGCATTGATCTGACCCACAGG
GACCAAATGGTCTTGGATACAGTCACTGTCCTTAGTCTCGATTCTGTCCCATCACCTG
GCACTAACACAGTCCTGGTGAAGCAGAAGGAGGAGGCCGACAGTTTTCCTTCAACT
CAAACAGAGTTTGAGAAGCAACTACAGCAAATGCTGTTAGATGCGGATGCATTGCA
ATCAGACGAAGAGCGTGTAAAGCTCAGAACAGTGCTGACTAAGTATCGCGCGTCTT
TCTCTCAAGATTCTATGGATTGTGGATTGACCCACATCCATATGATTCGTATTCCGAC
GCATCCCGACGCTGCTCCTGCTTACGTTCGCCAATACAAGATTCCACTTGCATCATA
CGGGCCGGTGCAAGAGATCATAGATGATCTCTTAGACAAAGGAATAATCCGACCAT
GCAACAGCACCTACTCTGCGCCACTTTGGCCGGTGTTGAAGCCCAATGGCAAATGGC
GTCTGACTATTGACTATCGCCGCCTAAATGACCAGGTCCCTCTTTCTAGATGGCCCAT
GACCCAATTGGAGCAAGAACTCCCTAGGGTCAGAGATGCCAAATATTTTTCAACCTT
AGATGTGGCATCGGGGTTCTGGACCATTCCGGTCCATGTGGAAGACCAACACAAAC
TGGCATTCACATTCGCCGGCCGTCAATTTACATTCACCCGCTGCCCATTCGGATATTC
CAATTCACCAGCTGAATTCAACATCTTTCTCAACAAAGCTTGTCCTGATGCAAGAGA
ACGAGGAACCTTGATTTATGTTGACGATGTTCTTATCCGGAACAATTCTCTCGATGCT
CACTTGGAAGAGATTGATCATGTGCTAGATCAGCTCACAAAAGCAGGTGCGAAAAT
ATCCCTAGCCAAATGTCAGTGGTGCAAAACTAAGGTCAACTATGTGGGCCTTTTGGT
TGGTCCAGATGGTGTTCTCCCGCAACCATGTCGTGCCCAAGGCATCGTGGATATCGC
GGAACCAAAAACCATCCACGCCCTTCGCAGCTTCTTGGGGGTATGCAACTACTCCCG
CCAGTTCATTGAAAACTTTGCGGAACTGGCCAAGCCACTGTACCAGCTGTTAAAGCA
GGATGTACCTTTCATTTGGGGTGAGGCCCAGGCACAAGCCATGCAAACCCTCAAAG
ACAAACTGGCATANGACAAACTGGCATTCACATTCGCCGGCCGTCAATTTACATTCA
CCCGCTGCCCATTCGGATATTCCAATTCACCAGCTGAATTCAACATCTTTCTCAACA
AAGCTTGTCCTGATGCAAGAGAACGAGGAACCTTGATTTATGTTGACGATGTTCTTA
TCCGGAACAATTCTCTCGATGCTCACTTGGAAGAGATTGATCATGTGCTAGATCAGC
TCACAAAAGCAGGTGCGAAAATATCCCTAGCCAAATGTCAGTGGTGCAAAACTAAG
GTCAACTATGTGGGCCTTTTGGTTGGTCCAGATGGTGTTCTCCCGCAACCATGTCGTG
CCCAAGGCATCGTGGATATCGCGGAACCAAAAACCATCCACGCCCTTCGCAGCTTCT
TGGGGGTATGCAACTACTCCCGCCAGTTCATTGAAAACTTTGCGGAACTGGCCAAGC

CACTGTACCAGCTGTTAAAGCAGGATGTACCTTTCATTTGGGGTGAGGCCCAGGCAC
AAGCCATGCAAACCCTCAAAGACAAACTGGCATCGGCACCCTGTTTAACCTATCCG
GACCACAGCAGAGAATTCTACCTTCAAGTTGGGTTTTCAGAGCACTGTGTGAGTGCT
GGTCTATACCAGGTACACGATCGAGACAAGAGAGTGGTGGCTTATGCAAGCAAAGC
TTTGATGGCTCCGGAACTCAAATACTCTGACTGTGAGAAAGCACTTCTGGCCACGGT
CTGGGCAGTGAAACACTTCTCCAACTACCTTGGGGGACAAAAGATCATTGTAGAAA
CGAACCACCAGCCAGTGGTCTTTCTCAACAGTCAGCGCATTCGAGAAGGTGTGGTCA
CCAACGCACGAGTCGCCTCCTGGTTAATGGCCCTACAAAGTTTTGAGGTTGAGGTGC
GCTATGCGCAAAACTCGAGATTACCTCTAGGCACTGATCTGGCGGCATGTCAGCGAT
GTGAAACTGACATACCGGCCACATCAACGCCCATTCCCAATCTCGCCTCCCAAAAAC
CCACAAACCATCGTTACTTTGACCCCAAGGAGTGTGAAAACATTCCCACGGTCTACG
TAGATGGATGCTCTTTCCGACATGACCAAGAAGGCCTCAAAGCCGGTGCAGGAATC
GTCTGGCTTGATGACAATCCTTGCGAACCACAACAGTTCAAACTAGGATCACAAACT
TCTCAGTATGCCGAAATCGCTGCCATACTCATCACAATTCAACTTGCTATCGACCAG
GGGGTAAAAACCCTGGTAATCTGTACAGATTCCAACTACGCACGTCTAAGTTTCACA
TGTCACTTGCCAATCTGGAAAACCAAAGGTTTTCTCACATCTGGCAGGAAAGCAGTC
AAACACACAGAGCTGTTCACAGCCGCGGACTACCTGGTAGTTCGGCATGACATGCTT
GTCTACTGGAAAAAGGTGAGGGGACACTCACGTGTCCCAGGGACAGACAAGACATA
CAATGATCAAGCTGATTCACTGGCCAAAAGAGGTGCCTTGGAAGGAGTTTCATGGG
TCTTTGACCCGTTGAAATACCCAACCCAACCCAACCCCACTGTTTTAGCCGTGACTA
GAGCACAAGCTAAACAAACCTCAACAACAGAAATCCCACCCTGCGCTGCGGTCTCC
ATTGATCCAGAAATCACAGATGCTGATTTGATCACGCTACAAGACGCAGACCCCGA
CATCAAATCTATTAAGGCTTTCTTGCTTGATCCCACCAACAATCCCATAACCAGCCA
AATGCTGGAGGCATCAATTCCGCTTAAACAACTCTTGGACAACCGCGCATTCTTGAA
GGTGGTCAAGGGCTTGCTTGTTCATGTCACTGAGACCCACACTTCTCCTGCGTTCGTG
GTGCCTCCGTGCCACAGGGGGGTAATGTTGGGGCACGCCCATGACTCTCCATCAGCA
GGTCACAAAGGGATCAAGGAAACCTATAGGACACTCAAACAGGTGGCCTTTTGGCC
ACGAATGCGGGAGCATGTAGCTAGTTACATTAAAGGATGCTTGGTGTGCTGCCAGTT
CCAGCCGGCGAATCCTCTGCACAGGGCTCCCTTACAGAGAAAGGGGTTAACCTTCCC
GTGGTCGGATCTGCAAATGGATTGGGTGGGCCCGCTCACAAAGTCTGCTCGAGGGA
ACAAATATTTCCTCACAGTCACATGTGCATTCACCAAGTGGGTGGAGTGTCTTCCAG
CCGCCAATGACACCGCTATAACCACAGCGTGTCTCCTGGTTAATCATATCTTCACCC
GGTTTGGTCTTCCTTGTAGGGTTGACTCAGATCGAGGTACTCATTTCACGGCAGAAG
TCATGCAACAAATGTGGCACGTCCTCGGTGTAAAAGCCAAGTTGCACGTGAGCTACC
ACCCCCAGTCCTCTGGTCAGGTCGAACGGGCTAACCGCACAGTGGTCAGCATCCTCA
AGAAATTTGTGGCCACGCATCACCGCGACTGGGATGTGAAATTGCCGTTGGTACTCA
TGGCCATTAGGGCAACTCCTCACGATTCCACCGGCGTCTCTCCATTCGAGGTGATGA
CAGGCAGACAGATGACACTTCCGCTACATCTGTTATATCAGCCAGGTGGAACAGAT
ACCTCAGTTGCCTACACAACACAGCAGTACATAGACGATTTGCAGAAACACCTAAA
ATCCACTTTTGCTTTTGCCCAACAGCACATGGCACGAAGTGCCCAGGGTCGGAAGGC
CTACTATGATCAGAAGGCGTCCTATGACGAGCTTCAGGTGGGAGATAAGGTGTGGT
ACTACATTTTCGCTCAACCTGTCGAGAAACCCCAATCGAATACGGGTAGGCTGGCTC
GCAAACTCATGCCTCGCTGGACTGGACCCCACACAATCACGGACAAATTGTCCCCG
GTGGTGTACCAGATCCAGATCAGCCAAGGTTCTAAGGAGCCAACCCGCAAGTGGGT
ACACCGCAACCAAATTAAGCCACACCAAAGCCCTATGGGACAAGTCGGGGCCGCCA
ATGCCTTGATTAGAGTCTAATCAGGTGGACAAAACCCCTGCTAGACCCCTAGGTGTT

TAGCAACTCCTAGCTGTACAACTGGCAACTAAACTAATAACTTCATTTTTAAAATTT
AGGAATACATTGATCTTTGATCCATATACAGTAGTCTGTTGTTTGCAATTTCAACCCA
AGTACCATTACACATCTGCCGATCAAGCCAGCCAACAGAACCATCATTTTGTGCCTG
GGAATAGGATATCACTGATTGTTAGCCACTTATCTGTTAGATTCTACTTCATTGGTGG
TTCACACTAGTACACCCAGGTTAATCATGTACATTTGTCTCGTTAGGAAGATGAGGA
CTCTCCTGCTCCTGACGCTCGTGCTGGGGACGATCGTCGCGGCCCCTGATGTGATCT
CCCCCGGACCTACCTCAGGAGTAGTTCTTAGGGATAATTCGGGACTACTAATCACAA
ACTGCCGGTTACATACCCAAAGGGTGTTTGTACGACTGAACCCGGAGGAAGTATGT
AAAAAGAACTTCCCTGATTGGTCCCGCCAAGTGGGGTGGGCTGGGCATTGGTGGGT
ACGAGAGGCAGTAAGCCATGCCGAACGAGACACCACCCACATGCTACAACAGCTGC
AAAAATTTGCGGTCACACAGTCTGAACTGAATGGGGCTCAAAAGAGATCCAAACGG
TTTCTTGCAGGTTTATTAACTGCTGCCGCCGCTGTAGGCTCACTCCTTGGCGTCGGCA
CATCCGCTGTTAATGCCGTGGGCCTAGCCACGGTCCGCCGTAATGTAGGGGAACTAC
AGGCAGAACTACCACACATCAAGGACCAAATAGACACCCAACAGAGACAGCTACA
GGCAATTGGACAAACTTTACAGGGTACCATTGTGGTCCTCAATACCCACACGGTAGC
TTTAAACAAAACTCTGCAAACGATCACCTCCCTGACCCGGGCAATCCAGACAGATAT
GGCTCACATCCAAGTGATTAACATGCTTTTGAATGACATGCTACAGGAAGTGGGATC
CTCAGTAGACAGTCTAGCTATGGGCAAGATACCAGCTTATCTGGTGCCTTTAACCCT
GGTACAGGAAATCCTCAGTCGGGCTACCCGAGAGGAAGTTACTGATCTCCAAGCAC
ATTTAGCCTATACACTCGGAAGTGCAGTACCTCTTTATGTCAATCCAGAAAATAGAG
AAGTAGCTTTTCTGGTCAACTTACCTATAATAGCTGCGGAAAACATTTATCGCTTAA
AAAGAGTGGTAAATGTTGGGTTTTGGCAAAAAGAAGCACACATCAAAATTCAGACT
CCGACCTTGGTAGCTTACCAAGACATCAATCCTGGACTTTATCTGGCCCCAAATTTG
CGCATGTGTACCCTGTCCAAGGATATACACTATCTCTGTCCCAGTAAACCGTTCACT
CGAGAGGGTGCCGGCAATCTTTGTGGCCTTAAGCCCATGCTCACTAAGGACCAGTGT
CCTGCCACTGTCACTCATAGGTACCAACAGGTCGAGACTCAGGTTGAAATCGTAGGG
AGCCGGTGGTTGGTAAATACTGCGGCACGGGAAGCGGTGTTGTCCTATGACCAGCA
TGATATCGATGAACGTATTCCCCTTCCGGAACAAACATTTTGGGTAGACGTACCACC
TGGGGCTACTCTGCATGTTGGGGAATGGTCCCTGTACCACTTAGTACCTGATCAATA
TGAGAGTGAAATTGAGGTGTCTGATTTCTTTACTACCCACTCCTTAGAACTCAGCCCT
GACACCCTTGGTCGGATACAGTATGAGGGTCCCCAGTCTATCGATTTGACCCCCATC
AATAATGTCTTAAAAGAGATCGAGGCACGTCCCACATGGACAGTCCAACCTGTAAG
ATACTCCTGGTCATTGCCCGATACCTTGTTAGCTGTGCTGGTATGTCTGGGGTATGTA
GCAACCTTTGGTATAACGTGTTTTTATCGTAGGAAGACCGTGAAGCTACAGAGACAA
ATGGATTCATGGTCGGGTAAGGTGGTTAGATTTGTGAGACGTAGAAGGGGGGGTCA
AGAAGTGTCAGTGGAAACTCAGGAGGACGCAACTGATGAACCTATGCAGGTAGCTG
CAGTAAGGGAAGTGGTTTAGCTACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC

ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACACGCAGTTTGACCCTTTA
ACTGGTCAAGGACCCGTGGACCATTTAACAAGACAATACTGCCACCCGCAGGAGAA
GAAAGGAACTTTTCTTGGAGCACGAGGCTATGATCGAAAAATTACATATGTATATGT
TTTTACAATAACTTAAATAACTGTTGGTTAATTCTAATATTGTGAGAAGCAGGGTAT
CAAGCAATGTGATGAAACACTTTCTAATCAAGAGAACAATTTTGAGGTTTACCTTCA
TTTTCTTCTACAGGTTCCAGGCTGAAGAATTATGTCTTGTTTGTCATTGGTTGAACTA
TTGCAACTAGTCAGCGTAGTTTTAACTTTTGTKTTGGTTATATAGATACATTTTCCKG
TTTAGTAAAGAGCAAAGCATAAAAGTATAAAGGTATAGTGAATTTCCGATTTAAAG
GGTTCTGAAATGATGTTAACCAGAGCAATAGGACCTGCATTCCTTAGTGGGGGAGA
GGTGTGAATTACGACGGCCACAGATTTACATCTTTGAATACACAGGCGAGTCTGCCA
TTTCAAAAGGGAAATTTAAGGCTGAAAAATGTTAAAACATTAAAAACAAAATGTGT
TCTTAAGACGTGAATGAACATTAACTACAAAAAGAACACGGAGGGCAACAAACTAG
CAGTCTGGTTAAAATAATTACTTAAAAGCTATCAGGGCATGTTTCCACAATACTTTA
ATCAGGTTAATAACCAAATCTTGAATTAATAAGTTTAAAGTTCATAACGAGTTTCTT
TCTTTCCTTTCTAAGAGTCCGGTGCGTGTGAGTTAGTGTATTCCCTGGTCATGTGTTA
TATTTCTTTCCATTCTTACACCTCAAAATCAACTATGATTATCTAATTGTCTGGTTAG
GATTTCAGTTCTGAGATGTTTATTCATATCACTAAAATGCTATGCAACAGTGGGTGA
ATGTGGGGAAATTGTGGGTTTTCTATGTCAAAGTACATTTTAATAAGCAATGGTTTA
ATTTGATCATTTAAGAATGTTAAGGTTTAAGTTAATAAAGAAAAAAAGGGGGAAGT
CGTCTAGGCAAGACTCCACCTACCTGGAAGATGATTGGGTTTTTAGAAATTATACTT
GCAGTTAGGAAACCTATCACAAGTTAGGGCCAATACGTCACGAAGGGAACGTGTTT
TCCAATAAAGGGGGATGTTAGTTCATTCAGTAGTAGTTCGTGGTTCGCAGTTCGGAG
TTCGTGCGCTTTTTAGGTTTAGTTAAGTTCAATTTTTTTTTAGAAGTAATTTTAACTTA
GAGTTTTAGTTAGAATTTTGTACTCTCTCTTTTGTTTTACTTTCTCTTAGATTCAGCTT
AGTGATTCTTAAGTTCGTGAGTCATAGATCCTAGCTCTAGCTGCGCTTGCAGCGCAC
CCGCGCTGACTAACACAAGTCCAGCGATTATCAACAGAATTAGAGAATTAGGTGCTT
CCAGCTATTGAACCACCTTAGCGCTACGGAGCTGCCTATAGTTCACCGCTCATACTT
AGCGTTATACATCGCTCGCCCGACCAGACCCAGACGCGAAGAGAACCCTTGATCAA
TCAGCTGTCCGGAGCGAGGCTCCCGCTTGCCGTTGGAACCGAAGCCCTGCTTGCACC
GTGAAGTCCACCTGAAGAACCCGCTTCGGCCACTTGGATCGAGCTACCCTTCCTGGC
GGTCGACACCGACGGTTCCCTTCGCGGCAGAGACGACACCGGTTGCCCGAGGCCGG
TCAACGCAAAGTGAGGCGCTGGTTAAATCTCTAATTCTCTGGTGGCGTTCTGTTGGT
CTGGTTAGGTTGAGTCTTGCTAAGTTCTCAGTTATAGGTATTCCGGTTGAGATGGTG
GTGATCTCTGTCAGTTCCATAGTTCTAATGTAGTCTAGTTTAACTCAAGTAGCTTGGG
TTTAGGGCGTTTTTGCTTCTTTAGTAGTTAGTTCGTTTATAGCAGTATATAATCACTT
ATAATCTTGTTTAATCTCCTAATATTCTTATAGTGGGTTTAGTTATCCATATTTAGGTT

TATCTCCTAGTGGGTTATCATAATTACTTCCCCTTTGGTATATTAAACTGGGTATTGA
TTTGTTTATTTTTATCTACTTATACATTTCTGTTGTTTGGGTCTAAAGTGTTTTGAGTT
ATTCATATATGCGATGAGAAGGTTTGGTTGTAAGACTGTTATCATCAGCGTGCTTGA
ATTTCACTCTCAAATTGTTGTTTGAATTATGCCCCTTATCTAATTCTGCATATATCCAT
GTATTACTTTTATTCTAAATAAACATTACTTGATTTAGCCAATTCTTGATTTAAGTGT
GTTCCTCCCGTGGTTCAAAGACTTTCTGAAACCGACAAGAACTCCA
> Loki-Ame consensus sequence 5’LTR
TGTTAAGGTTTAGAAATTATACTTGCAGTTAGGAAACCTATCACAAGTTAGGGCCAA
TACGTCACGAAGGGAACGTGTTTTCCAATAAAGGGGGATGTTAGTTCATTCAGTAGT
AGTTCGTGGTTCGCAGTTCGGAGTTCGTGCGCTTTTTAGGTTTAGTTAAGTTCAATTT
TTTTTTTAAGTAATTTTAACTTAGAGTTTTAGTTAGAATTTTGTACTCTCTCTTTTGGT
TTTACTTTCTCTTAGATTCAGCTTAGTGATTCTTAAGTTCGTGAGTCATAGATCCTAG
CTCTAGCTGCGCTTGCAGCGCACCCGGGCTGACTAACACAAGTCCAGCGATTATCAA
CAGAATTAGAGAATTAGGTGCTTCCAGCTATTGAACCACCTTAGCGCTACGGAGCTG
CCTATAGTTCACCGCTCATACTTAGCGTTATACATCGCTCGCCGTACCAGACCCAGA
CGCGAAGAGAATCCTGAATCAATCAGCTGCCCGGAGCGAGGCTCCCGCTTGCCGTT
GGAACCGAAGCCCTGCTTGCACCGTGAAGTCCACCTGAAGAACCCGCTTGGCCACC
TGGATCGAGCTACCCTTCCTGGCGGTCGACACCGACGGTTCCCTTCGCGGCAGAGAC
GACACCGGTTGCCGAGGCCGGTCAACGCAAAGTGAGGCGCTGGTTAAATCTCTAAT
TCTCTGGTGGCGTTCTGTTGGTCTGGTTAGGTTGAGTCTTGCTAAGTTTTCAGTTATA
GGTATTCCGGTTGAGATGGTGGTGATCTCTGTCAGTTCCATAGTTCTAATGTAATCTA
GTTTAACTCAAGTAGCTTGGGTTTAGGGCGTTTTTGCTTCTTTAGTAGTTAGTTCGTT
TATAGCAGTATATAATCACTTATAATCTTGTTTAATCTCCTAATATTCTTATAGTGGG
TTTAGTTATCCATATTTAGGTTTATCTCCTAGTGGGTTATCATAATTACTTCCCCTTTG
GTATATTAAACTGGGTATTGATTTGTTTATTTTTATCTACTTATACATTTCTGTTGTTT
GGGTCTAAAGTGTTTTGAGTTATTCATATATGCGATGAGAAGGTTTGGTTGTAAGAC
TGTTATCATCAGCGTGCTTGAATTTCACTCTCAAATTGTTGTTTGAATTATGCCCCTT
ATCTAATTCTGCATATATCCATGTATTACTTTTATTCTAAATAAACATTACTTGATTT
AGCCAATTCTTGACTTAAGTGTGTTCCTCCCGTGGTTCAAAGACTTTCTGAAACCGA
CA
> Loki-Ame consensus sequence 3’LTR
TGATTGGGTTTTTAGAAATTATACTTGCAGTTAGGAAACCTATCACAAGTTAGGGCC
AATACGTCACGAAGGGAACGTGTTTTCCAATAAAGGGGGATGTTAGTTCATTCAGTA
GTAGTTCGTGGTTCGCAGTTCGGAGTTCGTGCGCTTTTTAGGTTTAGTTAAGTTCAAT
TTTTTTTTAGAAGTAATTTTAACTTAGAGTTTTAGTTAGAATTTTGTACTCTCTCTTTT
GTTTTACTTTCTCTTAGATTCAGCTTAGTGATTCTTAAGTTCGTGAGTCATAGATCCT
AGCTCTAGCTGCGCTTGCAGCGCACCCGCGCTGACTAACACAAGTCCAGCGATTATC
AACAGAATTAGAGAATTAGGTGCTTCCAGCTATTGAACCACCTTAGCGCTACGGAGC
TGCCTATAGTTCACCGCTCATACTTAGCGTTATACATCGCTCGCCCGACCAGACCCA
GACGCGAAGAGAACCCTTGATCAATCAGCTGTCCGGAGCGAGGCTCCCGCTTGCCG
TTGGAACCGAAGCCCTGCTTGCACCGTGAAGTCCACCTGAAGAACCCGCTTCGGCCA
CTTGGATCGAGCTACCCTTCCTGGCGGTCGACACCGACGGTTCCCTTCGCGGCAGAG
ACGACACCGGTTGCCCGAGGCCGGTCAACGCAAAGTGAGGCGCTGGTTAAATCTCT
AATTCTCTGGTGGCGTTCTGTTGGTCTGGTTAGGTTGAGTCTTGCTAAGTTCTCAGTT
ATAGGTATTCCGGTTGAGATGGTGGTGATCTCTGTCAGTTCCATAGTTCTAATGTAGT
CTAGTTTAACTCAAGTAGCTTGGGTTTAGGGCGTTTTTGCTTCTTTAGTAGTTAGTTC

GTTTATAGCAGTATATAATCACTTATAATCTTGTTTAATCTCCTAATATTCTTATAGT
GGGTTTAGTTATCCATATTTAGGTTTATCTCCTAGTGGGTTATCATAATTACTTCCCC
TTTGGTATATTAAACTGGGTATTGATTTGTTTATTTTTATCTACTTATACATTTCTGTT
GTTTGGGTCTAAAGTGTTTTGAGTTATTCATATATGCGATGAGAAGGTTTGGTTGTA
AGACTGTTATCATCAGCGTGCTTGAATTTCACTCTCAAATTGTTGTTTGAATTATGCC
CCTTATCTAATTCTGCATATATCCATGTATTACTTTTATTCTAAATAAACATTACTTG
ATTTAGCCAATTCTTGATTTAAGTGTGTTCCTCCCGTGGTTCAAAGACTTTCTGAAAC
CGACA
> Loki-Ame consensus sequence *smc-gag-pol* gene
ATGAGTTCTCCAAGAGAGGCAGGGTCCCCGCTCCTCAGCGAGGCTGAGTCGGTGGA
CCAAGTAGAAGCCACTCTTCTTGATTTAGTAAGCTGCCGCATTCCAGACTATGCTGA
AGGCGTGCAACAGGCAGATGATGCCCACCTTCAGAGTTTGATGACTAGCTGTGTGG
ACGACCACCTAACTAAATCCCTAAAAATAAAAGAACAAATAGACACCGTGTACAAA
ATGGCCCTCATACTCTGGAGACAACTCCAACTCTCACAAGAGAGAGAGCAACATGC
TACTAACCACTGTACAGCTCTCCAGCAAACTAACCAAAAACTCAGGGACGAATTGG
CGCAATGGCATAATACTGGTTCAGAGATACAATCCCAACAAGAAGCCCTTGAACAT
AAAGTAAAGTCACTGCAGGAAGAGGCCGCAACAGCCACAAAGCTTGCCATCCAGTC
ACGAGGTAAGCTGGAGGAAGCCCAGGAGCAGCAGACCGAACTCGAGGCTGAACTA
ACACAGGCTATGAGTGCTTTAAGATTAGCTCGTATAGAAAAAAGGGATGCAGAAAC
GCAGTTGGCAGAGTTCAAATTAGAACGGGATGAGACATATTCCCGGGGGCCAAAGG
TAAGCCCAACAAAGGGTAGACCAGCGACCCCACATCTCTTTAAGACGGAGGTACAT
ACTCCTACGCCCGTAACTTCCATGGGTTACCAGGGAGGAGCTGAAGCTCCTCGATCA
CCCATCAGAGTCAAACCGGAACCGGCTGGCCCCAAAAAGGTGAACCGGACCCTGTT
TGGCTGGGACGAATCTGAACACACCCGACGACCCCAAACAAATGCAGATATGGCCA
GGCCCATAACACCCTCAGATAGAGATTTGGATAAAGTTGCCCGCAATATCAGCCGGT
TTGAACCTAAGCCTGGCTGCCCTAATGATATCCACCTTTATCTAAATGACATTGATTA
CTATCTACGGCGCTTCCCTCTGGCCACCATAGACGACAGAATCTATCTCATTAAAAT
AACATCCAGCCGCGACGTTAGCAGTTTCATTGAACGCCAACCACCACACGTAAGGG
GGAACTATAGGGCTCTCTGCCAAGCATTGGAGGCAGAATTCTCTAGCTTCTCAACCC
AGTTAGGCCTCTCCGCTGCATTCACGGTGAGGCAGAACAGGCAGGAGGGTCCCGAA
CAGTACTACTATCGCCTCAGGCGCGCTTACTTTGGCTTCAAGAATGAACCAGAGATG
GAGGAGGACGAAAATTTCAAGGTCCTGTTTGTCCAGAACCTCCATCCTTCAACCAGC
CACCATTTAGGTGTCTCGGCTTGTCCCCGCACCCTGACTAGTCGGCAGCTAAGAGAT
CTAGCTCTTAGAGGCTTTGCAAAAGTTCAGCAAAGTTTGGGTAAAAAAACCGACTCC
GAGGTCCTTGCAGTAGCCAGTAACCCTTGTATGGAACTAGAGGGGAAGCACAGCTT
CCAGGGAACAGATAAGACTCACGGGTCCAGACCCTATAGACAAGAACCAGCCCAAA
GTCAGCAGGTCGATCCACAGCGACGCCGCTCACCTACTTATAAGGTGAATCGCGGC
CAGGGCAGACCCCATCGAGGTCGAGCCTCCCACCATCAACGTGGTGGCAGGCCACA
CCATCATGGGAATAACAGGCACTCTGGACCAAACCAGACTGCAAAAAGCCCTAAAT
CATCACTTGACAAAACCCTTAACGGCCTGGAGGCAGGGGATCTTGAGTTAATCCTGC
AATGGGTTAACCAAATGCGCAAGGAGAAAAGGGGTAACCTCAACCAAGGGCGCAA
CATCCAGGAGCAGTCGAGCACTGCTCCCCATGAAAAAGACCCCAAACGAGCTACTT
GGGAGCATGCCCAAGACACTGCACCTAGTGCAGTCCTAGTTGTGGAACTAGACACT
GACCACTCTCTCAAGGAAGAAGACCCATTTCTAATTCGGGGGGAGCTTCCTTTCCGA
CAGTTCTTGGGCCATCTGAGCGAAAGGGGGGTGGCCAGAAAATTCTATCTTGCTACT
ACTTTGGAGGATGAATTTGTGCAAGATGCTCTTCTTGACACGGCATCAGACATCACA

CTGATGTCCATGTCCCTTTTTAACCGAGTGCGGGCATCCGCTCGGCGACGGAATAGG
GAGATACAGCTGAGTAACTGTTCTCTCAATATCCAGCCCTACTCACACACGGGCATC
ACCATACATTCGATGGCCCTAATTCAACTTACGATTGGGCCCGTGACTGTTGTTCATC
CAATTCATGTGTCTCCCTTGGAGTCCATCCCATTCCTCGCAGGAAAAGATCTTCTTAA
TCGGTTCGAGCCCCTTATCGACTTTCAAGGTCTTAAGATCTGGGCGCAGGTTCGGCA
ACCCCTTCCTGTTCTTCCCTCTCGTGAGGAGGAGATCCAGTGCTACCGTCTTTGTGGA
CAGTTCCAACCATGCAAGGAACTGATCGACCCCGTGGTTCCTATGGAGATTGAGGA
AGAGCCTACAGCCGCCGTTTCACAGTTTACAAATCCTTCTTGGCCTCCAAGCTCAGT
CTTGGAACAACGTGACACCTTCTTGTGTACCATGGTCCCATCTGAAAATCTTGGGGA
ATATTGTCCCAGAGTGGAGAAGGGGGTCACATTGGACGACTTCCAGGTGGATGACG
TGGTCATCGCCCTGTGGTCGGACAAATCCGCAGTCAGCCAAGACTTAGTAAACTCAT
TATCGGCAACCCATGCTAGCTTTCATTTCATTCAATCCCAGGGAGCTTTTCCGCTTGG
GCCCGTTCCTCAACCCACCTTGCCGGCAGTCGGCACTTGTTCTTTGACCGTCCGTTGG
AACAAGAAAGAGTTCACCCAGCTTTTTCTTGTGGTCCCAGACTTACCTCATTCCATCT
ATGTGGGCGGTGACGTGTTGGTCAGGTTGGATGTTCAGGTAGACACCATCAATAACA
TCCTTTGGTCCCTCACAAACACTCAGGCAGATGGCTCTTCCCCTGAGCCAGAGTATC
TGCACTCGGGTCAGACCATACCTGAAGTTTGTCAGGTGGCAAACAACCAAGCCATC
ACCCTACCGGCAAATACCAAAAATATCGGCCTCTACCTGACCATCCTACCTGGACAA
CAACTATGTCACCCCCATGCACTGTTCCAACCCTCACATCGATTCTTTGATCTTGGTC
TGACCGTGGAAGCCACACCTCTTGTGGAGTTACGTGCTAGGTCAGTCTACTTGTTCG
TTCAAAATTGTACAACCCAGGACATAGTGATTCCCCAACTCACAGAACTGGGTTGGC
TGGTTAGCACCAAATTTCATGACTTTGAACTGCGGCTTCCAGTCATCGGTCCGATTCC
TCAGCAGTTGCTCCCAAAGCAAGAGTCCAAGACATTATACACCCGACCGTCTGGGG
CAATTGCTATTTTTCCTGTCTTGGAGACCAATGAAGATGCGTTGTACCGCATTGATCT
GACCCACAGGGACCAAATGGTCTTGGATACAGTCACTGTCCTTAGTCTCGATTCTGT
CCCATCACCTGGCACTAACACAGTCCTGGTGAAGCAGAAGGAGGAGGCCGACAGTT
TTCCTTCAACTCAAACAGAGTTTGAGAAGCAACTACAGCAAATGCTGTTAGATGCGG
ATGCATTGCAATCAGACGAAGAGCGTGTAAAGCTCAGAACAGTGCTGACTAAGTAT
CGCGCGTCTTTCTCTCAAGATTCTATGGATTGTGGATTGACCCACATCCATATGATTC
GTATTCCGACGCATCCCGACGCTGCTCCTGCTTACGTTCGCCAATACAAGATTCCAC
TTGCATCATACGGGCCGGTGCAAGAGATCATAGATGATCTCTTAGACAAAGGAATA
ATCCGACCATGCAACAGCACCTACTCTGCGCCACTTTGGCCGGTGTTGAAGCCCAAT
GGCAAATGGCGTCTGACTATTGACTATCGCCGCCTAAATGACCAGGTCCCTCTTTCT
AGATGGCCCATGACCCAATTGGAGCAAGAACTCCCTAGGGTCAGAGATGCCAAATA
TTTTTCAACCTTAGATGTGGCATCGGGGTTCTGGACCATTCCGGTCCATGTGGAAGA
CCAACACAAACTGGCATTCACATTCGCCGGCCGTCAATTTACATTCACCCGCTGCCC
ATTCGGATATTCCAATTCACCAGCTGAATTCAACATCTTTCTCAACAAAGCTTGTCCT
GATGCAAGAGAACGAGGAACCTTGATTTATGTTGACGATGTTCTTATCCGGAACAAT
TCTCTCGATGCTCACTTGGAAGAGATTGATCATGTGCTAGATCAGCTCACAAAAGCA
GGTGCGAAAATATCCCTAGCCAAATGTCAGTGGTGCAAAACTAAGGTCAACTATGT
GGGCCTTTTGGTTGGTCCAGATGGTGTTCTCCCGCAACCATGTCGTGCCCAAGGCAT
CGTGGATATCGCGGAACCAAAAACCATCCACGCCCTTCGCAGCTTCTTGGGGGTATG
CAACTACTCCCGCCAGTTCATTGAAAACTTTGCGGAACTGGCCAAGCCACTGTACCA
GCTGTTAAAGCAGGATGTACCTTTCATTTGGGGTGAGGCCCAGGCACAAGCCATGCA
AACCCTCAAAGACAAACTGGCATANGACAAACTGGCATTCACATTCGCCGGCCGTC
AATTTACATTCACCCGCTGCCCATTCGGATATTCCAATTCACCAGCTGAATTCAACAT

CTTTCTCAACAAAGCTTGTCCTGATGCAAGAGAACGAGGAACCTTGATTTATGTTGA
CGATGTTCTTATCCGGAACAATTCTCTCGATGCTCACTTGGAAGAGATTGATCATGT
GCTAGATCAGCTCACAAAAGCAGGTGCGAAAATATCCCTAGCCAAATGTCAGTGGT
GCAAAACTAAGGTCAACTATGTGGGCCTTTTGGTTGGTCCAGATGGTGTTCTCCCGC
AACCATGTCGTGCCCAAGGCATCGTGGATATCGCGGAACCAAAAACCATCCACGCC
CTTCGCAGCTTCTTGGGGGTATGCAACTACTCCCGCCAGTTCATTGAAAACTTTGCG
GAACTGGCCAAGCCACTGTACCAGCTGTTAAAGCAGGATGTACCTTTCATTTGGGGT
GAGGCCCAGGCACAAGCCATGCAAACCCTCAAAGACAAACTGGCATCGGCACCCTG
TTTAACCTATCCGGACCACAGCAGAGAATTCTACCTTCAAGTTGGGTTTTCAGAGCA
CTGTGTGAGTGCTGGTCTATACCAGGTACACGATCGAGACAAGAGAGTGGTGGCTT
ATGCAAGCAAAGCTTTGATGGCTCCGGAACTCAAATACTCTGACTGTGAGAAAGCA
CTTCTGGCCACGGTCTGGGCAGTGAAACACTTCTCCAACTACCTTGGGGGACAAAAG
ATCATTGTAGAAACGAACCACCAGCCAGTGGTCTTTCTCAACAGTCAGCGCATTCGA
GAAGGTGTGGTCACCAACGCACGAGTCGCCTCCTGGTTAATGGCCCTACAAAGTTTT
GAGGTTGAGGTGCGCTATGCGCAAAACTCGAGATTACCTCTAGGCACTGATCTGGCG
GCATGTCAGCGATGTGAAACTGACATACCGGCCACATCAACGCCCATTCCCAATCTC
GCCTCCCAAAAACCCACAAACCATCGTTACTTTGACCCCAAGGAGTGTGAAAACATT
CCCACGGTCTACGTAGATGGATGCTCTTTCCGACATGACCAAGAAGGCCTCAAAGCC
GGTGCAGGAATCGTCTGGCTTGATGACAATCCTTGCGAACCACAACAGTTCAAACTA
GGATCACAAACTTCTCAGTATGCCGAAATCGCTGCCATACTCATCACAATTCAACTT
GCTATCGACCAGGGGGTAAAAACCCTGGTAATCTGTACAGATTCCAACTACGCACGT
CTAAGTTTCACATGTCACTTGCCAATCTGGAAAACCAAAGGTTTTCTCACATCTGGC
AGGAAAGCAGTCAAACACACAGAGCTGTTCACAGCCGCGGACTACCTGGTAGTTCG
GCATGACATGCTTGTCTACTGGAAAAAGGTGAGGGGACACTCACGTGTCCCAGGGA
CAGACAAGACATACAATGATCAAGCTGATTCACTGGCCAAAAGAGGTGCCTTGGAA
GGAGTTTCATGGGTCTTTGACCCGTTGAAATACCCAACCCAACCCAACCCCACTGTT
TTAGCCGTGACTAGAGCACAAGCTAAACAAACCTCAACAACAGAAATCCCACCCTG
CGCTGCGGTCTCCATTGATCCAGAAATCACAGATGCTGATTTGATCACGCTACAAGA
CGCAGACCCCGACATCAAATCTATTAAGGCTTTCTTGCTTGATCCCACCAACAATCC
CATAACCAGCCAAATGCTGGAGGCATCAATTCCGCTTAAACAACTCTTGGACAACC
GCGCATTCTTGAAGGTGGTCAAGGGCTTGCTTGTTCATGTCACTGAGACCCACACTT
CTCCTGCGTTCGTGGTGCCTCCGTGCCACAGGGGGGTAATGTTGGGGCACGCCCATG
ACTCTCCATCAGCAGGTCACAAAGGGATCAAGGAAACCTATAGGACACTCAAACAG
GTGGCCTTTTGGCCACGAATGCGGGAGCATGTAGCTAGTTACATTAAAGGATGCTTG
GTGTGCTGCCAGTTCCAGCCGGCGAATCCTCTGCACAGGGCTCCCTTACAGAGAAAG
GGGTTAACCTTCCCGTGGTCGGATCTGCAAATGGATTGGGTGGGCCCGCTCACAAAG
TCTGCTCGAGGGAACAAATATTTCCTCACAGTCACATGTGCATTCACCAAGTGGGTG
GAGTGTCTTCCAGCCGCCAATGACACCGCTATAACCACAGCGTGTCTCCTGGTTAAT
CATATCTTCACCCGGTTTGGTCTTCCTTGTAGGGTTGACTCAGATCGAGGTACTCATT
TCACGGCAGAAGTCATGCAACAAATGTGGCACGTCCTCGGTGTAAAAGCCAAGTTG
CACGTGAGCTACCACCCCCAGTCCTCTGGTCAGGTCGAACGGGCTAACCGCACAGT
GGTCAGCATCCTCAAGAAATTTGTGGCCACGCATCACCGCGACTGGGATGTGAAATT
GCCGTTGGTACTCATGGCCATTAGGGCAACTCCTCACGATTCCACCGGCGTCTCTCC
ATTCGAGGTGATGACAGGCAGACAGATGACACTTCCGCTACATCTGTTATATCAGCC
AGGTGGAACAGATACCTCAGTTGCCTACACAACACAGCAGTACATAGACGATTTGC
AGAAACACCTAAAATCCACTTTTGCTTTTGCCCAACAGCACATGGCACGAAGTGCCC

AGGGTCGGAAGGCCTACTATGATCAGAAGGCGTCCTATGACGAGCTTCAGGTGGGA
GATAAGGTGTGGTACTACATTTTCGCTCAACCTGTCGAGAAACCCCAATCGAATACG
GGTAGGCTGGCTCGCAAACTCATGCCTCGCTGGACTGGACCCCACACAATCACGGA
CAAATTGTCCCCGGTGGTGTACCAGATCCAGATCAGCCAAGGTTCTAAGGAGCCAA
CCCGCAAGTGGGTACACCGCAACCAAATTAAGCCACACCAAAGCCCTATGGGACAA
GTCGGGGCCGCCAATGCCTTGATTAGAGTCTAA
> Loki-Ame consensus sequence Smc-Gag-Pol polyprotein
MSSPREAGSPLLSEAESVDQVEATLLDLVSCRIPDYAEGVQQADDAHLQSLMTSCVDDH
LTKSLKIKEQIDTVYKMALILWRQLQLSQEREQHATNHCTALQQTNQKLRDELAQWHN
TGSEIQSQQEALEHKVKSLQEEAATATKLAIQSRGKLEEAQEQQTELEAELTQAMSALR
LARIEKRDAETQLAEFKLERDETYSRGPKVSPTKGRPATPHLFKTEVHTPTPVTSMGYQ
GGAEAPRSPIRVKPEPAGPKKVNRTLFGWDESEHTRRPQTNADMARPITPSDRDLDKVA
RNISRFEPKPGCPNDIHLYLNDIDYYLRRFPLATIDDRIYLIKITSSRDVSSFIERQPPHVRG
NYRALCQALEAEFSSFSTQLGLSAAFTVRQNRQEGPEQYYYRLRRAYFGFKNEPEMEED
ENFKVLFVQNLHPSTSHHLGVSACPRTLTSRQLRDLALRGFAKVQQSLGKKTDSEVLAV
ASNPCMELEGKHSFQGTDKTHGSRPYRQEPAQSQQVDPQRRRSPTYKVNRGQGRPHRG
RASHHQRGGRPHHHGNNRHSGPNQTAKSPKSSLDKTLNGLEAGDLELILQWVNQMRK
EKRGNLNQGRNIQEQSSTAPHEKDPKRATWEHAQDTAPSAVLVVELDTDHSLKEEDPFL
IRGELPFRQFLGHLSERGVARKFYLATTLEDEFVQDALLDTASDITLMSMSLFNRVRASA
RRRNREIQLSNCSLNIQPYSHTGITIHSMALIQLTIGPVTVVHPIHVSPLESIPFLAGKDLLN
RFEPLIDFQGLKIWAQVRQPLPVLPSREEEIQCYRLCGQFQPCKELIDPVVPMEIEEEPTA
AVSQFTNPSWPPSSVLEQRDTFLCTMVPSENLGEYCPRVEKGVTLDDFQVDDVVIALWS
DKSAVSQDLVNSLSATHASFHFIQSQGAFPLGPVPQPTLPAVGTCSLTVRWNKKEFTQLF
LVVPDLPHSIYVGGDVLVRLDVQVDTINNILWSLTNTQADGSSPEPEYLHSGQTIPEVCQ
VANNQAITLPANTKNIGLYLTILPGQQLCHPHALFQPSHRFFDLGLTVEATPLVELRARS
VYLFVQNCTTQDIVIPQLTELGWLVSTKFHDFELRLPVIGPIPQQLLPKQESKTLYTRPSG
AIAIFPVLETNEDALYRIDLTHRDQMVLDTVTVLSLDSVPSPGTNTVLVKQKEEADSFPS
TQTEFEKQLQQMLLDADALQSDEERVKLRTVLTKYRASFSQDSMDCGLTHIHMIRIPTH
PDAAPAYVRQYKIPLASYGPVQEIIDDLLDKGIIRPCNSTYSAPLWPVLKPNGKWRLTID
YRRLNDQVPLSRWPMTQLEQELPRVRDAKYFSTLDVASGFWTIPVHVEDQHKLAFTFA
GRQFTFTRCPFGYSNSPAEFNIFLNKACPDARERGTLIYVDDVLIRNNSLDAHLEEIDHVL
DQLTKAGAKISLAKCQWCKTKVNYVGLLVGPDGVLPQPCRAQGIVDIAEPKTIHALRSF
LGVCNYSRQFIENFAELAKPLYQLLKQDVPFIWGEAQAQAMQTLKDKLAXDKLAFTFA
GRQFTFTRCPFGYSNSPAEFNIFLNKACPDARERGTLIYVDDVLIRNNSLDAHLEEIDHVL
DQLTKAGAKISLAKCQWCKTKVNYVGLLVGPDGVLPQPCRAQGIVDIAEPKTIHALRSF
LGVCNYSRQFIENFAELAKPLYQLLKQDVPFIWGEAQAQAMQTLKDKLASAPCLTYPD
HSREFYLQVGFSEHCVSAGLYQVHDRDKRVVAYASKALMAPELKYSDCEKALLATVW
AVKHFSNYLGGQKIIVETNHQPVVFLNSQRIREGVVTNARVASWLMALQSFEVEVRYA
QNSRLPLGTDLAACQRCETDIPATSTPIPNLASQKPTNHRYFDPKECENIPTVYVDGCSFR
HDQEGLKAGAGIVWLDDNPCEPQQFKLGSQTSQYAEIAAILITIQLAIDQGVKTLVICTDS
NYARLSFTCHLPIWKTKGFLTSGRKAVKHTELFTAADYLVVRHDMLVYWKKVRGHSR
VPGTDKTYNDQADSLAKRGALEGVSWVFDPLKYPTQPNPTVLAVTRAQAKQTSTTEIPP
CAAVSIDPEITDADLITLQDADPDIKSIKAFLLDPTNNPITSQMLEASIPLKQLLDNRAFLK
VVKGLLVHVTETHTSPAFVVPPCHRGVMLGHAHDSPSAGHKGIKETYRTLKQVAFWPR
MREHVASYIKGCLVCCQFQPANPLHRAPLQRKGLTFPWSDLQMDWVGPLTKSARGNK
YFLTVTCAFTKWVECLPAANDTAITTACLLVNHIFTRFGLPCRVDSDRGTHFTAEVMQQ

MWHVLGVKAKLHVSYHPQSSGQVERANRTVVSILKKFVATHHRDWDVKLPLVLMAIR
ATPHDSTGVSPFEVMTGRQMTLPLHLLYQPGGTDTSVAYTTQQYIDDLQKHLKSTFAFA
QQHMARSAQGRKAYYDQKASYDELQVGDKVWYYIFAQPVEKPQSNTGRLARKLMPR
WTGPHTITDKLSPVVYQIQISQGSKEPTRKWVHRNQIKPHQSPMGQVGAANALIRV*
> Loki-Ame consensus sequence *env* gene
TTGTTAGCCACTTATCTGTTAGATTCTACTTCATTGGTGGTTCACACTAGTACACCCA
GGTTAATCATGTACATTTGTCTCGTTAGGAAGATGAGGACTCTCCTGCTCCTGACGC
TCGTGCTGGGGACGATCGTCGCGGCCCCTGATGTGATCTCCCCCGGACCTACCTCAG
GAGTAGTTCTTAGGGATAATTCGGGACTACTAATCACAAACTGCCGGTTACATACCC
AAAGGGTGTTTGTACGACTGAACCCGGAGGAAGTATGTAAAAAGAACTTCCCTGAT
TGGTCCCGCCAAGTGGGGTGGGCTGGGCATTGGTGGGTACGAGAGGCAGTAAGCCA
TGCCGAACGAGACACCACCCACATGCTACAACAGCTGCAAAAATTTGCGGTCACAC
AGTCTGAACTGAATGGGGCTCAAAAGAGATCCAAACGGTTTCTTGCAGGTTTATTAA
CTGCTGCCGCCGCTGTAGGCTCACTCCTTGGCGTCGGCACATCCGCTGTTAATGCCG
TGGGCCTAGCCACGGTCCGCCGTAATGTAGGGGAACTACAGGCAGAACTACCACAC
ATCAAGGACCAAATAGACACCCAACAGAGACAGCTACAGGCAATTGGACAAACTTT
ACAGGGTACCATTGTGGTCCTCAATACCCACACGGTAGCTTTAAACAAAACTCTGCA
AACGATCACCTCCCTGACCCGGGCAATCCAGACAGATATGGCTCACATCCAAGTGA
TTAACATGCTTTTGAATGACATGCTACAGGAAGTGGGATCCTCAGTAGACAGTCTAG
CTATGGGCAAGATACCAGCTTATCTGGTGCCTTTAACCCTGGTACAGGAAATCCTCA
GTCGGGCTACCCGAGAGGAAGTTACTGATCTCCAAGCACATTTAGCCTATACACTCG
GAAGTGCAGTACCTCTTTATGTCAATCCAGAAAATAGAGAAGTAGCTTTTCTGGTCA
ACTTACCTATAATAGCTGCGGAAAACATTTATCGCTTAAAAAGAGTGGTAAATGTTG
GGTTTTGGCAAAAAGAAGCACACATCAAAATTCAGACTCCGACCTTGGTAGCTTACC
AAGACATCAATCCTGGACTTTATCTGGCCCCAAATTTGCGCATGTGTACCCTGTCCA
AGGATATACACTATCTCTGTCCCAGTAAACCGTTCACTCGAGAGGGTGCCGGCAATC
TTTGTGGCCTTAAGCCCATGCTCACTAAGGACCAGTGTCCTGCCACTGTCACTCATA
GGTACCAACAGGTCGAGACTCAGGTTGAAATCGTAGGGAGCCGGTGGTTGGTAAAT
ACTGCGGCACGGGAAGCGGTGTTGTCCTATGACCAGCATGATATCGATGAACGTATT
CCCCTTCCGGAACAAACATTTTGGGTAGACGTACCACCTGGGGCTACTCTGCATGTT
GGGGAATGGTCCCTGTACCACTTAGTACCTGATCAATATGAGAGTGAAATTGAGGTG
TCTGATTTCTTTACTACCCACTCCTTAGAACTCAGCCCTGACACCCTTGGTCGGATAC
AGTATGAGGGTCCCCAGTCTATCGATTTGACCCCCATCAATAATGTCTTAAAAGAGA
TCGAGGCACGTCCCACATGGACAGTCCAACCTGTAAGATACTCCTGGTCATTGCCCG
ATACCTTGTTAGCTGTGCTGGTATGTCTGGGGTATGTAGCAACCTTTGGTATAACGT
GTTTTTATCGTAGGAAGACCGTGAAGCTACAGAGACAAATGGATTCATGGTCGGGT
AAGGTGGTTAGATTTGTGAGACGTAGAAGGGGGGGTCAAGAAGTGTCAGTGGAAAC
TCAGGAGGACGCAACTGATGAACCTATGCAGGTAGCTGCAGTAAGGGAAGTGGTTT
AG
> Loki-Ame consensus sequence Env protein
LLATYLLDSTSLVVHTSTPRLIMYICLVRKMRTLLLLTLVLGTIVAAPDVISPGPTSGVVL
RDNSGLLITNCRLHTQRVFVRLNPEEVCKKNFPDWSRQVGWAGHWWVREAVSHAERD
TTHMLQQLQKFAVTQSELNGAQKRSKRFLAGLLTAAAAVGSLLGVGTSAVNAVGLAT
VRRNVGELQAELPHIKDQIDTQQRQLQAIGQTLQGTIVVLNTHTVALNKTLQTITSLTRA
IQTDMAHIQVINMLLNDMLQEVGSSVDSLAMGKIPAYLVPLTLVQEILSRATREEVTDL
QAHLAYTLGSAVPLYVNPENREVAFLVNLPIIAAENIYRLKRVVNVGFWQKEAHIKIQTP

TLVAYQDINPGLYLAPNLRMCTLSKDIHYLCPSKPFTREGAGNLCGLKPMLTKDQCPAT
VTHRYQQVETQVEIVGSRWLVNTAAREAVLSYDQHDIDERIPLPEQTFWVDVPPGATLH
VGEWSLYHLVPDQYESEIEVSDFFTTHSLELSPDTLGRIQYEGPQSIDLTPINNVLKEIEAR
PTWTVQPVRYSWSLPDTLLAVLVCLGYVATFGITCFYRRKTVKLQRQMDSWSGKVVRF
VRRRRGGQEVSVETQEDATDEPMQVAAVREVV*
>Loki-Dcl consensus sequence
TGCTACGCACTGTAGCGTCTGCGTAGGATCACAGATGGTTGATCTTCAAAGAACCTT
TTAACCTGTGACATTTACATACCAATTACAACCCTGGACCCTCACTCTGATCGTCCC
GTCTTCCCCCACTGTCACAAATGTAGCGTCTGCGTAGGATCATAGATGGTTGATCTT
CAAAGAACCTTTTAACCTTTGACATTTACATACAAATTACAACCTGGACCCTCAGCC
TGATTGTCCCGTCTGCCCCCACTGTCACCAAGACGATGCGAAACAGACGGGGTCTGC
AAACTGCTCATCGGGGCTCCAAGCTTCAAAGCATTCCAAAGACCTCCTGCCTGGCCC
CACCGTCGCAATACACAATCACGCACACACTTAAAACATATTGTCTGTATCTTGTGC
TAGGAATATGTGATTATAAGTGGATCCCGTATATGCACGCGTTGTGTGTACGTTCCC
TTGTATGAATTATACTAGTTACAACTTTCTATTGTACATTAGGTAAACTTTAATTACC
TATAATTAAGTGTATAAGTGTATCTCTACTTTCCTACCTCCTAAAGGTCCCTTTCTTG
GTGTCAGAAGGAACTTCTCTTTACTACTGGAACAGTGATGCATACCATGTGATCACA
AAATACATCATACAATTTCTATAAGGTAGAAATTGTAACTGCAACAAACTACAGTCT
CTCTCTGGACCAGCCATAAACATTCCGGTCAGTCGTAAACCCAGACAAAGGGAACT
TCTCCCGCCAAAATGTCACACTTGTGATTGGCTATTCAGGCCACATGATGGGCGTGC
CAGAAAACATTCAAAAAGGAGCGTCACACAGGCACGAATGCTCAGAACTTCTCTTC
TTCCTTCCGAGACACCTTTTCTTTCAACACTTTGTTACTCTTTGGCAAAATTTACCTTC
AAGGAGAACCTCCAGCGCCAGACCTAAGGATCGAGGATCGTGACCCTACCGCATCA
GGACTCTTTTGCATGAAATAAGGACCAATGCAAGTAACTATTTAGATGCACGTTTAA
ACCCCTTTTTAAGGTGAACTTAAGAATTCTCTGACGATTCTCATTAGGCTGTGGTTAA
TTGATGTGTAATACTGCCATTCTCTTTTCATCACTTTCCTTTACCCTGTATGTATATGT
TTGTCTAATTGTTGTATGTTAGCTATAGCCGTGTTTATTAAATTCTTTAAATTCACAA
TCGTTTGTTTCTGTGTGCTGCTCACAGCTGAGTCACTGAACGTTGACCTTTTAAGCTA
CATGCTAAGTTACTGCTAGCACGCAGAGTCGGGAAAGCTACCTTTCCTGGCCAAGAG
GGGGGATAATCGTTCCCAGAATTGATAAATAATTATATTATCTGCTCCGTGGACGGA
CAGATCAAGTAATTGTAACATTAATTCTGCTACAGTTTATCCCAAAATCTAATCTAA
ATGTTTATTATTAATTATAACCAGTTATAATTAATTCTAGATATTTCCTCTTGAGCTA
ATTCGTTACATTTAATGGTGGAGAATGCGGGCACGTTCAAACTTATTTTGTGAGCAA
ACTCAGTAACAATCTTTTTGTGTTTTTATGCAAGCTTAAAGTGAGGCAAGACTCACT
CTCGTGATTGGTTGGTTCAAGACCACCTCCATATGTGTATTTCAACAGCCATTGTCTG
TTAAAGTACACAGCCATTAGTTCATAATATTAAAATTGTGCTGTAAAGACGAAGAAA
TCTCTTTGCTGTACATTTTAATATTATCAGATCGCCTGTTGGAACGAAGAAATCTTCC
TTCATCGCGATAAAGATTGATTTGAGCGACGTTCCTCTTGTCATCTTTTTATAGGCCT
TATTCGTAAATGAGAGCCCTAAATGCATATGTGTAGAAACAGCTAAGGAACCCTGTT
TTTAATTGTTAGTGTGTAAGAGGCGGCGAAGAAATCCTGCCCACACATTGCTGTGCA
ACTAACTGCATTTGTGTAAAAACGGCGAAGAAATCCCGTTTTTAATTTTTAGTGTGT
AACAGGCAGCGAAGAAATCCTGCCACACACCCCTTTTGGTGCAGCTAAATTGCATTT
GTTTTAAAGCAGCGAAGAATTCCTGTTTTTAATTGTTAGTGTGTGAAAGGCAGCGAA
GAAATCCTGCCACACCCTGTCTAACACTCCTAAATTCCAACTCCACCCCCTCTGTTCT
AACCCTTGGGGCTCACCCACGCCACCTTGTGGCCGGTCTCAGTAGGTGGGTTTTCTT
GGTTGGTTGGCCTAAGCCTGCCAGTTTATCTGAAAGGCTTAGTCTGACTCTTTGCCTC

TTAACTTGATTCTTGTTATTACTGAGTCAAGCCTCTCTCTCTCCTGTTTAGATTTATTT
AAATTTAGTCTAAACAGTGTTTGAATTTAAATTTAGAATCGATTTCTATATTCTACAC
CTGTGTGTTTACTTATAGGTAGTGCTCCCCCTCAATCCTATACCTCTCCTCTGTGTTTG
GCGTATTTGTTTGTTTGTCTTCACCTGGTTTCTGTTTGTTTACCAAACCTTTGAGAATC
ATCAGGGAAGGCTGAGTAATACATTGACAACCCAAGCTGCCAAAGTCTTAAGACTG
TCACGCGGTCAAATCTAAGTACTATAGGCGCAGTCATTAACACTGGTAGGACTTGCA
TTGGTTTTGTTTTGTGCTGTCCCTCTCTCTCGCTCCCTTTCTCTCCCTCTCTCGCCTGA
CCAGCAAGCCAACATGTCTCGATCATCCAGCCCTGAGGCCCACTGGGATCGTCTAGA
AGTCTGGCTGAACGCCCTGACTAGCAATCTTCTTCATGAACCCACCGGGGACCTGCA
GAACCTGAGCCAAGAGCAGATTGACGACAACCTGAGCATCTTGATGGCCCGCGATC
CGACGCAGGACTACAGCCACAAGGAGCTGGCCAAGATCACCGCGTCACTGGCTCAC
AACCTCCTCGGCCAGGCAAAGCTGAGTGAAGCAACAATCCACCACCAGGAACGGCA
GACTGCAGCACTAAAGCTTCAGGATGAGAGAGCACAGAGGAATCTGACGCTCGTTC
GAAGACAACTGGAGCAACTGGAGACTCAGGACAGGACAGCACACAAACAGCACCC
AGAGCTACAGGAGGAATTGGAGAGACTCCAGAACGCCTTGACTGATCTCCGCGCGG
ACACAGCACACCGGGAAGAGCAGGAAAAGGACGCCAGAGAAGAGCTGAGGGAAG
AACTCAACGAGGCTAAGACCCTCCTGATGAGAGCAGAGACTGAACTTAAAGAAAGA
GAAGCCTGGGCAAGGGCCTGTGAGGGCCACCTGCAAACTGCCCGGGCCGAGGTCAG
TACTCTGTCCCAACAAAGAGACTATCTGAAAGATGAACTTGACACAGTTCGCAGAG
AGCTCAAGCACTCATATAGACTGCACGATGACCCTAGAAGAGAGATGCACCCCGCA
GAGTTTCCCCTAACCAGCAGCAAAAGGGAGGAGAGCCCAGTGCTCAGGACATCACC
TGCCAGCATCCCAAACCCCCCGCCAGCAGCAAGGGGGTGGGAACCCTCTCAAAGAC
TACATTCCAGTCATGGCGTGTCAACCAAAGAACTGGATAAGCTGGCGAGGAACATT
CCTATTTTCACTCCCAACCCTGCAGGAGGCCATGATGTCCACTCATACCTCGATGAT
GTGGACTTCCACTTGCAAACTGTTGCTAACGTTACCACACGCGACAGACTTTATCTG
CTCAGAATCACTGCTAGCCGTGAGGTGCGCAGTTTCCTGGACCGACAGCCAGAGAG
CGTCAAAATGGATCATCAGCAACTGCAGCAAGCGCTAATCAGGGAGTTCTCTGACC
CAGAGTCAGAGCATGGACTGATCACTGCCATGGACCTTAAACAGGGCAGACAGGAA
ACCACGCAAGCCTACTACAACAGACTCCGACAGGCCTACTTCGGAGCACGCAACGA
GCCAGGAATGGAGGACGATTTCAATTTCAGATCCCTGTTCCTCCGGAACCTGCACCC
CACAGTAAGTCACTATCTGGGGGTTCTGGCATGCCCACGGACGATGTCCACCCAACA
GCTGAGAGACTTGGCCCACAAGGCCTTCGCCAAACAAAGAGCTGTTTCGGAAAAGA
CAGCGAAAACCCACACCATCTATCCTGTCTCCACTCAGAGTGCAGAACTCACTCTAG
AGGGCGCCGAGCAAAGCCACCATGACAACAGATGGTCACGAACCTTTCCAGCCAAC
AGGGGGCAGCGTGGACATGACGGGGCCCGACACAAACACCAAACTGGTCATTCCGG
AAGATCCTGGGACAGGTCACAACCACCCTCCAACCCAAGGGATGAAGCCACATGGG
AAGGAAAGAGACGACCCAGAAGGAGCCAATTTTCAGCAACACAAGCGTCAAGATC
AGATTGGCAGCAGCCGAACCCTTCTCAACACTCTCTGGATCAGTCCAGATTTAAATC
ACCAGCAGGACCAAACTGGGAAGCCACATCTGAAGCAGCAGAGATTATGAGAGTTT
TGAAGGAACTCCTCCAGGAAAGGCACCACAGGGGGGACAAAAAGGATGAGTCAGA
TCAGGCATGACTAAGCGTGCAGTCTGACCCTATCACCCAATCAAAGTCCCCTCAACA
TAATCCTCCCAACCAACTCAAACATAATCCTCTGCTGACCACAGAACGAGAGGTCCC
ACCACCAAGTGATATCACGCAACCACAGCAGATGACTACAGAACCACCAGAGCCAG
CCAATTTCCAGCATCACACCTCAAGCCAGGACCTTGATGTACCTGAGAGTGCTGTCT
TGGTGGTCTGTCTGCCGGAAGAGTCACGGGAGACTGGTCCTGACTGCTTACCGATTA
CCACACAAGCCCCAATACCAAGACTGTTGGGGAACCTAATTGAAAGGGGGGTAGCC

AAGAAACTGTACCTGACCACCACTTTGGAGAATGAGGTAAGGCTAGAGGCCCTGGT
GGACACGGGAGCAGACCTGACTCTGATGTCAGCCCAGCTTTTCGACAGACTTCAGAT
TGCTGCCAAGAGACAGAATCGAACTCTCGAATATAAGAACTGTGTGCTGAACGTAC
AATCATACAGCCAAAATGATGTACAACTAGAACAACTAGCTCCCATTCACCTAACC
ATCGGCCCTATGACCTTAGTACACCCTGTGTACATCTCTCCATTTGACACATACCCTC
TCCTTATTGGAAAAGACCTGCTAGACCGCTTTGAGCCCCTATTGGACTTCAAACATC
TAAAAATTTGGGCTCAAGTGCGAGAACCTCTGCCTCCTCAGTCACCAAGATCACGCC
AGACAGACTGCCAGGTCACAAAGGCCACAGAAGGTCTCGCCAACAACTGCGGGAAC
ACACCGACCCAAGAGCAAGGTTCAAGGTCATCACTATGCACTTTCCAGCTAGCTAA
AGACTCTGACTCCCACGGCCCTCAGGTTATGGGGGCCTTACAGCTGAATGATGTCAC
AACAACAGACTCTGTGCTTGCACTGTGGGCAGACAACTCAGCAATCAGCTCCTCACT
GTTCAATGCTCTCATACAGCAAACAGGGAACGTCCCTACTGCCAGCAAACGTACCC
GCTTCCCTCTGAATCCCTGCCTGTCAACTATGGTAACCGCCAAGTGCATCTGTGCTCT
GAACCTGAAGTGGAATGACAGATGTTTGACGCATTACTTCCTGGTCCTCCCAGACAT
ATCACATGACGTGTACATTGGAGCGGACATTTTGATTCGCCTCAGGGCTCATGTGGA
CATGATCAATGACGTGGTATGGGCTCCAACAACCTCACAGGCCCCCACTGTTCCCGT
CAAACTTAAAAACCTCCAGTCAGGTCAAACCATCCCTGAGGCCTGCACAATGGTCA
GTGAAGAAGAAACTGTAATACCAGCCTACACCAAAGGGGTCACTGTTCGACTCAAC
CTGCGATCTGGTCAGACCTGCAGCCACACATTGGGCTTCTTCCAACCCTTGCCTGAG
TGTGGGGAACTGGGCCTCACTCTAGAAGCTACACCCCTTGTTGAAATAACATCCCGG
GCAGTGTACATACTGTTCAATAACTGCACCGCAACGGATGCCAAAGTGCCTAAGGC
CCACCGCCTGGGTTGGCTAGTGAGTTACGACTTTCATGACTTCGAACTGATCCTCCC
AGTAATTGGACCCATCCCAGAGTCAATGATTCCGGAAGAAAGCTCGGGCAACTCAG
TGTCAACAGCACCATTCAAGACCATCACCATAGCTTCAGTCATGCCAGTAGTCAAAG
AGCAGGTGTGTAGAACGGAGCTGACAGAAGACCAACATCTTGCAGTGTATACAGTC
TCTAACCACACTACCAACAAAATGACAGAAGAGCCATACCCAGGTTTTGAGACCCA
AGTCCAGCAAATTCTGAAAGACGCCGATGCCATACAGGAGGAGGCAGACCGCCAAA
AGCTCCGCCAAGTCCTGTACAAATACAAAGAATCATTTGCCAAGGACCCTTTAGACT
GTGGCCTCACTAATCTTCACGTGATGCGCATACCAACACACCCGAAAGCTCCTCCCA
CTTTCGTCAAGCAATACAAGATACCCTTAGAATCGTACGAACCCGTACAAGAGATCA
TCGATTCTATGCTCGAAAAAGGCGTTATTCGGCCATGCAACAGCACCTATTCTGCCC
CAATTTGGCCTGTCCTCAAACCAAATGGCAAGTGGCGACCAACCATCGACTACAGA
AAACTGAACCAGCAGGTACCATTATCACGGTGGCCCATGACCCAGCTCGAACAGGA
AATCCCAAGAATCAGAGGGGCCACCATCTTCTCCACGCTTGATGTGGCCTCAGGATT
TTGGACTATACCGGTACATCCAGATGACCAACATAAGCTGGCTTTCACCTTCAGTAA
CCGGCAATACACCTTCACAAGGTGCCCATTCGGCTATGCCAACTCACCTGCCGAATT
TAATATTTTTCTTAACAAGGCGTGTCCAGATGCAAGAGTGAGAGGGAACCTAATCTA
TGTTGATGATGTTCTCATGAAAAGCATGACCGTGGCCGATCACCTGAAAGAAATTGA
CCACGTGTTGCACCAACTGACTGCCGCTGGTGCCAAGATCGCCCTTCACAAGGGACA
ATGGGGAAAGTCTAAGGTTAACTATGTAGGTCTACTCATCGGACCACAAGGTATCG
AGCCTCAATCCAGCCGCATTCAAGCCATCCAGACCATCAAGCCACCTGTCAATGTCT
CAGAACTCAGAAGTTTTCTCGGCGTGTGCAACTATTCACGGCAATTTATCGAAAACT
ACTCAGACATTGCACGCCCTCTCACTACACTGTTGAAGAAGGACTGCCCATTTGCCT
GGACAGAGGATCAGCAAGAAGCTATGGATGAACTGAAACGACACCTGTGCACTGCA
CCATGTCTGGCCTATCCTGATCCACAAAACGAGTTCTATTTGGACGCTGGATTCTCC
AGCCACTGTCTCAGCGCCGGGTTATACCAACGACTGGACCAGGACAAAAGGGTGGT

AGCATATGCCAGCAAGACTCTTCTTCCCCCAGAGTGTAAATACTCAGACTGTGAAAA
AGCCCTGCTCTGCACTGTGTGGGCCCTACAAAGATTCTCTAACTACATTGGAGCACA
GAAAGTTATAATAGAAACCTGCCACCAGCCGGTCACATTTCTCAACAGCCAGCGCA
TCAGAGACGGAGTGGTCACAAACGCACGCATAGCCACATGGTTAATGGCCCTCCAG
GGGCGCAATGTGGAAGCACGATACGCCCAAAACCACAAGTCACCGCTGGGTGACGG
CCTGGCAGCCTGCCAGAGTTGCTCCGATGACTCACAAACAGAGAACACAGAAGCAG
AACGGACCCAAGTACAGCCGTCTACCTGCCACAGGTATTTTGACGAGAATGTGTGCC
AAGGCATCCCCACGGCCTACGTTGATGGATGTTCCTACAACCAAGAAGGCAAACTA
AAAGCAGGTGCCGGGGTTCTCTGGCTAGACAATCGACCTTGCCCACCACAACACTTC
AAACTGGGCCCACAATCATCACAATACGCAGAGATTGCAGCCATTCTCATAACCCTG
CAAATCGCCTCGACTCATAACATCAGAGAACTGCTCATCTGCACAGACTCGAACTAT
GCCAGACTCAGCTTCATGTGCCATCTTCCAAACTGGAAACAAAATGGTTTCAAGACC
GCAAACAAAAAACCAGTTAAGCACCAACAACTGTTCCAGGCTTGTGACAACATCAC
AAGGGAACACGATATAATGGTCTACTGGAAAAAAGTGAAAGGACACTCTAAACAAC
CTGGTTTAGACAAGGACCTTAACGACCAAACAGATGCACTAGCTACAGCAGGTGCA
CTACACGGCGAGTTATGGTATCCACCTGCCCACCCACCCACCCTTAATGTGGCTGCC
ATAACCCGCAGCAAACAGACTCAGCCCACTTCAATGACCACCTCACAGCCGCTTAC
GCTGGCTCCTCAAATCACCACCATGGACATTGCTGACATGCAGGCCTCTGACCCAGC
TATACACACTATCCCTCACCACCTCTCTGACCCCTCCAACCGCCCTATCACAGCTGC
ACAAATGGCCAATACGCCAACCCTCAAAGCACTACTCAACGTGAAACCTATGCTGT
GTGTGACGCACAATATTCTTGGGTACTGCCCTGATGACGTCACTGCGCCAAGGCTTG
TGGTGCCACAGTGCCTAAGGGGGGTCATGCTAATATATGCTCACGACACACCAAGC
GCGAGACACCACGGCACAAGTACCACGTATGAGACACTAGAACAAGCGGCATACTG
GCCCAGCATGCAGCAAGATTTGGCAGAGTATGTGAAAGAATTCTTTCCTTGTTACCA
ATTCCAACCAACAAAAACAAACCACAGAGCCTCACTGCAGCGCAGAGGAGTCACAT
CTCCTTGGTCAGACCTCCAAAGTGACTGGGTTGGACCCCTGCCTAGGTCAACTAGAG
GCTACAAGTATTTCCTCACAGTCGTATTCCAGTTCACAAAGTGGGTAGAATGCCTCC
CGGCACCGAATGACACTGCACAGACTACAGCATACCTCCTGATGAATCACATATTCT
CCAGGTTCGGGCTGCCACTGAGAGTCAACTCGGACAGGGGCACCCACTTCACAGCT
GAGATCATGCAGCAAACCTGGAAACTTCTGGGAATACAAGCACGCCTTCACATAAG
TCAACACCCTATCGCGTCCAGACAAGTGGAACGTATGAACCGGACCGTAGTCAGCA
TGCTACGAAAGTTTGTGTCCCCCAACCAAAAGGACTGGGATGTGAAACTTCCACTGG
TGCTCATGGCCATCAGAGCAACTCCAAATGAAGCAACAGGTGTGTCTCCCTTCGAAC
TGATGACAGGCCGGCAAATGACATTGCCATTACACCTGCTATACCAGCCAGGAGAC
ACCAACCTTGTCACCGCCTACACCACTCACCAGTACCTTGATGAACTGCACCGGCAT
TTGAGAACCACATTTGCTTTTGCCCAACAACGACTCCAAAAGAGTGCAGAAGGCCA
GAAAGCATATTACGACCGGAAGGCATCACACCACGAACTGGACGTGGGTGACAAGG
TCTGGTATTACCGCTACACACAGCTGCCTCAGACATCCTCTCATCAGCTGTCAAAGA
AGTTTCTGCCACACTGGACTGGACCGCATGAGATCATGGACAAACTCTCTCCTGTCG
CATACCGGATCAAGCTAAGCCGGGGACAGAAAGAACCAGCTTTCAAATGGGTCCAT
CGTAACCAGATCAAGAGACATGTGGCTACTGGCAGGGGAAGACAAGGGGGGGCTA
GTACTGACTAGACCTTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACT
GAAGACCGCTCTGAACAGGCGAGAACTTCCTTTGCTGATCCACTGCCTTTCTCCTTC
CTCTCTGAAACAGGATGCTGCTGTGGCTCTCCATGCTATGCCTCCAGCTAACCAGAG
GACAACCACAACCTGACATCACATCCCCGGGCCCAGCATCTGGTATTGCACTAAGG
GACCAACCTGGACTTCTCGTCACCAACTGCAGAACATATACCCAAAAGGTGTATGTC

CGTCTGGACCCCCGAAATGTCTACAGGAAACATTACCCTGAGGCAACTGCACGCTA
CAGTTGGGCAGGTGCCCGGTGGACAGAGGATGCCCTAGGTCACGCAGAAGCTGATG
TCAAACACATGCTCACTCAACTAAAAAAGATGACTGTCACACAAGCAGAGCTAAAT
GGACACAACAAACGTAACAAACGTTTCCTGGGGGCCTTGCTCGGAGCTGCTGCTGC
CGTAGGAACCTTGTTCAACCTAGGAGTGACCAGCGTCAACTCCGTGCGCCTGTCTAC
CCTGAAGCAACATGTAGAGGAAATCCAAACCGAGGTTCCCCAACTAAGAGAACAAC
TTGCACTACAGGGACAAACCCTACAGACTATCGGCAAATCATTGAAGGGCACCTTG
GTGGTTCTGAACACTCACAGTGTCCTTTTAAACAAAACTGTGAACTCTGTGAAACAG
CTATTCTCCATTATACAGAATGACGTTGCCCAAGCTCAGCTGGTCACCACACTGATG
TCCGACATGCTACGTGAGGTGAGCTCTTCCATTGACAGCATAGCTATGGGTAGAATT
CCCCCGTATCTCGTGCCCCTGAGCCTTGTACAAGCAGTGTTATCCTCCGCTACCACTC
ATCCACCGGACTCGCTCCAAGCCCACCTGGCCTACTCCCTGGGAGGTTCCATCCTGT
TACATGTTGACCCTGAACAGAGTGAGATGGCATTTCTCCTAAACCTCCCCATCATTG
ACTCTGACAACATTTACAGACTCAAGGACATTGTCAATGTCGGATTTTGGCAAGGAA
ACACACACATCAGGATACTCACGCCAGATGTTGTAGCCTACCACGACAGTAACCCA
CAGTTATACCTGGCCCCAAACCTGCGCATGTGTACCCTGACCAAAGACATCCATTAT
CTCTGCCCCAGTAAACCCTTCCTCCGTGACAATACGGAAGGAATCTGTGGCTTACAG
GCCATGCCGAAAGAGACACTCTGCCCTGCTGAGGCCAAAACTCGAGCACACGTCAC
AGAAACGCATGCAGAGATTGTGGGGAATCAGTGGCTGGTTAACACACCCATCAGTA
CCGCCACACTGACCTATGATCAACATGACACTGCCACCCGGATCAACGTGCCCAACC
AAAGTATGTGGATTCAAGTGCCGAAAGGTGCCATCCTCCACATCGGGGACCTGGCG
TTGTACCACCTTCCGAGCGATGAGTACCAGTCCGAGGTGGAACTTCCATCTTTCTTC
AAGGATCACAGTCTCACCCTGGAGCCCGAATTGGAACAGAGCATTGAGGAAGGGGG
GTCTCAGATAATTGATATCACCCCAATCGACACAGCCCTCCAAGCATTGTCTAGGCA
ACCTGTCCAGACCAACTTTCCTGTTGCCCGATCCTGGACTGCGGCCGACACAGCTCT
TTGCCTCGCCACAGCGATCGGTTGCACCCTAACTCTCGGCCTTGCCTTCGTCCTGTTT
AACAGAGTAAATTTGGTGCAAAGGTCCATGAACCGGTGCACGGCTGCATTTCCCCG
GACCTTCAAGAAAAACCGCCGTAGGAGAGTACCCCAGGCGGAGGCAGTGCTGAACT
TGACCAAGGGGAACACCCGGGTGGAAGACCCAAAGGCTTGCTGGTCCACCTGCCAT
CTGACTTACAGGATCCATCTGTCATCCCACCTTCATGATCCAATGTGTCTGATAAGGT
TTCTAGTTGAACACTTTTCAGCCACAAGGGGGGATTTGTAGCGTCTGCGTAGGATCA
CAGATGGTTGATCTTCAAAGAACCTTTTAACCTTTGACATTTACATACCAATTACAA
CCCTGGACCCTCACTCTGATCGTCCCGTCTTCCCCCACTGTCACAAATGTAGCGTCTG
CGTAGGATCATAGATGGTTGATCTTCAAAGAACCTTTTAACCTTTGACATTTACATA
CAAATTACAACCTGGACCCTCAGCCTGATTGTCCCGTCTGCCCCCACTGTCACCAAG
ACGATGCGAAGCAGACGGGGTCTGCAAACTGCTCATCGGGGCTCCAAGCTTCAAAG
CATTCCAAAGACCTCCTGCCTGGCCCCACCGTCGCAATACACAATCACGCACACACT
TAAAACATATTGTCTGTATCTTGTGCTAGGAATATGTGATTATAAGTGGATCCCGTA
TATGCACGCGTTGTGTGTACGTTCCCTTGTATGAATTATACTAGTTACAACTTTCTAT
TGTACATTAGGTAAACTTTAATTACCTATAATTAAGTGTATAAGTGTATCTCTACTTT
CCTACCTCCTAAAGGTCCCTTTCTTGGTGTCAGAAGGAACTTCTCTTTACTACTGGAA
CAGTGATGCATACCATGTGATCACAAAATACATCATACAATTTCTATAAGGTAGAAA
TTGTAACTGCAACAAACTACAGTCTCTCTCTGGACCAGCCATAAACATTCCGGTCAG
TCGTAAACCCAGACAAAGGGAACTTCTCCCGCCAAAATGTCACACTTGTGATTGGCT
ATTCAGGCCACATGATGGGCGTGCCAGAAAACATTCAAAAAGGAGCGTCACACAGG
CACGAATGCTCAGAACTTCTCTTCTTCCTTCCGAGACACCTCTTTCAACACTTTGACT

CTTTGGCAAAATTTACCTTCAAGGAGAACCTCCAGCGCCAGACCTAAGGATCGAGG
ATCGTGACCCTACCGCATCAGGACTCTTTGCATGAAATAAGGACCAATGCAAGTAAC
TATTTAGATGCACGTCTAAACCCCTTTTTAAGGTGAACTTAAGAATTCTCTGACGATT
CTCATTAGGCTGTGGTTAATTGATGTGTAATACTGCCATTCTCTTTTCATCACTTTCCT
TTACCCTGTATGTATATGTTTGTCTAATTGTTGTATGTTAGCTATAGCCGTGTTTATTA
AAATCTTTAAATTCACAATCGTTTGTTTCTGTGTGCTGCTCACAGCTGAGTCACTGAA
CGTTGACCTTTTAAGCTACATGCTAAGTTACTGCTAGTACGCAGAGGGGAAAGCTAC
CTTTCCTGGCAAGAGGGGGGGATAATCGTTCCCAGAATTGATAAATAATTATATTAT
CTGCTCCGTGGACGGACAGATCAAGTAATTGTAACATTAATTCTGCTACAGTTTATC
CCCAAATCTAGTCTAAATGTTTATTATTAATTATAACCAGTTATAATTAATTCTAGAT
ATTTCCTCTTGAGCTAATTCGTTACAGCACCCTGA
> Loki-Dcl consensus sequence 5’LTR
TGTAGCGTCTGCGTAGGATCACAGATGGTTGATCTTCAAAGAACCTTTTAACCTGTG
ACATTTACATACCAATTACAACCCTGGACCCTCACTCTGATCGTCCCGTCTTCCCCCA
CTGTCACAAATGTAGCGTCTGCGTAGGATCATAGATGGTTGATCTTCAAAGAACCTT
TTAACCTTTGACATTTACATACAAATTACAACCTGGACCCTCAGCCTGATTGTCCCGT
CTGCCCCCACTGTCACCAAGACGATGCGAAACAGACGGGGTCTGCAAACTGCTCAT
CGGGGCTCCAAGCTTCAAAGCATTCCAAAGACCTCCTGCCTGGCCCCACCGTCGCAA
TACACAATCACGCACACACTTAAAACATATTGTCTGTATCTTGTGCTAGGAATATGT
GATTATAAGTGGATCCCGTATATGCACGCGTTGTGTGTACGTTCCCTTGTATGAATTA
TACTAGTTACAACTTTCTATTGTACATTAGGTAAACTTTAATTACCTATAATTAAGTG
TATAAGTGTATCTCTACTTTCCTACCTCCTAAAGGTCCCTTTCTTGGTGTCAGAAGGA
ACTTCTCTTTACTACTGGAACAGTGATGCATACCATGTGATCACAAAATACATCATA
CAATTTCTATAAGGTAGAAATTGTAACTGCAACAAACTACAGTCTCTCTCTGGACCA
GCCATAAACATTCCGGTCAGTCGTAAACCCAGACAAAGGGAACTTCTCCCGCCAAA
ATGTCACACTTGTGATTGGCTATTCAGGCCACATGATGGGCGTGCCAGAAAACATTC
AAAAAGGAGCGTCACACAGGCACGAATGCTCAGAACTTCTCTTCTTCCTTCCGAGAC
ACCTTTTCTTTCAACACTTTGTTACTCTTTGGCAAAATTTACCTTCAAGGAGAACCTC
CAGCGCCAGACCTAAGGATCGAGGATCGTGACCCTACCGCATCAGGACTCTTTTGCA
TGAAATAAGGACCAATGCAAGTAACTATTTAGATGCACGTTTAAACCCCTTTTTAAG
GTGAACTTAAGAATTCTCTGACGATTCTCATTAGGCTGTGGTTAATTGATGTGTAAT
ACTGCCATTCTCTTTTCATCACTTTCCTTTACCCTGTATGTATATGTTTGTCTAATTGT
TGTATGTTAGCTATAGCCGTGTTTATTAAATTCTTTAAATTCACAATCGTTTGTTTCT
GTGTGCTGCTCACAGCTGAGTCACTGAACGTTGACCTTTTAAGCTACATGCTAAGTT
ACTGCTAGCACGCAGAGTCGGGAAAGCTACCTTTCCTGGCCAAGAGGGGGGATAAT
CGTTCCCAGAATTGATAAATAATTATATTATCTGCTCCGTGGACGGACAGATCAAGT
AATTGTAACATTAATTCTGCTACAGTTTATCCCAAAATCTAATCTAAATGTTTATTAT
TAATTATAACCAGTTATAATTAATTCTAGATATTTCCTCTTGAGCTAATTCGTTACA
> Loki-Dcl consensus sequence 3’LTR
TGTAGCGTCTGCGTAGGATCACAGATGGTTGATCTTCAAAGAACCTTTTAACCTTTG
ACATTTACATACCAATTACAACCCTGGACCCTCACTCTGATCGTCCCGTCTTCCCCCA
CTGTCACAAATGTAGCGTCTGCGTAGGATCATAGATGGTTGATCTTCAAAGAACCTT
TTAACCTTTGACATTTACATACAAATTACAACCTGGACCCTCAGCCTGATTGTCCCGT
CTGCCCCCACTGTCACCAAGACGATGCGAAGCAGACGGGGTCTGCAAACTGCTCAT
CGGGGCTCCAAGCTTCAAAGCATTCCAAAGACCTCCTGCCTGGCCCCACCGTCGCAA
TACACAATCACGCACACACTTAAAACATATTGTCTGTATCTTGTGCTAGGAATATGT

GATTATAAGTGGATCCCGTATATGCACGCGTTGTGTGTACGTTCCCTTGTATGAATTA
TACTAGTTACAACTTTCTATTGTACATTAGGTAAACTTTAATTACCTATAATTAAGTG
TATAAGTGTATCTCTACTTTCCTACCTCCTAAAGGTCCCTTTCTTGGTGTCAGAAGGA
ACTTCTCTTTACTACTGGAACAGTGATGCATACCATGTGATCACAAAATACATCATA
CAATTTCTATAAGGTAGAAATTGTAACTGCAACAAACTACAGTCTCTCTCTGGACCA
GCCATAAACATTCCGGTCAGTCGTAAACCCAGACAAAGGGAACTTCTCCCGCCAAA
ATGTCACACTTGTGATTGGCTATTCAGGCCACATGATGGGCGTGCCAGAAAACATTC
AAAAAGGAGCGTCACACAGGCACGAATGCTCAGAACTTCTCTTCTTCCTTCCGAGAC
ACCTCTTTCAACACTTTGACTCTTTGGCAAAATTTACCTTCAAGGAGAACCTCCAGC
GCCAGACCTAAGGATCGAGGATCGTGACCCTACCGCATCAGGACTCTTTGCATGAA
ATAAGGACCAATGCAAGTAACTATTTAGATGCACGTCTAAACCCCTTTTTAAGGTGA
ACTTAAGAATTCTCTGACGATTCTCATTAGGCTGTGGTTAATTGATGTGTAATACTGC
CATTCTCTTTTCATCACTTTCCTTTACCCTGTATGTATATGTTTGTCTAATTGTTGTAT
GTTAGCTATAGCCGTGTTTATTAAAATCTTTAAATTCACAATCGTTTGTTTCTGTGTG
CTGCTCACAGCTGAGTCACTGAACGTTGACCTTTTAAGCTACATGCTAAGTTACTGC
TAGTACGCAGAGGGGAAAGCTACCTTTCCTGGCAAGAGGGGGGGATAATCGTTCCC
AGAATTGATAAATAATTATATTATCTGCTCCGTGGACGGACAGATCAAGTAATTGTA
ACATTAATTCTGCTACAGTTTATCCCCAAATCTAGTCTAAATGTTTATTATTAATTAT
AACCAGTTATAATTAATTCTAGATATTTCCTCTTGAGCTAATTCGTTACA
> Loki-Dcl consensus sequence *smc-gag* gene
CTGGTAGGACTTGCATTGGTTTTGTTTTGTGCTGTCCCTCTCTCTCGCTCCCTTTCTCT
CCCTCTCTCGCCTGACCAGCAAGCCAACATGTCTCGATCATCCAGCCCTGAGGCCCA
CTGGGATCGTCTAGAAGTCTGGCTGAACGCCCTGACTAGCAATCTTCTTCATGAACC
CACCGGGGACCTGCAGAACCTGAGCCAAGAGCAGATTGACGACAACCTGAGCATCT
TGATGGCCCGCGATCCGACGCAGGACTACAGCCACAAGGAGCTGGCCAAGATCACC
GCGTCACTGGCTCACAACCTCCTCGGCCAGGCAAAGCTGAGTGAAGCAACAATCCA
CCACCAGGAACGGCAGACTGCAGCACTAAAGCTTCAGGATGAGAGAGCACAGAGG
AATCTGACGCTCGTTCGAAGACAACTGGAGCAACTGGAGACTCAGGACAGGACAGC
ACACAAACAGCACCCAGAGCTACAGGAGGAATTGGAGAGACTCCAGAACGCCTTGA
CTGATCTCCGCGCGGACACAGCACACCGGGAAGAGCAGGAAAAGGACGCCAGAGA
AGAGCTGAGGGAAGAACTCAACGAGGCTAAGACCCTCCTGATGAGAGCAGAGACTG
AACTTAAAGAAAGAGAAGCCTGGGCAAGGGCCTGTGAGGGCCACCTGCAAACTGCC
CGGGCCGAGGTCAGTACTCTGTCCCAACAAAGAGACTATCTGAAAGATGAACTTGA
CACAGTTCGCAGAGAGCTCAAGCACTCATATAGACTGCACGATGACCCTAGAAGAG
AGATGCACCCCGCAGAGTTTCCCCTAACCAGCAGCAAAAGGGAGGAGAGCCCAGTG
CTCAGGACATCACCTGCCAGCATCCCAAACCCCCCGCCAGCAGCAAGGGGGTGGGA
ACCCTCTCAAAGACTACATTCCAGTCATGGCGTGTCAACCAAAGAACTGGATAAGCT
GGCGAGGAACATTCCTATTTTCACTCCCAACCCTGCAGGAGGCCATGATGTCCACTC
ATACCTCGATGATGTGGACTTCCACTTGCAAACTGTTGCTAACGTTACCACACGCGA
CAGACTTTATCTGCTCAGAATCACTGCTAGCCGTGAGGTGCGCAGTTTCCTGGACCG
ACAGCCAGAGAGCGTCAAAATGGATCATCAGCAACTGCAGCAAGCGCTAATCAGGG
AGTTCTCTGACCCAGAGTCAGAGCATGGACTGATCACTGCCATGGACCTTAAACAGG
GCAGACAGGAAACCACGCAAGCCTACTACAACAGACTCCGACAGGCCTACTTCGGA
GCACGCAACGAGCCAGGAATGGAGGACGATTTCAATTTCAGATCCCTGTTCCTCCGG
AACCTGCACCCCACAGTAAGTCACTATCTGGGGGTTCTGGCATGCCCACGGACGATG
TCCACCCAACAGCTGAGAGACTTGGCCCACAAGGCCTTCGCCAAACAAAGAGCTGT

TTCGGAAAAGACAGCGAAAACCCACACCATCTATCCTGTCTCCACTCAGAGTGCAG
AACTCACTCTAGAGGGCGCCGAGCAAAGCCACCATGACAACAGATGGTCACGAACC
TTTCCAGCCAACAGGGGGCAGCGTGGACATGACGGGGCCCGACACAAACACCAAAC
TGGTCATTCCGGAAGATCCTGGGACAGGTCACAACCACCCTCCAACCCAAGGGATG
AAGCCACATGGGAAGGAAAGAGACGACCCAGAAGGAGCCAATTTTCAGCAACACA
AGCGTCAAGATCAGATTGGCAGCAGCCGAACCCTTCTCAACACTCTCTGGATCAGTC
CAGATTTAAATCACCAGCAGGACCAAACTGGGAAGCCACATCTGAAGCAGCAGAGA
TTATGAGAGTTTTGAAGGAACTCCTCCAGGAAAGGCACCACAGGGGGGACAAAAAG
GATGAGTCAGATCAGGCATGA
\> Loki-Dcl consensus sequence Smc-Gag polyprotein
LVGLALVLFCAVPLSRSLSLPLSPDQQANMSRSSSPEAHWDRLEVWLNALTSNLLHEPT
GDLQNLSQEQIDDNLSILMARDPTQDYSHKELAKITASLAHNLLGQAKLSEATIHHQER
QTAALKLQDERAQRNLTLVRRQLEQLETQDRTAHKQHPELQEELERLQNALTDLRADT
AHREEQEKDAREELREELNEAKTLLMRAETELKEREAWARACEGHLQTARAEVSTLSQ
QRDYLKDELDTVRRELKHSYRLHDDPRREMHPAEFPLTSSKREESPVLRTSPASIPNPPP
AARGWEPSQRLHSSHGVSTKELDKLARNIPIFTPNPAGGHDVHSYLDDVDFHLQTVANV
TTRDRLYLLRITASREVRSFLDRQPESVKMDHQQLQQALIREFSDPESEHGLITAMDLKQ
GRQETTQAYYNRLRQAYFGARNEPGMEDDFNFRSLFLRNLHPTVSHYLGVLACPRTMS
TQQLRDLAHKAFAKQRAVSEKTAKTHTIYPVSTQSAELTLEGAEQSHHDNRWSRTFPA
NRGQRGHDGARHKHQTGHSGRSWDRSQPPSNPRDEATWEGKRRPRRSQFSATQASRS
DWQQPNPSQHSLDQSRFKSPAGPNWEATSEAAEIMRVLKELLQERHHRGDKKDESDQA
*
\> Loki-Dcl consensus sequence *pol* gene
CTGCTGACCACAGAACGAGAGGTCCCACCACCAAGTGATATCACGCAACCACAGCA
GATGACTACAGAACCACCAGAGCCAGCCAATTTCCAGCATCACACCTCAAGCCAGG
ACCTTGATGTACCTGAGAGTGCTGTCTTGGTGGTCTGTCTGCCGGAAGAGTCACGGG
AGACTGGTCCTGACTGCTTACCGATTACCACACAAGCCCCAATACCAAGACTGTTGG
GGAACCTAATTGAAAGGGGGGTAGCCAAGAAACTGTACCTGACCACCACTTTGGAG
AATGAGGTAAGGCTAGAGGCCCTGGTGGACACGGGAGCAGACCTGACTCTGATGTC
AGCCCAGCTTTTCGACAGACTTCAGATTGCTGCCAAGAGACAGAATCGAACTCTCGA
ATATAAGAACTGTGTGCTGAACGTACAATCATACAGCCAAAATGATGTACAACTAG
AACAACTAGCTCCCATTCACCTAACCATCGGCCCTATGACCTTAGTACACCCTGTGT
ACATCTCTCCATTTGACACATACCCTCTCCTTATTGGAAAAGACCTGCTAGACCGCTT
TGAGCCCCTATTGGACTTCAAACATCTAAAAATTTGGGCTCAAGTGCGAGAACCTCT
GCCTCCTCAGTCACCAAGATCACGCCAGACAGACTGCCAGGTCACAAAGGCCACAG
AAGGTCTCGCCAACAACTGCGGGAACACACCGACCCAAGAGCAAGGTTCAAGGTCA
TCACTATGCACTTTCCAGCTAGCTAAAGACTCTGACTCCCACGGCCCTCAGGTTATG
GGGGCCTTACAGCTGAATGATGTCACAACAACAGACTCTGTGCTTGCACTGTGGGCA
GACAACTCAGCAATCAGCTCCTCACTGTTCAATGCTCTCATACAGCAAACAGGGAAC
GTCCCTACTGCCAGCAAACGTACCCGCTTCCCTCTGAATCCCTGCCTGTCAACTATG
GTAACCGCCAAGTGCATCTGTGCTCTGAACCTGAAGTGGAATGACAGATGTTTGACG
CATTACTTCCTGGTCCTCCCAGACATATCACATGACGTGTACATTGGAGCGGACATT
TTGATTCGCCTCAGGGCTCATGTGGACATGATCAATGACGTGGTATGGGCTCCAACA
ACCTCACAGGCCCCCACTGTTCCCGTCAAACTTAAAAACCTCCAGTCAGGTCAAACC
ATCCCTGAGGCCTGCACAATGGTCAGTGAAGAAGAAACTGTAATACCAGCCTACAC
CAAAGGGGTCACTGTTCGACTCAACCTGCGATCTGGTCAGACCTGCAGCCACACATT

GGGCTTCTTCCAACCCTTGCCTGAGTGTGGGGAACTGGGCCTCACTCTAGAAGCTAC
ACCCCTTGTTGAAATAACATCCCGGGCAGTGTACATACTGTTCAATAACTGCACCGC
AACGGATGCCAAAGTGCCTAAGGCCCACCGCCTGGGTTGGCTAGTGAGTTACGACTT
TCATGACTTCGAACTGATCCTCCCAGTAATTGGACCCATCCCAGAGTCAATGATTCC
GGAAGAAAGCTCGGGCAACTCAGTGTCAACAGCACCATTCAAGACCATCACCATAG
CTTCAGTCATGCCAGTAGTCAAAGAGCAGGTGTGTAGAACGGAGCTGACAGAAGAC
CAACATCTTGCAGTGTATACAGTCTCTAACCACACTACCAACAAAATGACAGAAGA
GCCATACCCAGGTTTTGAGACCCAAGTCCAGCAAATTCTGAAAGACGCCGATGCCA
TACAGGAGGAGGCAGACCGCCAAAAGCTCCGCCAAGTCCTGTACAAATACAAAGAA
TCATTTGCCAAGGACCCTTTAGACTGTGGCCTCACTAATCTTCACGTGATGCGCATA
CCAACACACCCGAAAGCTCCTCCCACTTTCGTCAAGCAATACAAGATACCCTTAGAA
TCGTACGAACCCGTACAAGAGATCATCGATTCTATGCTCGAAAAAGGCGTTATTCGG
CCATGCAACAGCACCTATTCTGCCCCAATTTGGCCTGTCCTCAAACCAAATGGCAAG
TGGCGACCAACCATCGACTACAGAAAACTGAACCAGCAGGTACCATTATCACGGTG
GCCCATGACCCAGCTCGAACAGGAAATCCCAAGAATCAGAGGGGCCACCATCTTCT
CCACGCTTGATGTGGCCTCAGGATTTTGGACTATACCGGTACATCCAGATGACCAAC
ATAAGCTGGCTTTCACCTTCAGTAACCGGCAATACACCTTCACAAGGTGCCCATTCG
GCTATGCCAACTCACCTGCCGAATTTAATATTTTTCTTAACAAGGCGTGTCCAGATG
CAAGAGTGAGAGGGAACCTAATCTATGTTGATGATGTTCTCATGAAAAGCATGACC
GTGGCCGATCACCTGAAAGAAATTGACCACGTGTTGCACCAACTGACTGCCGCTGGT
GCCAAGATCGCCCTTCACAAGGGACAATGGGGAAAGTCTAAGGTTAACTATGTAGG
TCTACTCATCGGACCACAAGGTATCGAGCCTCAATCCAGCCGCATTCAAGCCATCCA
GACCATCAAGCCACCTGTCAATGTCTCAGAACTCAGAAGTTTTCTCGGCGTGTGCAA
CTATTCACGGCAATTTATCGAAAACTACTCAGACATTGCACGCCCTCTCACTACACT
GTTGAAGAAGGACTGCCCATTTGCCTGGACAGAGGATCAGCAAGAAGCTATGGATG
AACTGAAACGACACCTGTGCACTGCACCATGTCTGGCCTATCCTGATCCACAAAACG
AGTTCTATTTGGACGCTGGATTCTCCAGCCACTGTCTCAGCGCCGGGTTATACCAAC
GACTGGACCAGGACAAAAGGGTGGTAGCATATGCCAGCAAGACTCTTCTTCCCCCA
GAGTGTAAATACTCAGACTGTGAAAAAGCCCTGCTCTGCACTGTGTGGGCCCTACAA
AGATTCTCTAACTACATTGGAGCACAGAAAGTTATAATAGAAACCTGCCACCAGCC
GGTCACATTTCTCAACAGCCAGCGCATCAGAGACGGAGTGGTCACAAACGCACGCA
TAGCCACATGGTTAATGGCCCTCCAGGGGCGCAATGTGGAAGCACGATACGCCCAA
AACCACAAGTCACCGCTGGGTGACGGCCTGGCAGCCTGCCAGAGTTGCTCCGATGA
CTCACAAACAGAGAACACAGAAGCAGAACGGACCCAAGTACAGCCGTCTACCTGCC
ACAGGTATTTTGACGAGAATGTGTGCCAAGGCATCCCCACGGCCTACGTTGATGGAT
GTTCCTACAACCAAGAAGGCAAACTAAAAGCAGGTGCCGGGGTTCTCTGGCTAGAC
AATCGACCTTGCCCACCACAACACTTCAAACTGGGCCCACAATCATCACAATACGCA
GAGATTGCAGCCATTCTCATAACCCTGCAAATCGCCTCGACTCATAACATCAGAGAA
CTGCTCATCTGCACAGACTCGAACTATGCCAGACTCAGCTTCATGTGCCATCTTCCA
AACTGGAAACAAAATGGTTTCAAGACCGCAAACAAAAAACCAGTTAAGCACCAACA
ACTGTTCCAGGCTTGTGACAACATCACAAGGGAACACGATATAATGGTCTACTGGA
AAAAAGTGAAAGGACACTCTAAACAACCTGGTTTAGACAAGGACCTTAACGACCAA
ACAGATGCACTAGCTACAGCAGGTGCACTACACGGCGAGTTATGGTATCCACCTGC
CCACCCACCCACCCTTAATGTGGCTGCCATAACCCGCAGCAAACAGACTCAGCCCAC
TTCAATGACCACCTCACAGCCGCTTACGCTGGCTCCTCAAATCACCACCATGGACAT
TGCTGACATGCAGGCCTCTGACCCAGCTATACACACTATCCCTCACCACCTCTCTGA

CCCCTCCAACCGCCCTATCACAGCTGCACAAATGGCCAATACGCCAACCCTCAAAGC
ACTACTCAACGTGAAACCTATGCTGTGTGTGACGCACAATATTCTTGGGTACTGCCC
TGATGACGTCACTGCGCCAAGGCTTGTGGTGCCACAGTGCCTAAGGGGGGTCATGCT
AATATATGCTCACGACACACCAAGCGCGAGACACCACGGCACAAGTACCACGTATG
AGACACTAGAACAAGCGGCATACTGGCCCAGCATGCAGCAAGATTTGGCAGAGTAT
GTGAAAGAATTCTTTCCTTGTTACCAATTCCAACCAACAAAAACAAACCACAGAGCC
TCACTGCAGCGCAGAGGAGTCACATCTCCTTGGTCAGACCTCCAAAGTGACTGGGTT
GGACCCCTGCCTAGGTCAACTAGAGGCTACAAGTATTTCCTCACAGTCGTATTCCAG
TTCACAAAGTGGGTAGAATGCCTCCCGGCACCGAATGACACTGCACAGACTACAGC
ATACCTCCTGATGAATCACATATTCTCCAGGTTCGGGCTGCCACTGAGAGTCAACTC
GGACAGGGGCACCCACTTCACAGCTGAGATCATGCAGCAAACCTGGAAACTTCTGG
GAATACAAGCACGCCTTCACATAAGTCAACACCCTATCGCGTCCAGACAAGTGGAA
CGTATGAACCGGACCGTAGTCAGCATGCTACGAAAGTTTGTGTCCCCCAACCAAAA
GGACTGGGATGTGAAACTTCCACTGGTGCTCATGGCCATCAGAGCAACTCCAAATG
AAGCAACAGGTGTGTCTCCCTTCGAACTGATGACAGGCCGGCAAATGACATTGCCAT
TACACCTGCTATACCAGCCAGGAGACACCAACCTTGTCACCGCCTACACCACTCACC
AGTACCTTGATGAACTGCACCGGCATTTGAGAACCACATTTGCTTTTGCCCAACAAC
GACTCCAAAAGAGTGCAGAAGGCCAGAAAGCATATTACGACCGGAAGGCATCACA
CCACGAACTGGACGTGGGTGACAAGGTCTGGTATTACCGCTACACACAGCTGCCTC
AGACATCCTCTCATCAGCTGTCAAAGAAGTTTCTGCCACACTGGACTGGACCGCATG
AGATCATGGACAAACTCTCTCCTGTCGCATACCGGATCAAGCTAAGCCGGGGACAG
AAAGAACCAGCTTTCAAATGGGTCCATCGTAACCAGATCAAGAGACATGTGGCTAC
TGGCAGGGGAAGACAAGGGGGGGCTAGTACTGACTAG
> Loki-Dcl consensus sequence Pol protein
LLTTEREVPPPSDITQPQQMTTEPPEPANFQHHTSSQDLDVPESAVLVVCLPEESRETGPD
CLPITTQAPIPRLLGNLIERGVAKKLYLTTTLENEVRLEALVDTGADLTLMSAQLFDRLQI
AAKRQNRTLEYKNCVLNVQSYSQNDVQLEQLAPIHLTIGPMTLVHPVYISPFDTYPLLIG
KDLLDRFEPLLDFKHLKIWAQVREPLPPQSPRSRQTDCQVTKATEGLANNCGNTPTQEQ
GSRSSLCTFQLAKDSDSHGPQVMGALQLNDVTTTDSVLALWADNSAISSSLFNALIQQT
GNVPTASKRTRFPLNPCLSTMVTAKCICALNLKWNDRCLTHYFLVLPDISHDVYIGADIL
IRLRAHVDMINDVVWAPTTSQAPTVPVKLKNLQSGQTIPEACTMVSEEETVIPAYTKGV
TVRLNLRSGQTCSHTLGFFQPLPECGELGLTLEATPLVEITSRAVYILFNNCTATDAKVPK
AHRLGWLVSYDFHDFELILPVIGPIPESMIPEESSGNSVSTAPFKTITIASVMPVVKEQVCR
TELTEDQHLAVYTVSNHTTNKMTEEPYPGFETQVQQILKDADAIQEEADRQKLRQVLY
KYKESFAKDPLDCGLTNLHVMRIPTHPKAPPTFVKQYKIPLESYEPVQEIIDSMLEKGVIR
PCNSTYSAPIWPVLKPNGKWRPTIDYRKLNQQVPLSRWPMTQLEQEIPRIRGATIFSTLD
VASGFWTIPVHPDDQHKLAFTFSNRQYTFTRCPFGYANSPAEFNIFLNKACPDARVRGN
LIYVDDVLMKSMTVADHLKEIDHVLHQLTAAGAKIALHKGQWGKSKVNYVGLLIGPQ
GIEPQSSRIQAIQTIKPPVNVSELRSFLGVCNYSRQFIENYSDIARPLTTLLKKDCPFAWTE
DQQEAMDELKRHLCTAPCLAYPDPQNEFYLDAGFSSHCLSAGLYQRLDQDKRVVAYA
SKTLLPPECKYSDCEKALLCTVWALQRFSNYIGAQKVIIETCHQPVTFLNSQRIRDGVVT
NARIATWLMALQGRNVEARYAQNHKSPLGDGLAACQSCSDDSQTENTEAERTQVQPST
CHRYFDENVCQGIPTAYVDGCSYNQEGKLKAGAGVLWLDNRPCPPQHFKLGPQSSQYA
EIAAILITLQIASTHNIRELLICTDSNYARLSFMCHLPNWKQNGFKTANKKPVKHQQLFQ
ACDNITREHDIMVYWKKVKGHSKQPGLDKDLNDQTDALATAGALHGELWYPPAHPPT
LNVAAITRSKQTQPTSMTTSQPLTLAPQITTMDIADMQASDPAIHTIPHHLSDPSNRPITA

AQMANTPTLKALLNVKPMLCVTHNILGYCPDDVTAPRLVVPQCLRGVMLIYAHDTPSA
RHHGTSTTYETLEQAAYWPSMQQDLAEYVKEFFPCYQFQPTKTNHRASLQRRGVTSPW
SDLQSDWVGPLPRSTRGYKYFLTVVFQFTKWVECLPAPNDTAQTTAYLLMNHIFSRFGL
PLRVNSDRGTHFTAEIMQQTWKLLGIQARLHISQHPIASRQVERMNRTVVSMLRKFVSP
NQKDWDVKLPLVLMAIRATPNEATGVSPFELMTGRQMTLPLHLLYQPGDTNLVTAYTT
HQYLDELHRHLRTTFAFAQQRLQKSAEGQKAYYDRKASHHELDVGDKVWYYRYTQLP
QTSSHQLSKKFLPHWTGPHEIMDKLSPVAYRIKLSRGQKEPAFKWVHRNQIKRHVATGR
GRQGGASTD*
> Loki-Dcl consensus sequence *env* gene
ATGCTGCTGTGGCTCTCCATGCTATGCCTCCAGCTAACCAGAGGACAACCACAACCT
GACATCACATCCCCGGGCCCAGCATCTGGTATTGCACTAAGGGACCAACCTGGACTT
CTCGTCACCAACTGCAGAACATATACCCAAAAGGTGTATGTCCGTCTGGACCCCCGA
AATGTCTACAGGAAACATTACCCTGAGGCAACTGCACGCTACAGTTGGGCAGGTGC
CCGGTGGACAGAGGATGCCCTAGGTCACGCAGAAGCTGATGTCAAACACATGCTCA
CTCAACTAAAAAAGATGACTGTCACACAAGCAGAGCTAAATGGACACAACAAACGT
AACAAACGTTTCCTGGGGGCCTTGCTCGGAGCTGCTGCCGTAGGAACCTTGTTC
AACCTAGGAGTGACCAGCGTCAACTCCGTGCGCCTGTCTACCCTGAAGCAACATGTA
GAGGAAATCCAAACCGAGGTTCCCCAACTAAGAGAACAACTTGCACTACAGGGACA
AACCCTACAGACTATCGGCAAATCATTGAAGGGCACCTTGGTGGTTCTGAACACTCA
CAGTGTCCTTTTAAACAAAACTGTGAACTCTGTGAAACAGCTATTCTCCATTATACA
GAATGACGTTGCCCAAGCTCAGCTGGTCACCACACTGATGTCCGACATGCTACGTGA
GGTGAGCTCTTCCATTGACAGCATAGCTATGGGTAGAATTCCCCCGTATCTCGTGCC
CCTGAGCCTTGTACAAGCAGTGTTATCCTCCGCTACCACTCATCCACCGGACTCGCT
CCAAGCCCACCTGGCCTACTCCCTGGGAGGTTCCATCCTGTTACATGTTGACCCTGA
ACAGAGTGAGATGGCATTTCTCCTAAACCTCCCCATCATTGACTCTGACAACATTTA
CAGACTCAAGGACATTGTCAATGTCGGATTTTGGCAAGGAAACACACACATCAGGA
TACTCACGCCAGATGTTGTAGCCTACCACGACAGTAACCCACAGTTATACCTGGCCC
CAAACCTGCGCATGTGTACCCTGACCAAAGACATCCATTATCTCTGCCCCAGTAAAC
CCTTCCTCCGTGACAATACGGAAGGAATCTGTGGCTTACAGGCCATGCCGAAAGAG
ACACTCTGCCCTGCTGAGGCCAAAACTCGAGCACACGTCACAGAAACGCATGCAGA
GATTGTGGGGAATCAGTGGCTGGTTAACACACCCATCAGTACCGCCACACTGACCTA
TGATCAACATGACACTGCCACCCGGATCAACGTGCCCAACCAAAGTATGTGGATTC
AAGTGCCGAAAGGTGCCATCCTCCACATCGGGGACCTGGCGTTGTACCACCTTCCGA
GCGATGAGTACCAGTCCGAGGTGGAACTTCCATCTTTCTTCAAGGATCACAGTCTCA
CCCTGGAGCCCGAATTGGAACAGAGCATTGAGGAAGGGGGGTCTCAGATAATTGAT
ATCACCCCAATCGACACAGCCCTCCAAGCATTGTCTAGGCAACCTGTCCAGACCAAC
TTTCCTGTTGCCCGATCCTGGACTGCGGCCGACACAGCTCTTTGCCTCGCCACAGCG
ATCGGTTGCACCCTAACTCTCGGCCTTGCCTTCGTCCTGTTTAACAGAGTAAATTTGG
TGCAAAGGTCCATGAACCGGTGCACGGCTGCATTTCCCCGGACCTTCAAGAAAAAC
CGCCGTAGGAGAGTACCCCAGGCGGAGGCAGTGCTGAACTTGACCAAGGGGAACAC
CCGGGTGGAAGACCCAAAGGCTTGCTGGTCCACCTGCCATCTGACTTACAGGATCCA
TCTGTCATCCCACCTTCATGATCCAATGTGTCTGATAAGGTTTCTAGTTGAACACTTT
TCAGCCACAAGGGGGGATTTGTAG
> Loki-Dcl consensus sequence Env protein
MLLWLSMLCLQLTRGQPQPDITSPGPASGIALRDQPGLLVTNCRTYTQKVYVRLDPRNV
YRKHYPEATARYSWAGARWTEDALGHAEADVKHMLTQLKKMTVTQAELNGHNKRN

KRFLGALLGAAAAVGTLFNLGVTSVNSVRLSTLKQHVEEIQTEVPQLREQLALQGQTLQ
TIGKSLKGTLVVLNTHSVLLNKTVNSVKQLFSIIQNDVAQAQLVTTLMSDMLREVSSSID
SIAMGRIPPYLVPLSLVQAVLSSATTHPPDSLQAHLAYSLGGSILLHVDPEQSEMAFLLNL
PIIDSDNIYRLKDIVNVGFWQGNTHIRILTPDVVAYHDSNPQLYLAPNLRMCTLTKDIHY
LCPSKPFLRDNTEGICGLQAMPKETLCPAEAKTRAHVTETHAEIVGNQWLVNTPISTATL
TYDQHDTATRINVPNQSMWIQVPKGAILHIGDLALYHLPSDEYQSEVELPSFFKDHSLTL
EPELEQSIEEGGSQIIDITPIDTALQALSRQPVQTNFPVARSWTAADTALCLATAIGCTLTL
GLAFVLFNRVNLVQRSMNRCTAAFPRTFKKNRRRRVPQAEAVLNLTKGNTRVEDPKAC
WSTCHLTYRIHLSSHLHDPMCLIRFLVEHFSATRGDL*
>Loki-Lch consensus sequence
TCCAGACAGAGAAAGTAGATCAATTGACTAGTGATACAAGTAAAAATACAGATTTG
AAGTTAACTGGATTAAGGAAACATAAATTTTTCCTTGGGAACCTTATTCCAAAAGGT
GAAAGATACTATGTAAAAGTTGATTTGCAAGAAATGTATAGTTGTGAGGCTTTGATA
GACACAGCTTCAGAAGTGACTTTGATTACTCAAAGTATGTTTGATAAATTGATTGAT
GTCATTGGTAAAAAGAAGAATTTTAGAATTCTTCCTTGTGATCTCCATTTAATAGATT
ATGGAGGCAATTCTCAAAAACCTTGTGGAAAAGTCTGGATAAAAGTAAAAATAGGG
AAAATGAGTGTTACACATCCTGTGTATATCACAAATAAATGTGAAGAAGAGTTTTTA
TTGGGAAATGATTTGCTGAAAAGACTTTCTCCAGTCATGGATTATTTACATGAGATT
TTGTGGTCTCAAGTTAATAGACCAATACCATATGGAAAATATCTTCAAGATCGTGAA
AAGAAGGTAGAAACTTGTTTTATTGAACAAGAAGATTCCAAGTTAAAATTGAATTTT
ATTCGTTCAGCAAATCCTACCTTACTAGTGTTAGGAGATCCAAATAAATTAGTGACT
TGGAAAGACAGAGATAATTTTGTTCCTATATCAAAGGAAAAAATTGTTAAAACTGA
ACTGGAACCTAATGGATTAAATTTCTTTTTATCTCCAACAGAAGCAGATGCTTTTGA
CTGTTTAACACAGCAACAGAAAACTGTGAGTAATGTTGATGTAGCAAACAGCTCAA
GTTTGCAAGAAAATGACAAAGGAAAGTCATTAACACATCCTAGAGAGTTGTGTTGC
ATTGAAAAGGAACAGGGTTTTCTTTACATACCTGTTAAAATAGAAAAGTTACAAGCA
ATGAAAACTTTGTTGTGCATGGAATATGAGAATTCACTAATAAGTGAGACTTTCTTT
AGAAAGTTAAAAAGGCAAGAATCTGATTTACTTTTTAAAAGAAATAAAAAGTGGGT
TACAAATGTACCTTGTGAAGAAAAAGATTTGATTAGTCCTGGGATCTGTGCTTTACA
AATCCAAATAGGAAACAAGAATATGATGCACTACTTTCAAGTAGTTAGGGGGTTAA
GTTATCCAATTTGTATTGGATCAGATTGTCTTGTACGAATGGGATCCTATGTAGATTT
AATTAACCATATGTTATGGTCTCGATTGGAAAATGATCCCAAACCTCTTAAAGGTAC
TATACAGTCTTTGAAATCAGGACAAACTGTGCCACAGGCATGTCAGGTAGTGAGTG
AGTTTACATTTACTATACTGAGTAATGTAATGAATTTTCCTCTCCAAATTAATATAAA
AAGGGGTCAGGAAATATCTGATTCACTGGCTTTCTTTGAACCAAGTAAAGAATTTGT
TAATATGGGTTTAAAAGTTTATGGAACTCCTTGTTTAGATGTGAGTAAAAGGTTCAC
ATATTTGTTGGTTCAAAATGCAACTCCCAAGACTATTGTTTTGGAGAAAGATACTGA
AATTGGAGTATTAGTTGAAAAATCATTTTATGATTTAGAATTGGCCATCCCTGTTATT
GGAAAATTACCAGAAAACATAACTGGGGGGGGGGTGAGTTTGAAGAAAGTGTTAA
AACTTATCCTAAAGGCTTAATTACAATCCAATCAATCTGGCCATATGATAAGGATCA
TACTTATCAATGTAAAATAGAACCTGATTGTTTGATCTTAGAACGAAAAGAGTTGTT
GGAAAATGAGACTGATACAGTTAAAACTGTAAATGCAGTTACTGTTGAAGATTTTTC
AGAAAAAATTGAAGGAAAAATTGAGTTAGGTGAAAACCAAAATTTTCCTGAATTTG
AACAAATGGTAAAACAACAAATTGAAGCTGCTGATGCTTTAGAATCTGATTCTGAA
AGGGAACAATTAAAGAAAATGTTCCTTGATTTTAAAGATCTTTTTGCTAAAGATGCA
TATGATTGTGGAGAAACAGATTTGCATATAACTAAAATACCTACAGATTGCAACAA

AAAACCCACATTTGTACGACAATACAGAATACCATTATCGTCCTATGAATCTGTGCA
AGAAATTTTACATACTTTGATGGAAAGAGAAATTATTAGAGAATGTAATTCTACATA
CAATTCTCCTATTTGGCCTGTTTTGAAACCAAATGGAAAGTGGCGACTTACAATAGA
TTACAGGCAATTAAATAAACAAGTGCCTTTGTCTCATTGGCCTATGACCTATCTGGA
TCAAAATTTGACAAAAATTGCTGGTGCAAAATATTTTACTACATTAGATATAGCTAA
TGGTTTTTGGACTGTAAAAGTAGATCCAAAAGATCAATATAAATTGGCTTTCTCTTTT
GGTAAGAGACAATATACTTGGAATAGATGTCCTTTTGGATATTTAAATTCCCCTGCA
GATTTCAATATTTTCTTACACAAAGCAATGCCAGATGCTGAAGAAAGAGGAAATTTG
ATATATGTGGATGATATATTGATAAAAAGCAACTCAGCTTCAGATCATTTAGATGAG
ATTAGGTATGTTCTTACTCAGCTGAAAAATGCTGGTGCTAAATTGTCATTAACAAAA
GGACAATGGTGTAGAAAGCGGGTAAATTATTTAGGATATGATGTTGGACCAGATGG
ACTTCGTCCACAACAGAACAGGATTAAAGCTATTCAAGCTATCAGAAGTCCTACTAA
TATAACAGAATTACGTAGTTTTCTTGGAATTTGTAATTATTCTCGTCAATTCATTGAG
AATTATGCAGAAATTTCACAACCACTCACAAAATTGTTAAGGAAAAATGAACCATG
GGTATGGGAATCTACACAAGAAAATGCTGTAAGTGAATTAAAGGAAAGGCTTAGTA
AAGTTCCATGCCTTGCCTATCCAGAGAAAGATAAAGAATTTTTCCTAGAAACAGGTT
ATACTGATTCTAGTATCAGTGCAGTATTGTACCAAAAACATGACACAGATAACAGA
GTAGTTGCTTATGCAAGTAAAACTCTTAGTGAGGTTGAAAAGAAATTTTCAGATTGT
GAAAAGGCACTTTTTTCTACAGTTTGGGCTGTACAATATTTTAGAAATTACATACAA
GGAGAAAAATTAATAATTGAAACATGTCATCAGCCCATAGCTTTTCTGGGGAGCCA
GAAAATTAGAGATGGTAATGTTTCTAGTAGTAGAATAGTGGCCTGGACTTTAGCACT
TCAGGGATTACCTTTTGAAGTTAAATATGCTAAAAATAAGAAAAACATTTTGGCACA
AGGAATAGCTGAATTACATGATTGTACTGCTACTGAACAAGAAAAACAAAATTGTG
GTTGTAATTTAGATTTGCCATTACCTTTGGAATCTAATAGACATCTTGTATATAATGA
AAATGTTTGCCAAAATCTTCCAAAAGCATATGTAGATGGTTGTTCTTTTCACCGTGA
AGGACAATTACAAGCAGGAATAGGAATATTTTGGGAAGGAAATTTTCCATGTACTC
CTTTGCAATATAAATTGGGACCCAAGTCTAGTCAATTTTCAGAAATTGCAGCTGTAT
TTGTAGCTATAGAAACTGCAGTGAAAAATAAATTGTCAGATATAGTGATATGTACTG
ATTCAGATTATGCTAGACATAGTTTTTGTAAATTTTTGCCAACATGGAAAGCTAATG
GAATGAAGAATTCTAGAGGAAAAACAGTAAAACATGCTGAACTTATATTGGCTGTT
GATCAATTGGTATTGGAAAATGAAATAACAGTTTACTGGAAAAAGGTAAAAGGCCA
TGCTAAAGACTCTAGTCAAGATAAGTATGGAAATGATTTGGCTGATAAGCTAGCAA
AAGAAGGAGCAATTACAGGTGAACTGTGGGAGTTTAAGCCTATTCCACAAGTAAAT
GTTAATGTGGTTACTAGACAAGAAAGTAAACAAAAGCCAATTAATACTGTGTTACA
GTTGTGCAGTGAAGTACCTAGTTTTGATTTAGTCAAAATGCAGAAGGAGGATGATAC
TTTAGGAAAATTTTATCAGTATATTGAAAATCCTGAAGAAAATGAGATTTCACCTTA
TGACCTTCAAACAAATTTTGAATTAAGAACAATGTTTAATAAAAAACAAAAATTTCA
GTTAAGAAATGGATTGTTAATTCGAATAGGTAAATCTGGAATGATTCAGTGGGTTGT
CCCTTTGTCTTACAGAGGGATAATGTTACAACATGCTCATGATGAAGAATGTGGAGG
TCACAGAGCAGATAGAGCTACTTATGAAACACTTAAACAAGTGGCCTACTGGCCAA
ATATGTACCAAGATGTACAAGATTATACAAAAGGTTGTTTAGTTTGTTGTAGGTTTC
AGTCTTCCTCTCCTAAACATAGAGCTCCATTGCAAACAAAAGGAATTACTATGCCTT
GGTCTGATTTACAGATAGATTGGATAGGACCAGTTACTAGATCTTCCAGAGGAAATA
AGTATCTACTTACAGTAACTTGTTTATTTACAAAATGGATAGAATGTTTACCAGCAC
CTAATGACACTGCTGAAACTACAGCTATGTTATTGATAAATCATGTGTTTAGTTGTTG
GGATTTACCAATGAGGATAGAGTCTGATAGAGGTTCTCATTTTACTGGAAAAGTGAT

AGAAAATGTTTGGAACACTCTTGGAGTTCAAAGGAAATTACATGTGTCACATCACCC
ACAATCCTCAGGACAAGTGGAGCGTGCAAACAAAACTATAGTGAATATCTTAAAGA
AATATATTTCTACTTCTGGTAAAAATTGGGATGTGAAGTTACCATTAGTCTTAATGG
CAATAAAAGCTACTCCTAATAAATCTACAGGTGTCACACCTTTTGAAATGATGACAG
GAAGAGAAATGACTTTACCCTTACATCTTTTGTACAGGACTTATGATGTAAATTTAG
CAAGTGCTACAACTACTTTTGACTATGTGACAAATTTAAAGCAACATCTTCAAAGTA
CTTTTGCTTTTGCTCAAAAGAATTTGGATAAAAGTGCTCAAGGTAGAAAAGCATATT
ATGATAAAAAGACAACTGTAAAAGAGTATGAAATTGGAGACAAAGTATTCTACTTT
AACTACAGTAGAAATAATGCAAAAGTAAAGAAATTTTTACCAAGCTGGAATGGTCC
ACATATTATTACAGATAAAGTCTCATCTGTGGCTTATCAAATTAAACTAAAAAAACC
TACAGGAGATGAAGTTTTCAAGTGGGTACATATAAACCAACTTAAACCCTATTATTC
TCAAAAGAGAACACAAACCAATGAAACAGAAAATACAGAATGAATTGCTATCTAAC
TATTGAGATTTTATTTTGCAGATGGCGTTCCAAACCCTCAATGAGGAGAGAACATTA
ATCCAACCAGTTAGATTGTTAATACAAGATTTGCTATGAGAAATTGACAGTAGCATA
AGTGATTTGATTAAGGGACAAATTCCTACTTATCTTGTTCCTCAAGAAATTGTTTATG
ATGCATTCAAACAAATTTCAAGGAATCCTATTGAACCAATTCAGATACAAGTTGCAT
ATAATATGGGAAGTGCTATTCCTATTTATCTGAATGCTGAAAATTTGGAAATTGCTTT
TGTTTTAAATTTGCCTTATATAATTCCAGATAATGTTTATCAACTAAAATCTGTCTTG
AATGTAGGTAATTGGCATAGAGATGTCTACGTAAAAGTGAATACTCCAGAATTGGT
AGCCTATCAAGAATATAACCCAGGACAGTATTTGATTCCAAATTTAGAAATGTGTAA
AAAGGCAAAAGATATTCATTGGTTATGTCCTGGAAAACCTTTCTTAAAAGATACTAC
TGAAGTTTTATGTGGTTTGACTTCAGAAAATGTTATGGAAAAGTGTAAAATTGGTAT
GACTAAATTGGATGATTTTTCTGAAACAAAAGTTGCTATTGCAAATGCACAGTGGTT
AGTTAGTATTCCTCTTAAAGAAATCACAGTATCATATAGTAAACATGATGTACTGAC
TAAAATGATTTTACCAAGTAATATTTCATTAATTACTGTTCCTAAAGACACTATCGTA
TATGTTGGTGATGTTTCTCTGTATTACTTAGAACAGAATATTCATGAAGCAAACATA
GAAATCCTTGATGTTTTCAAAGGACATGAACTGAACTGGGGAGATGACATCAATTAT
TTGTTGAACTATAAAAACTTCCACACTATAGAAATGAGTGTTAA
> Loki-Lch consensus sequence *pol* gene
TTGACTAGTGATACAAGTAAAAATACAGATTTGAAGTTAACTGGATTAAGGAAACA
TAAATTTTTCCTTGGGAACCTTATTCCAAAAGGTGAAAGATACTATGTAAAAGTTGA
TTTGCAAGAAATGTATAGTTGTGAGGCTTTGATAGACACAGCTTCAGAAGTGACTTT
GATTACTCAAAGTATGTTTGATAAATTGATTGATGTCATTGGTAAAAAGAAGAATTT
TAGAATTCTTCCTTGTGATCTCCATTTAATAGATTATGGAGGCAATTCTCAAAAACCT
TGTGGAAAAGTCTGGATAAAAGTAAAAATAGGGAAAATGAGTGTTACACATCCTGT
GTATATCACAAATAAATGTGAAGAAGAGTTTTTATTGGGAAATGATTTGCTGAAAAG
ACTTTCTCCAGTCATGGATTATTTACATGAGATTTTGTGGTCTCAAGTTAATAGACCA
ATACCATATGGAAAATATCTTCAAGATCGTGAAAAGAAGGTAGAAACTTGTTTTATT
GAACAAGAAGATTCCAAGTTAAAATTGAATTTTATTCGTTCAGCAAATCCTACCTTA
CTAGTGTTAGGAGATCCAAATAAATTAGTGACTTGGAAAGACAGAGATAATTTTGTT
CCTATATCAAAGGAAAAAATTGTTAAAACTGAACTGGAACCTAATGGATTAAATTTC
TTTTTATCTCCAACAGAAGCAGATGCTTTTGACTGTTTAACACAGCAACAGAAAACT
GTGAGTAATGTTGATGTAGCAAACAGCTCAAGTTTGCAAGAAAATGACAAAGGAAA
GTCATTAACACATCCTAGAGAGTTGTGTTGCATTGAAAAGGAACAGGGTTTTCTTTA
CATACCTGTTAAAATAGAAAAGTTACAAGCAATGAAAACTTTGTTGTGCATGGAATA
TGAGAATTCACTAATAAGTGAGACTTTCTTTAGAAAGTTAAAAAGGCAAGAATCTG

ATTTACTTTTTAAAAGAAATAAAAAGTGGGTTACAAATGTACCTTGTGAAGAAAAA
GATTTGATTAGTCCTGGGATCTGTGCTTTACAAATCCAAATAGGAAACAAGAATATG
ATGCACTACTTTCAAGTAGTTAGGGGGTTAAGTTATCCAATTTGTATTGGATCAGAT
TGTCTTGTACGAATGGGATCCTATGTAGATTTAATTAACCATATGTTATGGTCTCGAT
TGGAAAATGATCCCAAACCTCTTAAAGGTACTATACAGTCTTTGAAATCAGGACAAA
CTGTGCCACAGGCATGTCAGGTAGTGAGTGAGTTTACATTTACTATACTGAGTAATG
TAATGAATTTTCCTCTCCAAATTAATATAAAAAGGGGTCAGGAAATATCTGATTCAC
TGGCTTTCTTTGAACCAAGTAAAGAATTTGTTAATATGGGTTTAAAAGTTTATGGAA
CTCCTTGTTTAGATGTGAGTAAAAGGTTCACATATTTGTTGGTTCAAAATGCAACTCC
CAAGACTATTGTTTTGGAGAAAGATACTGAAATTGGAGTATTAGTTGAAAAATCATT
TTATGATTTAGAATTGGCCATCCCTGTTATTGGAAAATTACCAGAAAACATAACTGG
GGGGGGGGTGAGTTTGAAGAAAGTGTTAAAACTTATCCTAAAGGCTTAATTACAA
TCCAATCAATCTGGCCATATGATAAGGATCATACTTATCAATGTAAAATAGAACCTG
ATTGTTTGATCTTAGAACGAAAAGAGTTGTTGGAAAATGAGACTGATACAGTTAAA
ACTGTAAATGCAGTTACTGTTGAAGATTTTTCAGAAAAAATTGAAGGAAAAATTGA
GTTAGGTGAAAACCAAAATTTTCCTGAATTTGAACAAATGGTAAAACAACAAATTG
AAGCTGCTGATGCTTTAGAATCTGATTCTGAAAGGGAACAATTAAAGAAAATGTTCC
TTGATTTTAAAGATCTTTTTGCTAAAGATGCATATGATTGTGGAGAAACAGATTTGC
ATATAACTAAAATACCTACAGATTGCAACAAAAAACCCACATTTGTACGACAATAC
AGAATACCATTATCGTCCTATGAATCTGTGCAAGAAATTTTACATACTTTGATGGAA
AGAGAAATTATTAGAGAATGTAATTCTACATACAATTCTCCTATTTGGCCTGTTTTGA
AACCAAATGGAAAGTGGCGACTTACAATAGATTACAGGCAATTAAATAAACAAGTG
CCTTTGTCTCATTGGCCTATGACCTATCTGGATCAAAATTTGACAAAAATTGCTGGTG
CAAAATATTTTACTACATTAGATATAGCTAATGGTTTTTGGACTGTAAAAGTAGATC
CAAAAGATCAATATAAATTGGCTTTCTCTTTTGGTAAGAGACAATATACTTGGAATA
GATGTCCTTTTGGATATTTAAATTCCCCTGCAGATTTCAATATTTTCTTACACAAAGC
AATGCCAGATGCTGAAGAAAGAGGAAATTTGATATATGTGGATGATATATTGATAA
AAAGCAACTCAGCTTCAGATCATTTAGATGAGATTAGGTATGTTCTTACTCAGCTGA
AAAATGCTGGTGCTAAATTGTCATTAACAAAAGGACAATGGTGTAGAAAGCGGGTA
AATTATTTAGGATATGATGTTGGACCAGATGGACTTCGTCCACAACAGAACAGGATT
AAAGCTATTCAAGCTATCAGAAGTCCTACTAATATAACAGAATTACGTAGTTTTCTT
GGAATTTGTAATTATTCTCGTCAATTCATTGAGAATTATGCAGAAATTTCACAACCA
CTCACAAAATTGTTAAGGAAAAATGAACCATGGGTATGGGAATCTACACAAGAAAA
TGCTGTAAGTGAATTAAAGGAAAGGCTTAGTAAAGTTCCATGCCTTGCCTATCCAGA
GAAAGATAAAGAATTTTTCCTAGAAACAGGTTATACTGATTCTAGTATCAGTGCAGT
ATTGTACCAAAAACATGACACAGATAACAGAGTAGTTGCTTATGCAAGTAAAACTC
TTAGTGAGGTTGAAAAGAAATTTTCAGATTGTGAAAAGGCACTTTTTTCTACAGTTT
GGGCTGTACAATATTTTAGAAATTACATACAAGGAGAAAAATTAATAATTGAAACA
TGTCATCAGCCCATAGCTTTTCTGGGGAGCCAGAAAATTAGAGATGGTAATGTTTCT
AGTAGTAGAATAGTGGCCTGGACTTTAGCACTTCAGGGATTACCTTTTGAAGTTAAA
TATGCTAAAAATAAGAAAAACATTTTGGCACAAGGAATAGCTGAATTACATGATTG
TACTGCTACTGAACAAGAAAAACAAAATTGTGGTTGTAATTTAGATTTGCCATTACC
TTTGGAATCTAATAGACATCTTGTATATAATGAAAATGTTTGCCAAAATCTTCCAAA
AGCATATGTAGATGGTTGTTCTTTTCACCGTGAAGGACAATTACAAGCAGGAATAGG
AATATTTTGGGAAGGAAATTTTCCATGTACTCCTTTGCAATATAAATTGGGACCCAA
GTCTAGTCAATTTTCAGAAATTGCAGCTGTATTTGTAGCTATAGAAACTGCAGTGAA

AAATAAATTGTCAGATATAGTGATATGTACTGATTCAGATTATGCTAGACATAGTTT
TTGTAAATTTTTGCCAACATGGAAAGCTAATGGAATGAAGAATTCTAGAGGAAAAA
CAGTAAAACATGCTGAACTTATATTGGCTGTTGATCAATTGGTATTGGAAAATGAAA
TAACAGTTTACTGGAAAAAGGTAAAAGGCCATGCTAAAGACTCTAGTCAAGATAAG
TATGGAAATGATTTGGCTGATAAGCTAGCAAAAGAAGGAGCAATTACAGGTGAACT
GTGGGAGTTTAAGCCTATTCCACAAGTAAATGTTAATGTGGTTACTAGACAAGAAAG
TAAACAAAAGCCAATTAATACTGTGTTACAGTTGTGCAGTGAAGTACCTAGTTTTGA
TTTAGTCAAAATGCAGAAGGAGGATGATACTTTAGGAAAATTTTATCAGTATATTGA
AAATCCTGAAGAAAATGAGATTTCACCTTATGACCTTCAAACAAATTTTGAATTAAG
AACAATGTTTAATAAAAAACAAAAATTTCAGTTAAGAAATGGATTGTTAATTCGAAT
AGGTAAATCTGGAATGATTCAGTGGGTTGTCCCTTTGTCTTACAGAGGGATAATGTT
ACAACATGCTCATGATGAAGAATGTGGAGGTCACAGAGCAGATAGAGCTACTTATG
AAACACTTAAACAAGTGGCCTACTGGCCAAATATGTACCAAGATGTACAAGATTAT
ACAAAAGGTTGTTTAGTTTGTTGTAGGTTTCAGTCTTCCTCTCCTAAACATAGAGCTC
CATTGCAAACAAAAGGAATTACTATGCCTTGGTCTGATTTACAGATAGATTGGATAG
GACCAGTTACTAGATCTTCCAGAGGAAATAAGTATCTACTTACAGTAACTTGTTTAT
TTACAAAATGGATAGAATGTTTACCAGCACCTAATGACACTGCTGAAACTACAGCTA
TGTTATTGATAAATCATGTGTTTAGTTGTTGGGATTTACCAATGAGGATAGAGTCTG
ATAGAGGTTCTCATTTTACTGGAAAAGTGATAGAAAATGTTTGGAACACTCTTGGAG
TTCAAAGGAAATTACATGTGTCACATCACCCACAATCCTCAGGACAAGTGGAGCGT
GCAAACAAAACTATAGTGAATATCTTAAAGAAATATATTTCTACTTCTGGTAAAAAT
TGGGATGTGAAGTTACCATTAGTCTTAATGGCAATAAAAGCTACTCCTAATAAATCT
ACAGGTGTCACACCTTTTGAAATGATGACAGGAAGAGAAATGACTTTACCCTTACAT
CTTTTGTACAGGACTTATGATGTAAATTTAGCAAGTGCTACAACTACTTTTGACTATG
TGACAAATTTAAAGCAACATCTTCAAAGTACTTTTGCTTTTGCTCAAAAGAATTTGG
ATAAAAGTGCTCAAGGTAGAAAAGCATATTATGATAAAAAGACAACTGTAAAAGAG
TATGAAATTGGAGACAAAGTATTCTACTTTAACTACAGTAGAAATAATGCAAAAGT
AAAGAAATTTTTACCAAGCTGGAATGGTCCACATATTATTACAGATAAAGTCTCATC
TGTGGCTTATCAAATTAAACTAAAAAAACCTACAGGAGATGAAGTTTTCAAGTGGGT
ACATATAAACCAACTTAAACCCTATTATTCTCAAAAGAGAACACAAACCAATGAAA
CAGAAAATACAGAATGA
> Loki-Lch consensus sequence Pol protein
LTSDTSKNTDLKLTGLRKHKFFLGNLIPKGERYYVKVDLQEMYSCEALIDTASEVTLITQ
SMFDKLIDVIGKKKNFRILPCDLHLIDYGGNSQKPCGKVWIKVKIGKMSVTHPVYITNKC
EEEFLLGNDLLKRLSPVMDYLHEILWSQVNRPIPYGKYLQDREKKVETCFIEQEDSKLKL
NFIRSANPTLLVLGDPNKLVTWKDRDNFVPISKEKIVKTELEPNGLNFFLSPTEADAFDC
LTQQQKTVSNVDVANSSSLQENDKGKSLTHPRELCCIEKEQGFLYIPVKIEKLQAMKTLL
CMEYENSLISETFFRKLKRQESDLLFKRNKKWVTNVPCEEKDLISPGICALQIQIGNKNM
MHYFQVVRGLSYPICIGSDCLVRMGSYVDLINHMLWSRLENDPKPLKGTIQSLKSGQTV
PQACQVVSEFTFTILSNVMNFPLQINIKRGQEISDSLAFFEPSKEFVNMGLKVYGTPCLDV
SKRFTYLLVQNATPKTIVLEKDTEIGVLVEKSFYDLELAIPVIGKLPENITGGGGEFEESV
KTYPKGLITIQSIWPYDKDHTYQCKIEPDCLILERKELLENETDTVKTVNAVTVEDFSEKI
EGKIELGENQNFPEFEQMVKQQIEAADALESDSEREQLKKMFLDFKDLFAKDAYDCGET
DLHITKIPTDCNKKPTFVRQYRIPLSSYESVQEILHTLMEREIIRECNSTYNSPIWPVLKPN
GKWRLTIDYRQLNKQVPLSHWPMTYLDQNLTKIAGAKYFTTLDIANGFWTVKVDPKD
QYKLAFSFGKRQYTWNRCPFGYLNSPADFNIFLHKAMPDAEERGNLIYVDDILIKSNSAS

DHLDEIRYVLTQLKNAGAKLSLTKGQWCRKRVNYLGYDVGPDGLRPQQNRIKAIQAIR
SPTNITELRSFLGICNYSRQFIENYAEISQPLTKLLRKNEPWVWESTQENAVSELKERLSK
VPCLAYPEKDKEFFLETGYTDSSISAVLYQKHDTDNRVVAYASKTLSEVEKKFSDCEKA
LFSTVWAVQYFRNYIQGEKLIIETCHQPIAFLGSQKIRDGNVSSSRIVAWTLALQGLPFEV
KYAKNKKNILAQGIAELHDCTATEQEKQNCGCNLDLPLPLESNRHLVYNENVCQNLPK
AYVDGCSFHREGQLQAGIGIFWEGNFPCTPLQYKLGPKSSQFSEIAAVFVAIETAVKNKL
SDIVICTDSDYARHSFCKFLPTWKANGMKNSRGKTVKHAELILAVDQLVLENEITVYWK
KVKGHAKDSSQDKYGNDLADKLAKEGAITGELWEFKPIPQVNVNVVTRQESKQKPINT
VLQLCSEVPSFDLVKMQKEDDTLGKFYQYIENPEENEISPYDLQTNFELRTMFNKKQKF
QLRNGLLIRIGKSGMIQWVVPLSYRGIMLQHAHDEECGGHRADRATYETLKQVAYWPN
MYQDVQDYTKGCLVCCRFQSSSPKHRAPLQTKGITMPWSDLQIDWIGPVTRSSRGNKY
LLTVTCLFTKWIECLPAPNDTAETTAMLLINHVFSCWDLPMRIESDRGSHFTGKVIENVW
NTLGVQRKLHVSHHPQSSGQVERANKTIVNILKKYISTSGKNWDVKLPLVLMAIKATPN
KSTGVTPFEMMTGREMTLPLHLLYRTYDVNLASATTTFDYVTNLKQHLQSTFAFAQKN
LDKSAQGRKAYYDKKTTVKEYEIGDKVFYFNYSRNNAKVKKFLPSWNGPHIITDKVSS
VAYQIKLKKPTGDEVFKWVHINQLKPYYSQKRTQTNETENTE*
> Loki-Lch consensus sequence *env* gene
TTGATTAAGGGACAAATTCCTACTTATCTTGTTCCTCAAGAAATTGTTTATGATGCAT
TCAAACAAATTTCAAGGAATCCTATTGAACCAATTCAGATACAAGTTGCATATAATA
TGGGAAGTGCTATTCCTATTTATCTGAATGCTGAAAATTTGGAAATTGCTTTTGTTTT
AAATTTGCCTTATATAATTCCAGATAATGTTTATCAACTAAAATCTGTCTTGAATGTA
GGTAATTGGCATAGAGATGTCTACGTAAAAGTGAATACTCCAGAATTGGTAGCCTAT
CAAGAATATAACCCAGGACAGTATTTGATTCCAAATTTAGAAATGTGTAAAAAGGC
AAAAGATATTCATTGGTTATGTCCTGGAAAACCTTTCTTAAAAGATACTACTGAAGT
TTTATGTGGTTTGACTTCAGAAAATGTTATGGAAAAGTGTAAAATTGGTATGACTAA
ATTGGATGATTTTTCTGAAACAAAAGTTGCTATTGCAAATGCACAGTGGTTAGTTAG
TATTCCTCTTAAAGAAATCACAGTATCATATAGTAAACATGATGTACTGACTAAAAT
GATTTTACCAAGTAATATTTCATTAATTACTGTTCCTAAAGACACTATCGTATATGTT
GGTGATGTTTCTCTGTATTACTTAGAACAGAATATTCATGAAGCAAACATAGAAATC
CTTGATGTTTTCAAAGGACATGAACTGAACTGGGGAGATGACATCAATTATTTGTTG
AACTATAAAAACTTCCACACTATAGAAATGAGTGTTAA
> Loki-Lch consensus sequence Env protein
LIKGQIPTYLVPQEIVYDAFKQISRNPIEPIQIQVAYNMGSAIPIYLNAENLEIAFVLNLPYII
PDNVYQLKSVLNVGNWHRDVYVKVNTPELVAYQEYNPGQYLIPNLEMCKKAKDIHWL
CPGKPFLKDTTEVLCGLTSENVMEKCKIGMTKLDDFSETKVAIANAQWLVSIPLKEITVS
YSKHDVLTKMILPSNISLITVPKDTIVYVGDVSLYYLEQNIHEANIEILDVFKGHELNWGD
DINYLLNYKNFHTIEMSVX
>Loki-Lro consensus sequence
TGGTACTAGTTAATCTCTAAAGGTGTATTTTTCAATGTATGAATGTATCTCTGGTTTA
ATATCTCCTAGATGTTTCATTCTTGGTATCAAAACAATCTGCACTTTATTTTCTGAAC
AATGGTGTACACCATGTGATTAGAAAATGCATCATACTGTTTGTACTATATTGTGGG
TAAAACTGCATCAATTCCCCAGGCCAGCAATAAACATGCAAATGGGCCATAAATGG
AGACAAAGAGAAGTTCTCCCGCTCAGCAATACAAACTTTGTATTGGCCAATTCACCC
AAAGGTGGGTGTGGTAGGGAAACTATCAAAAGACCCTAACACCAACCCATCCAATG
CTCAAACCCTTCGAGACACTCTCCAAGACTCTTCCCTTGAGTCAATTTTCCGGTAGTT
TATCTTCAAGTAACTTTGTCGACTCGCTGTGGATTTTCCTATCTTCGAACGAACTAGA

GAAAACCTCCGGAACACCTAAAGGAATTCATCGAGAACTTCCATTGTCTACCGCATC
AGGACTCTTGTTCAGCAACAAAGCCAATGCAAGTAACCTTTACAACTATTTAGATGC
GCGTTTTAAGTTGAGCCCCTGTTTAAGGTGATTTTGCTTCTCTGATGATTCTCATGTG
GTCGCTGGTTAACTGATGCTAAATACTGCCATTCTTGCCTCTCTGTCTGCTCTCTCTC
TCTCTCTCTCCTTTCTCACTCTCTTTCTCTATGTATATGTACACGTGTTTTGCTTAGTT
AGTTGTATGTTAACTGTAGTTCCATTTATTAAATTGCATATATTCACAATTGATTGTC
TCTGTGTGCTGCTCACAATTTGAAGTCACTGAATATTTGATCTTGCTACATGCTACAT
TTTGAATTTACTACGAGTTAAGTTACTGCTAGTACTAACTGAAGTAGGAAATGTATT
CTTTCTTGGCCAGAAAGTAACCTTTCCTAGAATTGATGAATGATTATACTGTCTATTC
GGTGGACGAACAGATCAAATAATTGTAACGTTAATTCTGCTACAGTTGATTCTAAGA
TCTGATTAAATATTCATAATTAATTATAACCAGTTATAATTAATCATAAATATTTCCC
TTTTGAGCTAATTTATTACAGATTTATGGTGGAGAACGAGTGGCAATATTCAAACTT
TGTTTTGTGAGCAAACCTGGAATTAATCTAATTGTTAGATATATGTTTTTTATATTAC
GCGCGCTTAGAGTGAGACAAGTCGCAACGCGAACTCACTTTTCTGATTGGACGGCTT
AAAGCCACCCCCAGAACTTTTAACAGCATCTATTTGTCTAACATTTAGTCACTGCTA
ATTGTTGATATAAACTGCAGTTCATAATATTAAAATTGTTCTGTAAAGACGAAGAAA
TCTCTTTGCAGTACATTTTAATATTATTAGAGTCGCCTGTTGGAACGAAGAAATCTTC
CTTCACCGCGATAAAGATTGATTTGAGCGATGTTCCTCTCTTCATCTTTTTAAAAGGC
CTTATTTGATGATATCGCTAGATTGAAGTCATTACATGTTAATTTCATTCAACTAGAC
GATCTCACTCCCTTAGGTTTAATCATTAGAGTAGCTTACTAATGTGTGCCTAACGTAC
AGATCTGATCACAAGAACACTGCGAGCCAGTAACTGATTGGAATGTAAATGGGAGT
TTTTACAATCTTGATCAAGTGATAAATTTTGTAACAGCGAAGAAACCCTGTTGCATA
TAATTGTTATTGGGTAACAACAGCAAAGAAACCCTGTTGCTATAATTGTTATTGGGT
AACAACAGCGAAGAAACCCTGTTGCTTATAATTTTATTGTGTAACAACAGCGAAGA
AACCCTGTTGCTCATAATTTTTAGTGTGTACTGGCAGCGAAGAAATCCTGCTACACA
TAATTGTGCAGTTAAGTTGCACAAACTCTAAGAATCTGTGGGTTCTAAATCACATTG
TTCTTGTGATAGCTGCAATAAATCGGCCGTTAAGTCCTAACGCGTTCAGATTTATGC
ACGATTGGAGTTAAGGCACAGTACCTATGACTAACACTAGAGGGCAGGCTGTTTAA
GTTTCTAACACCTCTTAAACTCCACCTCAACCCTTTGGTTTCACGGTGCGTGGGCTCA
CCCACTCCATCTTGTGGGCATTCTCAGTTAGGTTAACTTCTCACTCCCCAAGTCCTTG
CATTTCATTTTGTTTGTTTATTTTCCATCTATTTGTTTTATCATTTATTATTTTCTCACC
ATCTTGGTTACCTCCAGTTTTATTTATCTACTTTATTATTGTTTTGTCCTTACCTTTTTA
TTTTATTTCATTATCAATTTAGTTTTATTTTTATTTTTTATTTCATTATATTTTATTTTTT
TTATTTACGTTTTCTCTCCCTCTTCTCTCTCATTTCACTACCCACTCATTTAAATTTAA
TTTAAACAGAGTTTGAATTTAAATTTAAAGTTAATTTCTGTATCTAACTACTTGTGTT
TGCCTATAGGTAGCACTCCCCCTAGTGCTTAGTTGGTGAAGTGACTGGACCGCGTCG
ATCCATCACTAAGCCTAGCCATAGAATACAGATCCATCCAGTTGGTGAAACCCCAA
ACTAGGTTTTTCCTCAATCTTTCAGGAACTTAGCCTGCTCCTGCTGTCCCAACAAGGC
CAAAACCCCGAATGTTATGACCCAGCAGTGCCCCCTACGCTGTGGGCCCTGTCATAA
CCCCCGACAGCAGTGAAAGGCCTAAACCTGCTCTCTCTTCTATCTCTATCTGTGTTTG
ATTTGTTGGTTTGTTTTAATTTTTGTTCTGTTTTGTTTGCATAGAGACCAACCTTTGAG
AATCATCAGGGAAGGCTGAGTAATAGATTGACGACCCAAGCTGCCAAAGTCTTTAG
ACGATCACGCTGTCAACAATCTTTAGTTAGACCCTTCCGGTGCCCCACAACAAGTAG
ACACACCTGACCATTCTCGGCCGAGATAACAGGTGCCCGGTCTACACTTCGCTAACT
GTTTGTACCCAAACCGTACCAAGCGAATATAGTTAATAGTATTGCAGCCGCTAACAC
TGTAAGGACTTGCATTGGCCCTTTTTGCTGTCTTTCTCTCTTCTCAGTGAGCCACCAT

GTCTCGTTCATCCAGCCCTGATGCCCACTGGGATCGACTAGAGGCCTGGCTGAGTGC
CATGACGGGTACCCTTCTGCCTGACGCCGCTGGTGGCCTGCAGCACCTGAGCCGGGA
ACAGCTTGACGACAACCTGAGTGGCTTGATGTCCCATGATCCCACGCAGGACTACA
ACCACAAAGACCTCGCCAAGATCATCGGGTCACTTGCCCACAACCTCCTCGCTCAGG
CAAGGCTGAGCGAAGCAAGCATCTCCCGCCTGGAACAGGAAGCAGCCGCGCTAAAG
CTCCGAGCTGAGGAGGCTCAGAGGGACCTGGCACTTGCCCAAAGTCAACTGGACCA
GCTGGAGACGGAGACACCGCAAAGTGACTCTGAAAGGGAAATGCACACCACAGAA
TTTCCCATAACCAGTAAACTAGTGCCTCCTAGTCAAGAACTCTCACTCCAGAAGGGG
GGTGAAAGCCCCCTTCCCAAGACATCTCCTGCCAGCATCCCAGAACACCCCACACCA
GCACGGGGGTGGGAGCCTTTCAAGGGAACACTGCCTACCAGTCACGGTGTAGCCCC
CAAAGAACTAGACAAGCTGGCCAGAAACATTCCCATGTTTACCCCGAACCCTGCAG
GAGGCCATGACGTGCACGCCTACCTGAAAGACATTAATTTCCACTTGCAAGGTGTGG
CCAATGTAACCACACGAGACAGACTTTACCTGCTCAGAATCACCTCCAGCCGTGAG
GTGCGAAGCTTCCTAGATCGACAGCCCGAAACGGTCAAAGCGGATTACCAGCAGCT
ACAACAGGCTTTGATCACGGAATTTTCTGACCCAGAGTCAGAACAAGGACTGATTAC
TGCCATGGACCTCAAACAGGCCCGACAAGAAACCTCGCAAGCCTACTACACCAGGC
TCCGACAGGCCTACTTCGGGGCACGGAACGAGCCAGGAATGGAAGAGGACTTTAAC
TTCAAGACTCTGTTCCTCCGGAACATGCACCCTACAGTGAGTCACCACTTGGGGGTC
TTGGCTTGCCCACGAAGCATGTCCATTCAACAGCTGCGAGACTTAGCCCACAAGGCC
TATGTCAAACAAAGAGCTGCTTCGGAAAAGACTGTGAAAACCCCCACAATTTGCCCT
GTCTCTGATCAGAGTCCAGAACTGGCTCTAGAGGGCGCCCAACAACACTACAGTGA
CAAACCTTTCAACAAGATGTCAAGGCCATTCCCTGCCAGCAGAGGGCAGTATGGCC
ATGGCGGTGCCCGTTCCAAGCATCGAACAGAGCGCCCTGAGACATCCTGGGACAAG
CCACAGCCAGCCTTCCACCCGAAGGGAGGAACCACGTGGGAAGGAAAAAGACGGC
ACAAGGGCAGTCGACCTTCCCCAGCTCAAGCATCAAGCTCGAACTGCCAGCTACAG
AATCCCTTTAAACATGCCCTAGCCCAGCCAAAGTCTGAACCATCAGGAGAGCAGGA
CAGAGAAGCTACAACTGAGTCAGCAGAGATCCTGAGGCTTTTAAGGGGACTTCTCC
GAGAAAAGCACAACAAGGGAGACAAGAGGGACAAGCCAGATCAAGCATGACTAAG
CATGGGATCTCACCTCAACTCCCACTCAGAGAACTCTCAAACTGAACCTCACAACCA
GATCACTGCTAATCCACTGTCAACCACAGAACCAGATACTTCCTCACCAAGTGATGT
CCGCACTCAACCGTCTCAGCAGACTGTTGAACTGCCAGAGCCAACCAACATCCAAA
GCCACACTTCCAGCCAGGATTCCAACGTGCCTGAGAGTGCAGTTTTGGTAGTCTGCC
TCCCCGACGAGCCACATGAGACCAGCCCTGACTGCCTGGCAGTCAACACACAAGTC
CCAGTGCCAAGACTCCTGGGAAACCTGATCGAACGGGGGGTGGCTAAAAAACTTTA
CCTGGTCATCACATTGGAAAAGGAGGTGAGACTGGAAGCCCTCGTAGACACTGGAG
CCGACCTGACTTTAATGTCCAGCCAGCTTTTCGACAGACTTCAGACTGAAGCCAAGA
GACAGAATCGAACTCTGAAATATCAGCAGTGTGTGCTGAATGTGCAGTCATACAGC
CAGAATGAGGTCCAACTTGAACAAGTAGCTCCCATTCACCTGACTATCGGCCCTATG
AGTTTGGTGCACCCTGTGTACATCGCTCCAGTTGACACCTACCCTCTTCTCATTGGCA
AAGACCTGCTAGACCGCTTTGAGCCCCTGTTAGACTTTAAACAGCTGAAAATCTGGG
CTCAAGTGCGAGAGCCTCTGCCTCCTCAGTCTCCAAGGTCACCCGAGACAGACTGCC
AGGCCACCAAGGTCACAGAAAATCCCACCGCCAGCCGTGGGAACGTCTCAACCCAA
GAGCAAGGTTCAAGGTCGATGCTCTGTGCTTTCCAGCTAGACAAAGACTTTGACTCA
GATGGCCCCCAAATTCTGGCAGGCCTACAGCTGAATGATGTCACAACAACAGACCT
CATACTGGCACTGTGGGCAGACAACTCGGCTATCAGCCTCACACTTTTCAACGCTCT
CAAGAGACACACAACAAACATCTCCTTTGCCAACAAGCGCACACGTTTCGCTTTGAA

CCCCGGCCTGTCAACCCTGGTGACTGCCAAGGTCATCTGTGCCCTGAACTTGAATTG
GAATGACAGGTGCTTAACGCATTACTTCCTGGTCCTGCCAGACCTACCGCATGATGT
TTACATCGGAGCGGACCTTCTGATTCGCCTCAAGGCTCACGTGGACCTGCTAAATGG
AGTAATATGGGCCCCAATGACCTCACAGGCCTCTGCTGTTCCTGTCAACCCCGAGAA
TCTCCTGTCAGGCCAAACCATTCCCGAGGCCTGCACACTAATCAGCGAACAAGAGA
CTGTAATACCAGCTTACACCAAAGGGGTCGCTGTTCGACTCAACCTGCGATGTGGCC
AAACCCTCAGCCACAGCCTAGGCTTCTTCCAGCCTTCACCCGAGTGTGTGGAACTGG
GCCTCACTCTAGAAGCCACCCCTCTCGTTGAAGTGACATCTCGGGCCTTGTACATCC
TGTTCAACAACTGCATGGCAATGGATGTCAGAGTGCCTAAAGCCTACCGTCTGGGTT
GGCTAGTGAGTTATGACTTCCAGGACTTTGAACTTGTTTTACCAGTAATTGGACCTCT
CCCTGTTTCCATGATTCCAGAGGGGAGCTCAAGCCACACAGTGTCGACAGTACCTTT
CAAGATCATCACCATCCTCTCAGTCGTGCCAGTGGTCAAGGAGCAGGTGTGCAGAA
CAGAGCTAACGGTGGACCAACACCTTGCTGTGTACACAGTCTCTAGCCATCCCACCT
GCGAGACTGACAACGCTACTGCACCCCCGACTGTCCACCAAACAGATGATGTAAAG
ACAGGAGAGCCGTACCCTGGTTTTGAGACACAAGTTCAGCAAATTCTGAGGGATGC
TGATGCCCTACGGGAGGATGCAGATCGCCACAAACTTCGTCAAGTCCTGTACAAATA
CAAAGACTCATTTGCCAAGGACTCTTTAGACTGTGGCCTCACTGACCTGCATGTGGT
GCGAATCCCAACGCACCCAAAAGCTCCACCCACCTTTGTCAAGCAGTACAAGATTCC
TCTCGAATCGTATGAACCTGTGCAGGACATCATCACCTCGATGCTTGAGAAGGGTGT
CATCCGTCCATGTAACAGCACCTACTCTGCCCCCATCTGGCCTGTCCTCAAACCAAA
CGGCAAGTGGCGGCCAACCATCGACTATAGGAAGCTGAATCAACAGGTGCCATTAT
CGCGGTGGCCCATGACACAGCTGGAACAGGAGATACCAAGAATCAGAGGAGCTAC
AATCTTCTCCACGCTTGATGTGGCCTCAGGATTTTGGACTATACCAGTGCATCCAGA
TGATCAACACAAGTTAGCCTTCCGTTTTGACAATCGTCAATACACGTTCACAAGGTG
CCCATTCGGCTATGCCAACTCACCCGCCGAATTTAACATCTTCCTCAACAAGGCGTG
CCCAGATGCCAGGGTAAGAGGAAACCTTGTTTTTGTGGATGATATCCTCCTGAAACA
CACAACGGTGGCTGATCATCTGAAGGAGATCGACCACATCCTGAACCAACTGACTA
CCGCTGGTGCCAAGATTGCCCTTCACAAAGGACAATGGTGCAAAATCAAAGTCAAC
TATGTGGGCCTGCTCATCGGACCACAGGGCATCGAGCCACAAACCAACCGTATCCA
GGCCATTCAAAACATCAAGCCTCCTGCCAACGTCTCAGAACTTCGCAGTTTCCTAGG
TGTGTGCAACTATTCACGTCAGTTCATTGAGAACTACTCCGACATTGCACGCCCTCTC
ACTGCACTCCTGAAAAAAGACTGCCCATTCATCTGGACAGACACTCAGCAAGATGC
CATGAATGAACTGAAGCAGCATTTGTGCACTGCACCATGCCTGGCTTACCCGGATCC
ACAGAAGGAGTTCTATTTGGATGCTGGATTCTCCAGCCACTGCCTTAGTGCCGGTCT
ATACCAACGGCATGACCAGGACAGACGCGTGGTCGCGTATGCCAGCAAGACCCTGC
TTCCACCAGAGTGCAAATACTCAGACTGTGAGAAAGCCCTGCTCTGTACAGTATGGG
CAATCCAGAGATTCTCCAACTACATCGGAGCACAGAAGGTGATCATAGAGACCTGC
CACCAGCCTGTCACCTTTCTCAACAGCCAACGTATCCGCGACGGCGTTGTCACAAAC
GCCCGCATAGCCACATGGCTAATGGCCCTCCAGGGGCGTAATGTGGAAGCACGCTA
TGCTCAAAACTACAAGTCGGCGCTGGGTAACGGCCTGGCAGCTTGCCAAAGCTGCT
CTGATGACACACCAACGGTGATCGTACAAGCTGAAACGCCCCCAGACCTGCCGCTT
ACCTGCCACAGGTATTTCGAAGAGAGCGTGTGCCAAGGCATGCCCACAGCCTACGT
CGATGGGTGTTCGTACAACCACGAAGGCCTCCTAAAGGCAGGTGCAGGCGTTCTCTG
GCTGAACGATCGACCGTGCCCACCACAACACTTCAAGTTGGGCCCCCAGTCATCACA
ATATGCTGAGGTAGCAGCTATTCTCATAACCCTGCAAATAGCATCAGCTCATAACAT
CAGAGACCTGCTTATTTGCACAGACTCTAACTATGCCAGACTTAGCTTCACATGCCA

TCTTCCAAACTGGAAACAAAATGGCTTCAAGACAGCAAACAACAAACCGGTGAAAC
ACCAACACCTGTTCCAGATCTGCGATGACCTCACAAGTAAACACGACATGATCATCT
ACTGGAAAAAGGTGAAAGGTCACTCAAAGCAACCTGGCACAGACAAAGACCTCAAT
GACCAAACAGACGCCCTTGCCAAAGCAGGGGCACTACACGGCGAGTTATGGACTCC
ACCAACCCACCCACCCACCCCTAACGTGGCAGCCATAACCCGCAGCAAACGAACTG
TACCCACCACAGTACCCACGTCACAGCCACTAATGCTGGCTCCTCAAATCTCCACGA
ATGACATTGCTGACATGCAAGCTACTGACTCCTCCATACAAACTATTCTTCACCACC
TCTCTGACCCCTCCAACCACCCTATCACACCTGCTCAGCTCGCTAATGCCCCAAGCC
TAAAACGGCTACACGACCTCAAACCCATGCTGCGTGTGATGGACGACATTCTTTGGT
ATGTCCCTGACGACAACGCTGCACCAAGGCTTGTGGTGCCGCAGGGCCAAAGGGGG
GTCTGGTTGATGAATGCACACGACGCAGCATGTGCGGGACATCATGGCACAAGAGC
CACGTACGAGACACTTAAACAAGTGGCATACTGGCCCGGCATGCAGCAAGATGTGG
CAGAGTATGTGAAAGGATGCCTTGTTTGCTGTCAATTCCAACCGGCAAACCCAAACC
ACAGAGCTCCACTGCAGCGCAGAGGAGTCACATTTCCTTGGTCTGACCTGCAAATTG
ACTGGGTAGGGCCCCTACCCAGATCAACAAGGGGCAACAAGTATTTCCTCACAGTT
GTATGCCAGTTCACCAAGTGGGTGGAGTGCCTTCCAGCACCCAATGACACGGCCCA
AACTACCGCATACCTCCTGATGAATCACATCTTCTCCAGGTTCGGGCTACCGCTGAA
AGTCAACTCAGACAGAGGCACCCACTTCACAGCTGAGATCATGCAGCAGATTTGGA
AACTTCTGGGAATACATGCTCAACTGCACATCAGCCATCACCCAATTTCTTCGGGTC
AAGTGGAACGATCTAATCAGACCGTGGTCAGCATGCTAAGGAAATATGTTTCTGCCA
ACCAGAAAGACTGGGATGTGAAACTTCCCTTAGTGCTCATGGCCATCAGAGCAACTC
CAAATGAATCAACAGGTGTGTCGCCCTTTGAACTAATGACCGGCCGGCTAATGACAT
TGCCACTGCACCTGCTGTACCAGCCAGGAGACTCAAACCTTGTCACCGCTTACACCA
CACATCAGTACCTTGATGAACTACGCCGTCATTTGAGATCCACATTCGCTTTCGCCC
AACAACGACTCCAAAAGAGTGCAGAAGGCCAAAAAGCATACTACGACCGCAAGGC
GTCACACGCAGAACTGGATGTGGGTGACAAGGTTTGGTACTACCAGTACGCACAGC
CACCTCAGACAACCTCTCATCGGCTGTCAAAGAAGTTCCTGCCACACTGGACAGGAC
CTCACGAGATTGTGGACAAACTATCCCCTGTTGCATACCGGATCAAACTGAGCAGGG
GACAGAAAGATCCAGTTTTCAAATGGGTCCACCGTAACCAGATCAAGAGACATGTG
GCTGCTGGCAGGGAAAGGCAGGGGGGGGCTAGTGCTTAGTAGGACCTGCTGTAGTT
CTGCAACAAAGAACTGTCCTGAACAGCCGAGAACCTCCCTCACTGACCCACTGCCTT
TCTTCTTCTTCTTTCTCTTCTTCTTCTTCTCTGAATCAGGATGCTGCTGTGGCTCTCCC
TTCTGTGTCTCCAGCTAACCCGAGGACAGCTGCAACCTGATATCACCTCACCGGGTC
CAGCATCAGGCATTGTACTGAGGGACCAAGCTGGACTCCTCATCACCAACTGCAGG
ACTCACACTCAGAAAGTGTTTGTCAGACTTGACCCTCGAAAGGTCTACCTAACACAC
TACCCCCAGGCGACAGCCCACTACAGTTGGAGAGCTGCCCTATGGACTGAGAATGC
CCTCCTTCACGCAGAGGCTGACACCAAACACATGCTCACTCAGTTGCAGAAGATGA
CTGTCACCCAGTCAGAGTTAAGTGGACACAACAGACGCAACAAACGTTTTCTCGGG
GCTCTGCTCGGAGCTGCCGCTGCTGTTGGAACTCTATTCAACCTGGGGGTGACCAGC
ATTAACTCTGTGAGCCTGTCTACCCTGAGACAACACGTGAAGGAGATCCAGACTGA
GATTCCCCAGCTGAGAGAACATCTCACGCAACAGGGCGAAACCCTACAGACTATAG
GCAAATCATTAAAGGGCACCTTGGTGGTTCTCAACACCCACAGTGTCCTGCTGAACC
AAACTGTGAACTCCATGAAACAGTTATTCTCTGTCATACAGAATGATGTTGCCCACA
TTCAGCTGGTAGCCTCCCTAATGTCAGACATGCTGCGGGAGGTGAGCTCCTCCATTG
ACAGCTTAGCCATGGGTAGGATTCCACCATATCTGGTGCCGCTAAGCCTAGTACAGA
CTGTGTTGTCCTCGATCACCACTCATCCACCTGACTCTCTTCAAGCCCACCTCGCCTA

```
CTCCCTGGGAGGGTCCATCCTGCTGCATGTCGACCCTGAACAGAGTGAGCTGGCTTT
TCTCCTGAATCTACCCATTGTTGAATCCAACAACATTTATAGACTCAAGGACATTGTT
AATGTTGGGTTTTGGCAAGGAAACACTCACATCACGATAAACACCCCAGATGTTGTG
GCCTATCATGACAGCAACCCACAGCTATACCTGGCTCCGAACCTGCGCATGTGTACC
CTGACCAAAGACATTCACTATCTCTGCCCCAGCAAGCCATTCCTTCGAGACAACACG
GAAGGAATCTGCGGTCTGCAACCCATGCAGAGAGACATGGTCTGTCCCGCTGAGGC
CAAGCCTCGCACCCAGGTCACAGAAACACGAGCAGAGATCGTGGGAAACCAGTGGC
TTGTCAACACTCCCATCCGCACAGCCACACTGACTTATGACCAGCACGACACTGCCA
CCCGTGTCAGCCTGCCCAACCAGAGTATGTGGATTCAAGTACCCAAGGGAGCCATC
CTCCACCTAGGTGACCTGGCACTATACCATCTCCCAAGTGAAGAATACCAATCCGAG
ATAGAAATTCCAGCCTTCTTCAGGAATCAGAATCTCACCCTGGAACCTGAATTGGAA
CTGAGCATTGAGGAAGGGGGGTCCCAGACAATCGACATCACCCCACTCGACACGGC
CCTCCAAGCTCTGTCCCGCCTGCCTGTCCTAAACAACTTTCCCGTTGCACGGTCCTGG
ACCGCCGCCGACACGGCACTTTGCCTCGCCACAGCAGTCGGGTACACCCTAACCCTT
GGCCTTGCCTTCGTCCTGTTCAACAGAGTCAACGGGGTGCAGAGGTCCATGAACAGG
TGCACCGCTGCCCTTCCTCGGACCTTCAAACGGAAACACCGGAAAATGCTATATTAG
CCCAGTTACCCCTCGTTATTGGTACTAGTTAATCTCTAAATGTGTATTTTTCAATGTA
TGAATGTATCTCTGGTTTAATACCTTCTAGATGTTTCATTCTTGGTATCAAAACGATC
TGCACTTTATTGTATGAACAATGGTGTACACCATGTGATTAGAAAATGCATCATACT
GTTTGTACTATATTGTGGGTAAAACTGCATCAATTCCCCAGGCCAGCAATAAACATG
CAATTGGGCCATAAATGGAGACAAAGAGAAGTTCTCCCGCTCAGCAATCAAACTTT
GTATTGGCCAATTCACCCAAAGGTGGGTGTGGTAGGGAAACTATCAAAAGACCCTA
ACACCAACCCATCCTTGCTCAAACCCTTCGAGACACTCTCCAGGACTCATCCTTTGA
GTCAATTTTCCGGTAGTTTATCTTCAAGTAACTTTGTCGACTCGCTGTGGATTTTCCT
ATCTTCGAACAAACTAGAGAAAACCTCCAGAACACCTAAAGGAATTCATCGAGAAC
TTCCATTGTCTACCGCATCAGGACTCTTGTTCAGCAACAAAGCCAATGCAAGTAACC
TTTACAACTATTTAGTTGCGCGTTTTAAGTTGAGCCCCTGTTTAAGGTGATTTTGCTT
CTCTGATGATTCTCATGTGGTCGCTGGTTAACTGATGCTAAATACTGCCATTCTTGCC
TTCTCTCTCTCTCTCTCTTTTCTCACTTTCTTTCTCTATGTATATGTACACGTGTTTTGC
TTAGTTAGTTGTATGTTAACTGCATTTCCATTTATTAAATTGCATATATTCACAATTG
ATTGTCTCCGTGTGCTGCTCACAATTTGAAGTCACTGAACATTTGATCTTGCTACATG
CTATATTTTGAGTTTACCACAAGTTAAGTTACTGCTAGTACTGACTGAAGTAGGAAA
GTTATTCTTTCTTGGCCCGAAAGTAACCTTTCCTAGAATTGATGAATGATTATACTGT
CTATTCGGTGGACGAACAGATCAAATAATTGTAACGTTAATTCTGCTACAGTTAATT
CTAAGATCTGATTAAAATATTCATAATTAATTATAACCAGTTATAATTAATCATATTT
CCCTTTTGAGCTAATTTGCTACA
> Loki-Lro consensus sequence 5’LTR
TGGTACTAGTTAATCTCTAAAGGTGTATTTTTCAATGTATGAATGTATCTCTGGTTTA
ATATCTCCTAGATGTTTCATTCTTGGTATCAAAACAATCTGCACTTTATTTTCTGAAC
AATGGTGTACACCATGTGATTAGAAAATGCATCATACTGTTTGTACTATATTGTGGG
TAAAACTGCATCAATTCCCCAGGCCAGCAATAAACATGCAAATGGGCCATAAATGG
AGACAAAGAGAAGTTCTCCCGCTCAGCAATACAAACTTTGTATTGGCCAATTCACCC
AAAGGTGGGTGTGGTAGGGAAACTATCAAAAGACCCTAACACCAACCCATCCAATG
CTCAAACCCTTCGAGACACTCTCCAAGACTCTTCCCTTGAGTCAATTTTCCGGTAGTT
TATCTTCAAGTAACTTTGTCGACTCGCTGTGGATTTTCCTATCTTCGAACGAACTAGA
GAAAACCTCCGGAACACCTAAAGGAATTCATCGAGAACTTCCATTGTCTACCGCATC
```

AGGACTCTTGTTCAGCAACAAAGCCAATGCAAGTAACCTTTACAACTATTTAGATGC
GCGTTTTAAGTTGAGCCCCTGTTTAAGGTGATTTTGCTTCTCTGATGATTCTCATGTG
GTCGCTGGTTAACTGATGCTAAATACTGCCATTCTTGCCTCTCTGTCTGCTCTCTCTC
TCTCTCTCTCCTTTCTCACTCTCTTTCTCTATGTATATGTACACGTGTTTTGCTTAGTT
AGTTGTATGTTAACTGTAGTTCCATTTATTAAATTGCATATATTCACAATTGATTGTC
TCTGTGTGCTGCTCACAATTTGAAGTCACTGAATATTTGATCTTGCTACATGCTACAT
TTTGAATTTACTACGAGTTAAGTTACTGCTAGTACTAACTGAAGTAGGAAATGTATT
CTTTCTTGGCCAGAAAGTAACCTTTCCTAGAATTGATGAATGATTATACTGTCTATTC
GGTGGACGAACAGATCAAATAATTGTAACGTTAATTCTGCTACAGTTGATTCTAAGA
TCTGATTAAATATTCATAATTAATTATAACCAGTTATAATTAATCATAAATATTTCCC
TTTTGAGCTAATTTATTACA
> Loki-Lro consensus sequence 3’LTR
TGGTACTAGTTAATCTCTAAATGTGTATTTTTCAATGTATGAATGTATCTCTGGTTTA
ATACCTTCTAGATGTTTCATTCTTGGTATCAAAACGATCTGCACTTTATTGTATGAAC
AATGGTGTACACCATGTGATTAGAAAATGCATCATACTGTTTGTACTATATTGTGGG
TAAAACTGCATCAATTCCCCAGGCCAGCAATAAACATGCAATTGGGCCATAAATGG
AGACAAAGAGAAGTTCTCCCGCTCAGCAATCAAACTTTGTATTGGCCAATTCACCCA
AAGGTGGGTGTGGTAGGGAAACTATCAAAAGACCCTAACACCAACCCATCCTTGCT
CAAACCCTTCGAGACACTCTCCAGGACTCATCCTTTGAGTCAATTTTCCGGTAGTTTA
TCTTCAAGTAACTTTGTCGACTCGCTGTGGATTTTCCTATCTTCGAACAAACTAGAGA
AAACCTCCAGAACACCTAAAGGAATTCATCGAGAACTTCCATTGTCTACCGCATCAG
GACTCTTGTTCAGCAACAAAGCCAATGCAAGTAACCTTTACAACTATTTAGTTGCGC
GTTTTAAGTTGAGCCCCTGTTTAAGGTGATTTTGCTTCTCTGATGATTCTCATGTGGT
CGCTGGTTAACTGATGCTAAATACTGCCATTCTTGCCTTCTCTCTCTCTCTCTCTTTTC
TCACTTTCTTTCTCTATGTATATGTACACGTGTTTTGCTTAGTTAGTTGTATGTTAACT
GCATTTCCATTTATTAAATTGCATATATTCACAATTGATTGTCTCCGTGTGCTGCTCA
CAATTTGAAGTCACTGAACATTTGATCTTGCTACATGCTATATTTTGAGTTTACCACA
AGTTAAGTTACTGCTAGTACTGACTGAAGTAGGAAAGTTATTCTTTCTTGGCCCGAA
AGTAACCTTTCCTAGAATTGATGAATGATTATACTGTCTATTCGGTGGACGAACAGA
TCAAATAATTGTAACGTTAATTCTGCTACAGTTAATTCTAAGATCTGATTAAAATATT
CATAATTAATTATAACCAGTTATAATTAATCATATTTCCCTTTTGAGCTAATTTGCTA
CA
> Loki-Lro consensus sequence *smc-gag* gene
ATGTCTCGTTCATCCAGCCCTGATGCCCACTGGGATCGACTAGAGGCCTGGCTGAGT
GCCATGACGGGTACCCTTCTGCCTGACGCCGCTGGTGGCCTGCAGCACCTGAGCCGG
GAACAGCTTGACGACAACCTGAGTGGCTTGATGTCCCATGATCCCACGCAGGACTA
CAACCACAAAGACCTCGCCAAGATCATCGGGTCACTTGCCCACAACCTCCTCGCTCA
GGCAAGGCTGAGCGAAGCAAGCATCTCCCGCCTGGAACAGGAAGCAGCCGCGCTAA
AGCTCCGAGCTGAGGAGGCTCAGAGGGACCTGGCACTTGCCCAAAGTCAACTGGAC
CAGCTGGAGACGGAGACACCGCAAAGTGACTCTGAAAGGGAAATGCACACCACAG
AATTTCCCATAACCAGTAAACTAGTGCCTCCTAGTCAAGAACTCTCACTCCAGAAGG
GGGGTGAAAGCCCCCTTCCCAAGACATCTCCTGCCAGCATCCCAGAACACCCCACA
CCAGCACGGGGGTGGGAGCCTTTCAAGGGAACACTGCCTACCAGTCACGGTGTAGC
CCCCAAAGAACTAGACAAGCTGGCCAGAAACATTCCCATGTTTACCCCGAACCCTG
CAGGAGGCCATGACGTGCACGCCTACCTGAAAGACATTAATTTCCACTTGCAAGGTG
TGGCCAATGTAACCACACGAGACAGACTTTACCTGCTCAGAATCACCTCCAGCCGTG

AGGTGCGAAGCTTCCTAGATCGACAGCCCGAAACGGTCAAAGCGGATTACCAGCAG
CTACAACAGGCTTTGATCACGGAATTTTCTGACCCAGAGTCAGAACAAGGACTGATT
ACTGCCATGGACCTCAAACAGGCCCGACAAGAAACCTCGCAAGCCTACTACACCAG
GCTCCGACAGGCCTACTTCGGGGCACGGAACGAGCCAGGAATGGAAGAGGACTTTA
ACTTCAAGACTCTGTTCCTCCGGAACATGCACCCTACAGTGAGTCACCACTTGGGGG
TCTTGGCTTGCCCACGAAGCATGTCCATTCAACAGCTGCGAGACTTAGCCCACAAGG
CCTATGTCAAACAAAGAGCTGCTTCGGAAAAGACTGTGAAAACCCCCACAATTTGC
CCTGTCTCTGATCAGAGTCCAGAACTGGCTCTAGAGGGCGCCCAACAACACTACAGT
GACAAACCTTTCAACAAGATGTCAAGGCCATTCCCTGCCAGCAGAGGGCAGTATGG
CCATGGCGGTGCCCGTTCCAAGCATCGAACAGAGCGCCCTGAGACATCCTGGGACA
AGCCACAGCCAGCCTTCCACCCGAAGGGAGGAACCACGTGGGAAGGAAAAAGACG
GCACAAGGGCAGTCGACCTTCCCCAGCTCAAGCATCAAGCTCGAACTGCCAGCTAC
AGAATCCCTTTAAACATGCCCTAGCCCAGCCAAAGTCTGAACCATCAGGAGAGCAG
GACAGAGAAGCTACAACTGAGTCAGCAGAGATCCTGAGGCTTTTAAGGGGACTTCT
CCGAGAAAAGCACAACAAGGGAGACAAGAGGGACAAGCCAGATCAAGCATGA
> Loki-Lro consensus sequence Smc-Gag polyprotein
MSRSSSPDAHWDRLEAWLSAMTGTLLPDAAGGLQHLSREQLDDNLSGLMSHDPTQDY
NHKDLAKIIGSLAHNLLAQARLSEASISRLEQEAAALKLRAEEAQRDLALAQSQLDQLET
ETPQSDSEREMHTTEFPITSKLVPPSQELSLQKGGESPLPKTSPASIPEHPTPARGWEPFKG
TLPTSHGVAPKELDKLARNIPMFTPNPAGGHDVHAYLKDINFHLQGVANVTTRDRLYLL
RITSSREVRSFLDRQPETVKADYQQLQQALITEFSDPESEQGLITAMDLKQARQETSQAY
YTRLRQAYFGARNEPGMEEDFNFKTLFLRNMHPTVSHHLGVLACPRSMSIQQLRDLAH
KAYVKQRAASEKTVKTPTICPVSDQSPELALEGAQQHYSDKPFNKMSRPFPASRGQYGH
GGARSKHRTERPETSWDKPQPAFHPKGGTTWEGKRRHKGSRPSPAQASSSNCQLQNPF
KHALAQPKSEPSGEQDREATTESAEILRLLRGLLREKHNKGDKRDKPDQA*
> Loki-Lro consensus sequence *pol* gene
ATGGGATCTCACCTCAACTCCCACTCAGAGAACTCTCAAACTGAACCTCACAACCAG
ATCACTGCTAATCCACTGTCAACCACAGAACCAGATACTTCCTCACCAAGTGATGTC
CGCACTCAACCGTCTCAGCAGACTGTTGAACTGCCAGAGCCAACCAACATCCAAAG
CCACACTTCCAGCCAGGATTCCAACGTGCCTGAGAGTGCAGTTTTGGTAGTCTGCCT
CCCCGACGAGCCACATGAGACCAGCCCTGACTGCCTGGCAGTCAACACACAAGTCC
CAGTGCCAAGACTCCTGGGAAACCTGATCGAACGGGGGGTGGCTAAAAAACTTTAC
CTGGTCATCACATTGGAAAAGGAGGTGAGACTGGAAGCCCTCGTAGACACTGGAGC
CGACCTGACTTTAATGTCCAGCCAGCTTTTCGACAGACTTCAGACTGAAGCCAAGAG
ACAGAATCGAACTCTGAAATATCAGCAGTGTGTGCTGAATGTGCAGTCATACAGCC
AGAATGAGGTCCAACTTGAACAAGTAGCTCCCATTCACCTGACTATCGGCCCTATGA
GTTTGGTGCACCCTGTGTACATCGCTCCAGTTGACACCTACCCTCTTCTCATTGGCAA
AGACCTGCTAGACCGCTTTGAGCCCCTGTTAGACTTTAAACAGCTGAAAATCTGGGC
TCAAGTGCGAGAGCCTCTGCCTCCTCAGTCTCCAAGGTCACCCGAGACAGACTGCCA
GGCCACCAAGGTCACAGAAAATCCCACCGCCAGCCGTGGGAACGTCTCAACCCAAG
AGCAAGGTTCAAGGTCGATGCTCTGTGCTTTCCAGCTAGACAAAGACTTTGACTCAG
ATGGCCCCCAAATTCTGGCAGGCCTACAGCTGAATGATGTCACAACAACAGACCTC
ATACTGGCACTGTGGGCAGACAACTCGGCTATCAGCCTCACACTTTTCAACGCTCTC
AAGAGACACACAACAAACATCTCCTTTGCCAACAAGCGCACACGTTTCGCTTTGAAC
CCCGGCCTGTCAACCCTGGTGACTGCCAAGGTCATCTGTGCCCTGAACTTGAATTGG
AATGACAGGTGCTTAACGCATTACTTCCTGGTCCTGCCAGACCTACCGCATGATGTT

TACATCGGAGCGGACCTTCTGATTCGCCTCAAGGCTCACGTGGACCTGCTAAATGGA
GTAATATGGGCCCCAATGACCTCACAGGCCTCTGCTGTTCCTGTCAACCCCGAGAAT
CTCCTGTCAGGCCAAACCATTCCCGAGGCCTGCACACTAATCAGCGAACAAGAGAC
TGTAATACCAGCTTACACCAAAGGGGTCGCTGTTCGACTCAACCTGCGATGTGGCCA
AACCCTCAGCCACAGCCTAGGCTTCTTCCAGCCTTCACCCGAGTGTGTGGAACTGGG
CCTCACTCTAGAAGCCACCCCTCTCGTTGAAGTGACATCTCGGGCCTTGTACATCCT
GTTCAACAACTGCATGGCAATGGATGTCAGAGTGCCTAAAGCCTACCGTCTGGGTTG
GCTAGTGAGTTATGACTTCCAGGACTTTGAACTTGTTTTACCAGTAATTGGACCTCTC
CCTGTTTCCATGATTCCAGAGGGGAGCTCAAGCCACACAGTGTCGACAGTACCTTTC
AAGATCATCACCATCCTCTCAGTCGTGCCAGTGGTCAAGGAGCAGGTGTGCAGAAC
AGAGCTAACGGTGGACCAACACCTTGCTGTGTACACAGTCTCTAGCCATCCCACCTG
CGAGACTGACAACGCTACTGCACCCCCGACTGTCCACCAAACAGATGATGTAAAGA
CAGGAGAGCCGTACCCTGGTTTTGAGACACAAGTTCAGCAAATTCTGAGGGATGCT
GATGCCCTACGGGAGGATGCAGATCGCCACAAACTTCGTCAAGTCCTGTACAAATA
CAAAGACTCATTTGCCAAGGACTCTTTAGACTGTGGCCTCACTGACCTGCATGTGGT
GCGAATCCCAACGCACCCAAAAGCTCCACCCACCTTTGTCAAGCAGTACAAGATTCC
TCTCGAATCGTATGAACCTGTGCAGGACATCATCACCTCGATGCTTGAGAAGGGTGT
CATCCGTCCATGTAACAGCACCTACTCTGCCCCCATCTGGCCTGTCCTCAAACCAAA
CGGCAAGTGGCGGCCAACCATCGACTATAGGAAGCTGAATCAACAGGTGCCATTAT
CGCGGTGGCCCATGACACAGCTGGAACAGGAGATACCAAGAATCAGAGGAGCTAC
AATCTTCTCCACGCTTGATGTGGCCTCAGGATTTTGGACTATACCAGTGCATCCAGA
TGATCAACACAAGTTAGCCTTCCGTTTTGACAATCGTCAATACACGTTCACAAGGTG
CCCATTCGGCTATGCCAACTCACCCGCCGAATTTAACATCTTCCTCAACAAGGCGTG
CCCAGATGCCAGGGTAAGAGGAAACCTTGTTTTTGTGGATGATATCCTCCTGAAACA
CACAACGGTGGCTGATCATCTGAAGGAGATCGACCACATCCTGAACCAACTGACTA
CCGCTGGTGCCAAGATTGCCCTTCACAAAGGACAATGGTGCAAAATCAAAGTCAAC
TATGTGGGCCTGCTCATCGGACCACAGGGCATCGAGCCACAAACCAACCGTATCCA
GGCCATTCAAAACATCAAGCCTCCTGCCAACGTCTCAGAACTTCGCAGTTTCCTAGG
TGTGTGCAACTATTCACGTCAGTTCATTGAGAACTACTCCGACATTGCACGCCCTCTC
ACTGCACTCCTGAAAAAAGACTGCCCATTCATCTGGACAGACACTCAGCAAGATGC
CATGAATGAACTGAAGCAGCATTTGTGCACTGCACCATGCCTGGCTTACCCGGATCC
ACAGAAGGAGTTCTATTTGGATGCTGGATTCTCCAGCCACTGCCTTAGTGCCGGTCT
ATACCAACGGCATGACCAGGACAGACGCGTGGTCGCGTATGCCAGCAAGACCCTGC
TTCCACCAGAGTGCAAATACTCAGACTGTGAGAAAGCCCTGCTCTGTACAGTATGGG
CAATCCAGAGATTCTCCAACTACATCGGAGCACAGAAGGTGATCATAGAGACCTGC
CACCAGCCTGTCACCTTTCTCAACAGCCAACGTATCCGCGACGGCGTTGTCACAAAC
GCCCGCATAGCCACATGGCTAATGGCCCTCCAGGGGCGTAATGTGGAAGCACGCTA
TGCTCAAAACTACAAGTCGGCGCTGGGTAACGGCCTGGCAGCTTGCCAAAGCTGCT
CTGATGACACACCAACGGTGATCGTACAAGCTGAAACGCCCCCAGACCTGCCGCTT
ACCTGCCACAGGTATTTCGAAGAGAGCGTGTGCCAAGGCATGCCCACAGCCTACGT
CGATGGGTGTTCGTACAACCACGAAGGCCTCCTAAAGGCAGGTGCAGGCGTTCTCTG
GCTGAACGATCGACCGTGCCCACCACAACACTTCAAGTTGGGCCCCCAGTCATCACA
ATATGCTGAGGTAGCAGCTATTCTCATAACCCTGCAAATAGCATCAGCTCATAACAT
CAGAGACCTGCTTATTTGCACAGACTCTAACTATGCCAGACTTAGCTTCACATGCCA
TCTTCCAAACTGGAAACAAAATGGCTTCAAGACAGCAAACAACAAACCGGTGAAAC
ACCAACACCTGTTCCAGATCTGCGATGACCTCACAAGTAAACACGACATGATCATCT

ACTGGAAAAAGGTGAAAGGTCACTCAAAGCAACCTGGCACAGACAAAGACCTCAAT
GACCAAACAGACGCCCTTGCCAAAGCAGGGGCACTACACGGCGAGTTATGGACTCC
ACCAACCCACCCACCCACCCCTAACGTGGCAGCCATAACCCGCAGCAAACGAACTG
TACCCACCACAGTACCCACGTCACAGCCACTAATGCTGGCTCCTCAAATCTCCACGA
ATGACATTGCTGACATGCAAGCTACTGACTCCTCCATACAAACTATTCTTCACCACC
TCTCTGACCCCTCCAACCACCCTATCACACCTGCTCAGCTCGCTAATGCCCCAAGCC
TAAAACGGCTACACGACCTCAAACCCATGCTGCGTGTGATGGACGACATTCTTTGGT
ATGTCCCTGACGACAACGCTGCACCAAGGCTTGTGGTGCCGCAGGGCCAAAGGGGG
GTCTGGTTGATGAATGCACACGACGCAGCATGTGCGGACATCATGGCACAAGAGC
CACGTACGAGACACTTAAACAAGTGGCATACTGGCCCGGCATGCAGCAAGATGTGG
CAGAGTATGTGAAAGGATGCCTTGTTTGCTGTCAATTCCAACCGGCAAACCCAAACC
ACAGAGCTCCACTGCAGCGCAGAGGAGTCACATTTCCTTGGTCTGACCTGCAAATTG
ACTGGGTAGGGCCCCTACCCAGATCAACAAGGGGCAACAAGTATTTCCTCACAGTT
GTATGCCAGTTCACCAAGTGGGTGGAGTGCCTTCCAGCACCCAATGACACGGCCCA
AACTACCGCATACCTCCTGATGAATCACATCTTCTCCAGGTTCGGGCTACCGCTGAA
AGTCAACTCAGACAGAGGCACCCACTTCACAGCTGAGATCATGCAGCAGATTTGGA
AACTTCTGGGAATACATGCTCAACTGCACATCAGCCATCACCCAATTTCTTCGGGTC
AAGTGGAACGATCTAATCAGACCGTGGTCAGCATGCTAAGGAAATATGTTTCTGCCA
ACCAGAAAGACTGGGATGTGAAACTTCCCTTAGTGCTCATGGCCATCAGAGCAACTC
CAAATGAATCAACAGGTGTGTCGCCCTTTGAACTAATGACCGGCCGGCTAATGACAT
TGCCACTGCACCTGCTGTACCAGCCAGGAGACTCAAACCTTGTCACCGCTTACACCA
CACATCAGTACCTTGATGAACTACGCCGTCATTTGAGATCCACATTCGCTTTCGCCC
AACAACGACTCCAAAAGAGTGCAGAAGGCCAAAAAGCATACTACGACCGCAAGGC
GTCACACGCAGAACTGGATGTGGGTGACAAGGTTTGGTACTACCAGTACGCACAGC
CACCTCAGACAACCTCTCATCGGCTGTCAAAGAAGTTCCTGCCACACTGGACAGGAC
CTCACGAGATTGTGGACAAACTATCCCCTGTTGCATACCGGATCAAACTGAGCAGGG
GACAGAAAGATCCAGTTTTCAAATGGGTCCACCGTAACCAGATCAAGAGACATGTG
GCTGCTGGCAGGGAAAGGCAGGGGGGGGCTAGTGCTTAG
> Loki-Lro consensus sequence Pol protein
MGSHLNSHSENSQTEPHNQITANPLSTTEPDTSSPSDVRTQPSQQTVELPEPTNIQSHTSS
QDSNVPESAVLVVCLPDEPHETSPDCLAVNTQVPVPRLLGNLIERGVAKKLYLVITLEKE
VRLEALVDTGADLTLMSSQLFDRLQTEAKRQNRTLKYQQCVLNVQSYSQNEVQLEQV
APIHLTIGPMSLVHPVYIAPVDTYPLLIGKDLLDRFEPLLDFKQLKIWAQVREPLPPQSPR
SPETDCQATKVTENPTASRGNVSTQEQGSRSMLCAFQLDKDFDSDGPQILAGLQLNDVT
TTDLILALWADNSAISLTLFNALKRHTTNISFANKRTRFALNPGLSTLVTAKVICALNLN
WNDRCLTHYFLVLPDLPHDVYIGADLLIRLKAHVDLLNGVIWAPMTSQASAVPVNPEN
LLSGQTIPEACTLISEQETVIPAYTKGVAVRLNLRCGQTLSHSLGFFQPSPECVELGLTLE
ATPLVEVTSRALYILFNNCMAMDVRVPKAYRLGWLVSYDFQDFELVLPVIGPLPVSMIP
EGSSSHTVSTVPFKIITILSVVPVVKEQVCRTELTVDQHLAVYTVSSHPTCETDNATAPPT
VHQTDDVKTGEPYPGFETQVQQILRDADALREDADRHKLRQVLYKYKDSFAKDSLDCG
LTDLHVVRIPTHPKAPPTFVKQYKIPLESYEPVQDIITSMLEKGVIRPCNSTYSAPIWPVLK
PNGKWRPTIDYRKLNQQVPLSRWPMTQLEQEIPRIRGATIFSTLDVASGFWTIPVHPDDQ
HKLAFRFDNRQYTFTRCPFGYANSPAEFNIFLNKACPDARVRGNLVFVDDILLKHTTVA
DHLKEIDHILNQLTTAGAKIALHKGQWCKIKVNYVGLLIGPQGIEPQTNRIQAIQNIKPPA
NVSELRSFLGVCNYSRQFIENYSDIARPLTALLKKDCPFIWTDTQQDAMNELKQHLCTA
PCLAYPDPQKEFYLDAGFSSHCLSAGLYQRHDQDRRVVAYASKTLLPPECKYSDCEKA

LLCTVWAIQRFSNYIGAQKVIIETCHQPVTFLNSQRIRDGVVTNARIATWLMALQGRNV
EARYAQNYKSALGNGLAACQSCSDDTPTVIVQAETPPDLPLTCHRYFEESVCQGMPTAY
VDGCSYNHEGLLKAGAGVLWLNDRPCPPQHFKLGPQSSQYAEVAAILITLQIASAHNIR
DLLICTDSNYARLSFTCHLPNWKQNGFKTANNKPVKHQHLFQICDDLTSKHDMIIYWKK
VKGHSKQPGTDKDLNDQTDALAKAGALHGELWTPPTHPPTPNVAAITRSKRTVPTTVPT
SQPLMLAPQISTNDIADMQATDSSIQTILHHLSDPSNHPITPAQLANAPSLKRLHDLKPML
RVMDDILWYVPDDNAAPRLVVPQGQRGVWLMNAHDAACAGHHGTRATYETLKQVA
YWPGMQQDVAEYVKGCLVCCQFQPANPNHRAPLQRRGVTFPWSDLQIDWVGPLPRST
RGNKYFLTVVCQFTKWVECLPAPNDTAQTTAYLLMNHIFSRFGLPLKVNSDRGTHFTA
EIMQQIWKLLGIHAQLHISHHPISSGQVERSNQTVVSMLRKYVSANQKDWDVKLPLVL
MAIRATPNESTGVSPFELMTGRLMTLPLHLLYQPGDSNLVTAYTTHQYLDELRRHLRST
FAFAQQRLQKSAEGQKAYYDRKASHAELDVGDKVWYYQYAQPPQTTSHRLSKKFLPH
WTGPHEIVDKLSPVAYRIKLSRGQKDPVFKWVHRNQIKRHVAAGRERQGGASA*
> Loki-Lro consensus sequence *env* gene
ATGCTGCTGTGGCTCTCCCTTCTGTGTCTCCAGCTAACCCGAGGACAGCTGCAACCT
GATATCACCTCACCGGGTCCAGCATCAGGCATTGTACTGAGGGACCAAGCTGGACT
CCTCATCACCAACTGCAGGACTCACACTCAGAAAGTGTTTGTCAGACTTGACCCTCG
AAAGGTCTACCTAACACACTACCCCCAGGCGACAGCCCACTACAGTTGGAGAGCTG
CCCTATGGACTGAGAATGCCCTCCTTCACGCAGAGGCTGACACCAAACACATGCTCA
CTCAGTTGCAGAAGATGACTGTCACCCAGTCAGAGTTAAGTGGACACAACAGACGC
AACAAACGTTTTCTCGGGGCTCTGCTCGGAGCTGCCGCTGCTGTTGGAACTCTATTC
AACCTGGGGGTGACCAGCATTAACTCTGTGAGCCTGTCTACCCTGAGACAACACGTG
AAGGAGATCCAGACTGAGATTCCCCAGCTGAGAGAACATCTCACGCAACAGGGCGA
AACCCTACAGACTATAGGCAAATCATTAAAGGGCACCTTGGTGGTTCTCAACACCCA
CAGTGTCCTGCTGAACCAAACTGTGAACTCCATGAAACAGTTATTCTCTGTCATACA
GAATGATGTTGCCCACATTCAGCTGGTAGCCTCCCTAATGTCAGACATGCTGCGGGA
GGTGAGCTCCTCCATTGACAGCTTAGCCATGGGTAGGATTCCACCATATCTGGTGCC
GCTAAGCCTAGTACAGACTGTGTTGTCCTCGATCACCACTCATCCACCTGACTCTCTT
CAAGCCCACCTCGCCTACTCCCTGGGAGGGTCCATCCTGCTGCATGTCGACCCTGAA
CAGAGTGAGCTGGCTTTTCTCCTGAATCTACCCATTGTTGAATCCAACAACATTTATA
GACTCAAGGACATTGTTAATGTTGGGTTTTGGCAAGGAAACACTCACATCACGATAA
ACACCCCAGATGTTGTGGCCTATCATGACAGCAACCCACAGCTATACCTGGCTCCGA
ACCTGCGCATGTGTACCCTGACCAAAGACATTCACTATCTCTGCCCCAGCAAGCCAT
TCCTTCGAGACAACACGGAAGGAATCTGCGGTCTGCAACCCATGCAGAGAGACATG
GTCTGTCCCGCTGAGGCCAAGCCTCGCACCCAGGTCACAGAAACACGAGCAGAGAT
CGTGGGAAACCAGTGGCTTGTCAACACTCCCATCCGCACAGCCACACTGACTTATGA
CCAGCACGACACTGCCACCCGTGTCAGCCTGCCCAACCAGAGTATGTGGATTCAAGT
ACCCAAGGGAGCCATCCTCCACCTAGGTGACCTGGCACTATACCATCTCCCAAGTGA
AGAATACCAATCCGAGATAGAAATTCCAGCCTTCTTCAGGAATCAGAATCTCACCCT
GGAACCTGAATTGGAACTGAGCATTGAGGAAGGGGGGTCCCAGACAATCGACATCA
CCCCACTCGACACGGCCCTCCAAGCTCTGTCCCGCCTGCCTGTCCTAAACAACTTTC
CCGTTGCACGGTCCTGGACCGCCGCCGACACGGCACTTTGCCTCGCCACAGCAGTCG
GGTACACCCTAACCCTTGGCCTTGCCTTCGTCCTGTTCAACAGAGTCAACGGGGTGC
AGAGGTCCATGAACAGGTGCACCGCTGCCCTTCCTCGGACCTTCAAACGGAAACAC
CGGAAAATGCTATATTAG
> Loki-Lro consensus sequence Env protein

MLLWLSLLCLQLTRGQLQPDITSPGPASGIVLRDQAGLLITNCRTHTQKVFVRLDPRKVY
LTHYPQATAHYSWRAALWTENALLHAEADTKHMLTQLQKMTVTQSELSGHNRRNKRF
LGALLGAAAAVGTLFNLGVTSINSVSLSTLRQHVKEIQTEIPQLREHLTQQGETLQTIGKS
LKGTLVVLNTHSVLLNQTVNSMKQLFSVIQNDVAHIQLVASLMSDMLREVSSSIDSLAM
GRIPPYLVPLSLVQTVLSSITTHPPDSLQAHLAYSLGGSILLHVDPEQSELAFLLNLPIVES
NNIYRLKDIVNVGFWQGNTHITINTPDVVAYHDSNPQLYLAPNLRMCTLTKDIHYLCPS
KPFLRDNTEGICGLQPMQRDMVCPAEAKPRTQVTETRAEIVGNQWLVNTPIRTATLTYD
QHDTATRVSLPNQSMWIQVPKGAILHLGDLALYHLPSEEYQSEIEIPAFFRNQNLTLEPEL
ELSIEEGGSQTIDITPLDTALQALSRLPVLNNFPVARSWTAADTALCLATAVGYTLTLGL
AFVLFNRVNGVQRSMNRCTAALPRTFKRKHRKMLY*
>Loki-Cne consensus sequence
CAGTGCTTTCGGCATACAGCAGAAAAGAAATGAGTCGCCGAGAGAGTTTTACCGTC
GTTTACGAGCTGCATATTTTCAGGGAGCGCCCCTGGACAGGAGGAAGACCACACAC
TAAAATATCTGTTCCTCCACAATCTTCATGAAAGCATCAAATACGATGTTACCATGT
ACTGCCGAACCAGAACCCTTACCATGCAAGAAATAAAGAAGTATGCTCAGATGGCA
TGGGAAACACGTGCTCGGGTCAACAAGGAAACTGACCGGAGCAATAGGGTTTTACA
GGTCCGGGCCGAAGGAGGTAAAGACCTAGAACTGGAGGGACGAGAGATTCCCTCCT
CTGACAACAGATCCAGGAACGAATCCAAAGGACGTCAGAAGTTCAACCCACAGGAG
AAGAAAGGGATGATGAGAGAAAGATGGAGTCGAGACGGGAAAGAAGGTAAGGAGT
CCAATGTGAGAGTCAAACCGGACAGAGCGAAGTCTGACAGCCACCATGAAAAGAA
GGTAAGGCCTAAAGATGCAGACGAGAAAGGACAACAACAGATGATGAGGGAGATG
GAGGCCACCTTACGCAGGCTGCTTACAGAAACGGTGCAACAGATGACCAACGCATC
CCAGCCGCCAACAAACTCCACCAATACCCGAAACAAACCAGATCCAAAACCTCCGT
CCGCATGACTAGGCGTGGATCCAGACCCATCCACCAAAAGGGTCTACCGGATTGAG
TGGGGAACACAGCCCCTCGAACCGGGTGAATCCTCCGAAACCCCACCACACAAAAC
ACCGTTCTTGAGGTTCTTGGGTGAGCTCCATCTCACCAAAAACGCACGCCGCCTCTA
TTGTTCTGCCGTAATAGGTGGGTGCATCACAGCTGATGCTCTTCTGGACACAGGTTC
AGAAATTAGCCTAATGTGCGCCATGCTTTTCGAAAAGGTGATGAGAGCCATGAACA
CCCTTGGAAAACCATTCCAGGTGGAGACCTGCAATGTAGGGATCACCAGCTACACT
CAACATAAGTCTTGCATCACCAAGCGGGTGTGGTTAGATGTCACCTTCCGAGACATG
ACTCTAATCCATCCTATCTACATTTGCACATTAGATACAGAACCCTTTCTGGTGGGCC
AAGACCTACTGGACAGACTGGCCCCACTCATTGACTGCAACCAGGGACACCTGTGG
GCCCAAGTGGAAAATCCAAAGCCGTTCCCCTCTACAGAGAAACAGCTTGTCCCTACC
CATCTGGTAACATGTTCCAACGAGCCAGTGAAACACCAAGAGCAACCACTTCCTCCT
GCAGAACCGCTTTTCGGGGTTTCAGCTTCAAAGACACAGACGGACAACCAAAGTGC
ATCTTTCCGCAGTCATGAATCCTTCCTCTGCACTCTCACACCTTCCGATAATGAACCC
TTTAATCCCAGAGTGACTGAAGGCATTCACTTGGAAGATGTTTTCCAACCAAATGCG
GTTTTGGCCCTATGGTTGGAAAAGTCAGCCATAAGCAGAGGACTGTACCTTTCGCGA
GGCGAAGCCTGCCTGGTTGAAGTAAGACGTTGCCACCACCTACTTTCAGACGATCGG
CCCCGCCGACTCCTGAAGGCTACTGGGGTATGTATGCTTAAGGTAAACATTGGATCC
AAACACATATCACATACCTTTAGCATCGTGGAACATCCACTCACCCCGGTTCTGATT
GGCGCCGACCTCCTGGTGCGACTAGGAGTACAACTTGACACATGCAACCAGGTCTT
GTGGGCCCGGGCCCGAACACCACCCTGTTCTTTCCTTGAAAAACCCGAGGTCATGAG
GTCAGGGCAGACCATTCCCCAAGCCTGCCAGGTTGCATGTGAAACAGACATGCTGA
TACCAGCATATACCACCGGTGTCCCAGTCAGGCTAGTTATCCTTAAAGGTCAGGAAA
TGGAGAACCCCCGGGTCCATTTCCAACCTCTGCCTCACTTTCACGAACTGAATCTAG

TGGTCCGGGGAAACCCCTTGCTTGAGTTGAACAATCGTTCAACCTACCTCTTGGTCG
AAAATCCAACCGCGGTCCCAATCCCACTAAAAACCGCTAAACCTCTCGGGGTGCTA
ATCGATAGTTCGTTCCACGACTTCGAGTTGAGCATTCCCGTCATTGGAGAACTACCT
TCTTCCCTGATCACCACACAGGGTTTCACAGACAGTGCCTCCACCTTCCCCTCGAAA
ATGATAACTGTCAGAAGGGGAGAGGCGCTGCCTGTCATATCCCTGTTTGGTGCCACT
CTTGAAGCCGAAGGGAACTTGGTAGTGTACTCGCTAGAAGCCCGTTCCACAAAAGA
CAGCCAAGTGAACCTGCCCGAGCCGTACCCAGGGTTCGAAGACGAAGTAAACCAAC
AGTTAGAGAAGGCTGACGCTTTGTCCACTGAAGCTGAAAAAGTGGCACTCAGGGAC
CTTTTCTACGAGTTCCAGACCATCTTATCCAGAGACTCTAATGACTGCGGGGTTACA
GGTTTGCATACGGTCCGTATTCCCACGAAACCAAACGCTCCTCCAACCTTCGTGCGT
CAGTATAAAATTCCTCTAGCCGCTTATGGCTCAATCCAAGAGATCCTCAATAAGCTG
CTAGAGAAAAACATAATTCGGGAATGCAACTCCACCTACAACTCCCCCATCTGGCC
GGTGTTAAAACCGACAGGAAAATGGCGTCTCACCATCGATTACCGTGCTTTAAACAA
GCAGGTACCGCTTTCGCGTTGGCCCATGATTCACTTAGACCAAGAACTGGCAAAAGT
GAAAGACGCCCGCTTCTTCTCAACCGTGGACGTGGCTAATGGCTTTTGGACAATGAA
GGTGGACCCGGCTGACCGGTATAAATTGGCGTTCTCTTTTGGGAACCGCCAATATNG
GAACAGGTGTCCTTTCGGATACTCCAACTCACCAGCCGAGTTCAACATCTTCCTTCA
TAAGGCGATGTCCGATGCTACAGAACGAGGAAACCTCATTTATGTCGATGACATCCT
AATGAGGAGCAGAACCTTCAAAGAACATCTTGAGGAAATCTGNCACGTTCTAGGCC
AGCTAGCGGCTGCAGGAGCCAAGTTAGCATTAAGCAAGGGTCAGTGGTGCAGGAAA
AAGGTGGACTACGTAGGGCTCACGGTCGGACCAAACGGCATTGAACCTCAAGCCGA
CAGAATCAGGGCCATCCAGGATATAGCAGCCCCGACCAACCTCTCCGAACTCCGGA
GCTTCCTCGGGGTCTGCAACTACTCCCGTCAGTTTATCGAAGACTATGCAGACATGG
CGAGGCCACTCACCGAGTTGTTACGCAAAAACACTCCGTTCGATTTGGGAGAACCTC
AGGAGCAGGCTTTCCAATGGATGAAGCAGAAGCTGGTGTCAGCCTCCTGCTTAGCAT
ACCCAGACAAGGACCGTGAGTTCTATCTGGAAGCCAGCTTCTCATCTCACTGCCTGA
GTGCAGCACTGAGACAGAAATACGACCAGGACATGCGAGTCGTAGCATATGCGAGC
AAACCACTAAGCGAGGTTGAAATAAAATTTTCAGACTGTGAAAAGGCCCTCCTCTCC
ACAGTCTGGGCGGTGGAACACTTCCGCAGCTACATAGGAGGACAAAAGGTGATTAT
CGAAACATGTCACCAGCCAGTAACTTTCCTTAACAGTCAGCGTCTGAGGGAAGGGA
GAGTGTCGAACAGCCGGATCGCGACATGGATGATGGCATTACAAGGCTACGACATT
GAGGTAAGATACGGTCAAAACAACCGGATGGCACTGGGACAAGGGCTAGCCGGAT
GTAACCACTGCGAGTGCGATACACCTCAACCCCAACCTGACATACCCACAGCTGTG
GGAACAACGCTCTCCGCCCACCGCTATTATGATGAGAACATATGCAAAGGGCTCCC
AAGAGTGTATACCGATGGCTGTGCGTTCATGCACGAACACCAACTCCGAGCTGGGG
TAGGTATAGTTTGGACAGGAATCGACTCAATCCTCATCACGATCCAACAAGCAGCTG
CCTTGAACCAGAAACAGCTAGTGATCTGCTCCGACTCCAACTACGCTAGGCTTAGTT
TCATCTCCCACCTACCCGTTTGGAAAACTAACGGAATGAAGAACGCTAGGAATAAA
GAAGTCAAACACTCAAATATATTCCTAGCATGTGAAAAACTCACCACAGAGAACGA
CATGACCATCTATTGGAAAAAGGTTAAAGGACATTCTCGCACCCTGGGCCAAGACA
AGGATGGAAACGACGAAGCAGACCGCCTTGCCAGGCTGGGGGCAGAGAGTGGAAC
TCCCTGGGAGTTTCGTGAGGACTGGGCTAAGCCCGTTGAAAACCACACAGTTTGCAC
CATAACGCGACAACAGGTAAGAGCCAACCAAAACAAGCTCGGAAACGCTAAAGAG
ATGCTCTGCCTCGGACGCAAACCAAATGACCAGGACTTGTGTACCATGCAGGACCA
AGATCCGACCCTGCACGCCATCAAGGAGTTGGTATGCAATGGTACGTCACCTGGAA
ATCAACCATCACCAACTGGAGAACCCAAGGAGCTAAAGTCGTTTCGTCATAGCTTAC

CCCACTTGAGGATAGAAAAGGGGTTGCTAGTCTACTCCAAAGATGGACACGACCCA
CTCCGATGGGTCGTTCCAATGGATCACAGAGGAATCATGCTCGCTCACGCACATGAT
TCGCCTGTGGGAGGACACAGAGGCTACAAAGCAACCCTACACACGCTTCAGCAGGT
GGCATACTGGCCTTTAATGGCAAATGATNAACGAAGGGGGGTTACATTCCCATGGTC
ACACATACAGATAGACTGGGTTGGACCGGTTCCGAGATCAGCAAGAGGCAACAAAT
ACTTCCTCACCGTAACCTGCAGCTTCACAAAGTGGGTTGAGTGCCTACCCGCACCCA
ACGACACAGCTACCACCACCGCGGTTCTCCTGCTGAACCATGTGTTCAGCCGTTGGG
GCCTCCCACTGTCAGTTGACTCAGACAGAGGGACCCACTTCACTTCGAATGTTATGA
CGGCACTCTTCGAGATGCTAGGGGTTGAGGTCAGGTTTCACATCTCATATCATCCCC
AGTCATCAGGACAGGTGGAGAGGATGAACCGCACGGTGGTTAATATGCTAAGAAAA
TATGTTGACAGCCATGCTAAGGACTGGGACGTCAAACTGCCCCTGGTCCTGATGGCT
ATCCGGTTCACACCGCAGACATCCACTGGCGTTACGCCCTTCGAGATGATGACAGGC
AGGCAAATGACTCTTCCGCTACACCTTATCTATCAGCCAGAATATGTTAGCGTCACG
ACTGCCTATACAGCACATCAATACGTCTCCGATCTGAAAGATCATCTTAGAGCAACA
TTCGCATGGGCTCAAAAGAAATTAGAGGCCAGCGTGAAAGGAGCCAAGGCCTACTA
TGACCGCAAGGCCACACACACAGAGTATGAAGTGGGAGATAAGATTTTCTACTTCA
AATTTGCTCAACCACCCGGCCGAGCCAAAAAGTTCCTTCCACGATGGGCTGGACCAT
TCGAGATCGTGGGAAAACTCTCACCAGTGGCCTATAGAATAAGGACCTCCAAACCC
AGGCAAGCACCTGAATACAGGTGGGTCCATGGGAATCAAATAAAACCTTAAAATCC
TCCTCCCAACCAACCAAACGCGACACCGCCACGGAGGAAGGAAAAGATGCATAGA
GCTGAAGTTGTGCATAACAAATAGGGCCAAAAGACACATAACCTGTACACACTCCG
TCATAATTTGGGTCTCTCTATTGCAGATGGAGATCTTCTGGGTAATCGGCTTCCTCCT
GGGATCAACCATGGGCCAACCATCCAGTGTAATAGAGACAGGACCTCCAACCGGCA
CCGCTCTCCAAGACACACCTGGACTACTCATCACACACTGTTANCTCCATACACAAC
GGATGTTCGTCCGACTCGATCCATGGGACGTGTACCGAAAACATATTAAACTACCGC
CTCAAATAGGAGAAGGACGACAACAAGGCAGTCAGATCCGTGACACAGTGAACCAT
GCCAAACATACCACAACCCACATCCTTAAACAACTCGAAAAGTATATGGTTACTGA
ACAGGATCTCGTCGGAAGTAAACGCCCGAAAAGATTTTTAGGCGGACTCATAGCCT
CTGTGGCCGCAGTCGGATCACTGTTTTCCATAGGCCTCTCAGCGGCAAATGCTGTTA
GTCTGAAAACAATTCAAAAGCATGTACAAGAACTGGACTCTGAAATGCCAGAAATC
CAGCAAAGGCTCCTCATACAACAAAAACTACTGAACATCCAGACAAAAACACTTCA
AGGCACAGTGGTCACGGTTAACCTACACTCTGCCCTTATAAACAACACCATGCATGC
ACTGGAAACACTGACAGAAGTGGTAAAGAATGATCTAATGTACACGCAAGTTGTGA
GAGATTTAATGCAGGATCTCATCCGAGAGGTTGGTGAATCGGTAAACAGCCTAGCT
GAGGGAAGGATCCCTCCTTACTTAGTCCCCGTGGAACTAGTAGAAAATGTGCTGAA
ATCCGCTGTCACTGATGCTCTGCAATCACCGCAAACTCACCTGGCTTATAGCTTAGG
TAGTGCAATACCAATCTTTGTAGATCCCGATAACTTAGAAATCGGATTCATGTTAAA
TCTCCCAATCATTGAAAGAGCTAATGAGTATAGACTAAAATCAGTACTAAACGTTGG
ATTTTGGAAAGATAAAACCCACGTCCACTTAAAAACTCCGTCCATGTTGGCTTTCAA
CGACGCCAACCCCCAGCTCTATTTGGTCCCTAACTTAGACTTATGCACCAAAACTAA
AGAAATCCACTGGGTCTGCCCCAGCAACCCATTCCTTAGAGATGTAACAGACCGCCT
CTGCGGACTTACAGAACAAAACTGCCAAGCTAAAATGACTCTAAAAGATGCAGGAG
TAGAAACAATGATAGAGCGCGCAGGTAATAAATGGCTAGTGAGTACACCAGCTACC
GAAGCACTGCTGTCCTACGAACTACACGACACAGTCACTAGAATTCACCTTCCCAAT
CAAACAATGTTGATTTCAGTTCCTCAAGGCGCCGTGATCCATACTAGAAGTGTGGTA
AATGTACAACAGATGTCTGTGTCTCAATGTTTAGAAATACTGANNATAGTCTTACAC

CATCTCAATGTTGAACGATACGACTCAGAAATTGAAATTCTAGACGTCTTCAAAGGA
TATAACTTCAGTATTGATGTCGAAATTGGAAAACAACTCCTAGCAGCAGGCACTCAG
CTAGTTAAATTCAGCCTCACACCATCTGAATTACTCTCTACCCCACATTTCCATTATC
GACAAGACAATGCCTGGCCAAATACCCACACCATGTCTATTCCGATCATTGTTTTA
TTTTTGGTTGGGTCATCACTACGATAATGACGTATAAGGCATACCAAAAGATAAAGG
TGTTACAGAATAAGTTCTATCAACCATCCATTCATTATCTGCAGGACATTATATTTCC
ACAACACTTTGTGCAAAGGACATTAAAGGTTTCCCTTACTGAATGTAGTGAAACCGT
TCAGGGGGTTTTCGGGACCTTTCTGACCTCCTTGTCGCAGTTCCTGTCCATTCATCAC
ATTCTTTTCATTTGTCATATCCATTTTTTTTGTCATATCCATCTATGATGCTGGGAGGT
CGAGGGAACCGATGGCACAGCCAGCAACCCTCCATATTCTAGGTTTGAATTGGGTTG
CTGTCTCCTTGGAATTCCACACACCTGTG
> Loki-Cne consensus sequence *pol* gene
TTGAGGTTCTTGGGTGAGCTCCATCTCACCAAAAACGCACGCCGCCTCTATTGTTCT
GCCGTAATAGGTGGGTGCATCACAGCTGATGCTCTTCTGGACACAGGTTCAGAAATT
AGCCTAATGTGCGCCATGCTTTTCGAAAAGGTGATGAGAGCCATGAACACCCTTGGA
AAACCATTCCAGGTGGAGACCTGCAATGTAGGGATCACCAGCTACACTCAACATAA
GTCTTGCATCACCAAGCGGGTGTGGTTAGATGTCACCTTCCGAGACATGACTCTAAT
CCATCCTATCTACATTTGCACATTAGATACAGAACCCTTTCTGGTGGGCCAAGACCT
ACTGGACAGACTGGCCCCACTCATTGACTGCAACCAGGGACACCTGTGGGCCCAAG
TGGAAAATCCAAAGCCGTTCCCCTCTACAGAGAAACAGCTTGTCCCTACCCATCTGG
TAACATGTTCCAACGAGCCAGTGAAACACCAAGAGCAACCACTTCCTCCTGCAGAA
CCGCTTTTCGGGGTTTCAGCTTCAAAGACACAGACGGACAACCAAAGTGCATCTTTC
CGCAGTCATGAATCCTTCCTCTGCACTCTCACACCTTCCGATAATGAACCCTTTAATC
CCAGAGTGACTGAAGGCATTCACTTGGAAGATGTTTTCCAACCAAATGCGGTTTTGG
CCCTATGGTTGGAAAAGTCAGCCATAAGCAGAGGACTGTACCTTTCGCGAGGCGAA
GCCTGCCTGGTTGAAGTAAGACGTTGCCACCACCTACTTTCAGACGATCGGCCCCGC
CGACTCCTGAAGGCTACTGGGGTATGTATGCTTAAGGTAAACATTGGATCCAAACAC
ATATCACATACCTTTAGCATCGTGGAACATCCACTCACCCCGGTTCTGATTGGCGCC
GACCTCCTGGTGCGACTAGGAGTACAACTTGACACATGCAACCAGGTCTTGTGGGCC
CGGGCCCGAACACCACCCTGTTCTTTCCTTGAAAAACCCGAGGTCATGAGGTCAGGG
CAGACCATTCCCCAAGCCTGCCAGGTTGCATGTGAAACAGACATGCTGATACCAGC
ATATACCACCGGTGTCCCAGTCAGGCTAGTTATCCTTAAAGGTCAGGAAATGGAGA
ACCCCCGGGTCCATTTCCAACCTCTGCCTCACTTTCACGAACTGAATCTAGTGGTCC
GGGGAAACCCCTTGCTTGAGTTGAACAATCGTTCAACCTACCTCTTGGTCGAAAATC
CAACCGCGGTCCCAATCCCACTAAAAACCGCTAAACCTCTCGGGGTGCTAATCGATA
GTTCGTTCCACGACTTCGAGTTGAGCATTCCCGTCATTGGAGAACTACCTTCTTCCCT
GATCACCACACAGGGTTTCACAGACAGTGCCTCCACCTTCCCCTCGAAAATGATAAC
TGTCAGAAGGGGAGAGGCGCTGCCTGTCATATCCCTGTTTGGTGCCACTCTTGAAGC
CGAAGGGAACTTGGTAGTGTACTCGCTAGAAGCCCGTTCCACAAAAGACAGCCAAG
TGAACCTGCCCGAGCCGTACCCAGGGTTCGAAGACGAAGTAAACCAACAGTTAGAG
AAGGCTGACGCTTTGTCCACTGAAGCTGAAAAAGTGGCACTCAGGGACCTTTTCTAC
GAGTTCCAGACCATCTTATCCAGAGACTCTAATGACTGCGGGGTTACAGGTTTGCAT
ACGGTCCGTATTCCCACGAAACCAAACGCTCCTCCAACCTTCGTGCGTCAGTATAAA
ATTCCTCTAGCCGCTTATGGCTCAATCCAAGAGATCCTCAATAAGCTGCTAGAGAAA
AACATAATTCGGGAATGCAACTCCACCTACAACTCCCCCATCTGGCCGGTGTTAAAA
CCGACAGGAAAATGGCGTCTCACCATCGATTACCGTGCTTTAAACAAGCAGGTACC

GCTTTCGCGTTGGCCCATGATTCACTTAGACCAAGAACTGGCAAAAGTGAAAGACG
CCCGCTTCTTCTCAACCGTGGACGTGGCTAATGGCTTTTGGACAATGAAGGTGGACC
CGGCTGACCGGTATAAATTGGCGTTCTCTTTTGGGAACCGCCAATATNGGAACAGGT
GTCCTTTCGGATACTCCAACTCACCAGCCGAGTTCAACATCTTCCTTCATAAGGCGA
TGTCCGATGCTACAGAACGAGGAAACCTCATTTATGTCGATGACATCCTAATGAGGA
GCAGAACCTTCAAAGAACATCTTGAGGAAATCTGNCACGTTCTAGGCCAGCTAGCG
GCTGCAGGAGCCAAGTTAGCATTAAGCAAGGGTCAGTGGTGCAGGAAAAAGGTGGA
CTACGTAGGGCTCACGGTCGGACCAAACGGCATTGAACCTCAAGCCGACAGAATCA
GGGCCATCCAGGATATAGCAGCCCCGACCAACCTCTCCGAACTCCGGAGCTTCCTCG
GGGTCTGCAACTACTCCCGTCAGTTTATCGAAGACTATGCAGACATGGCGAGGCCAC
TCACCGAGTTGTTACGCAAAAACACTCCGTTCGATTTGGGAGAACCTCAGGAGCAG
GCTTTCCAATGGATGAAGCAGAAGCTGGTGTCAGCCTCCTGCTTAGCATACCCAGAC
AAGGACCGTGAGTTCTATCTGGAAGCCAGCTTCTCATCTCACTGCCTGAGTGCAGCA
CTGAGACAGAAATACGACCAGGACATGCGAGTCGTAGCATATGCGAGCAAACCACT
AAGCGAGGTTGAAATAAAATTTTCAGACTGTGAAAAGGCCCTCCTCTCCACAGTCTG
GGCGGTGGAACACTTCCGCAGCTACATAGGAGGACAAAAGGTGATTATCGAAACAT
GTCACCAGCCAGTAACTTTCCTTAACAGTCAGCGTCTGAGGGAAGGGAGAGTGTCG
AACAGCCGGATCGCGACATGGATGATGGCATTACAAGGCTACGACATTGAGGTAAG
ATACGGTCAAAACAACCGGATGGCACTGGGACAAGGGCTAGCCGGATGTAACCACT
GCGAGTGCGATACACCTCAACCCCAACCTGACATACCCACAGCTGTGGGAACAACG
CTCTCCGCCCACCGCTATTATGATGAGAACATATGCAAAGGGCTCCCAAGAGTGTAT
ACCGATGGCTGTGCGTTCATGCACGAACACCAACTCCGAGCTGGGGTAGGTATAGTT
TGGACAGGAATCGACTCAATCCTCATCACGATCCAACAAGCAGCTGCCTTGAACCA
GAAACAGCTAGTGATCTGCTCCGACTCCAACTACGCTAGGCTTAGTTTCATCTCCCA
CCTACCCGTTTGGAAAACTAACGGAATGAAGAACGCTAGGAATAAAGAAGTCAAAC
ACTCAAATATATTCCTAGCATGTGAAAAACTCACCACAGAGAACGACATGACCATCT
ATTGGAAAAAGGTTAAAGGACATTCTCGCACCCTGGGCCAAGACAAGGATGGAAAC
GACGAAGCAGACCGCCTTGCCAGGCTGGGGGCAGAGAGTGGAACTCCCTGGGAGTT
TCGTGAGGACTGGGCTAAGCCCGTTGAAAACCACACAGTTTGCACCATAACGCGAC
AACAGGTAAGAGCCAACCAAAACAAGCTCGGAAACGCTAAAGAGATGCTCTGCCTC
GGACGCAAACCAAATGACCAGGACTTGTGTACCATGCAGGACCAAGATCCGACCCT
GCACGCCATCAAGGAGTTGGTATGCAATGGTACGTCACCTGGAAATCAACCATCAC
CAACTGGAGAACCCAAGGAGCTAAAGTCGTTTCGTCATAGCTTACCCCACTTGAGGA
TAGAAAAGGGGTTGCTAGTCTACTCCAAAGATGGACACGACCCACTCCGATGGGTC
GTTCCAATGGATCACAGAGGAATCATGCTCGCTCACGCACATGATTCGCCTGTGGGA
GGACACAGAGGCTACAAAGCAACCCTACACACGCTTCAGCAGGTGGCATACTGGCC
TTTAATGGCAAATGATNAACGAAGGGGGGTTACATTCCCATGGTCACACATACAGA
TAGACTGGGTTGGACCGGTTCCGAGATCAGCAAGAGGCAACAAATACTTCCTCACC
GTAACCTGCAGCTTCACAAAGTGGGTTGAGTGCCTACCCGCACCCAACGACACAGC
TACCACCACCGCGGTTCTCCTGCTGAACCATGTGTTCAGCCGTTGGGGCCTCCCACT
GTCAGTTGACTCAGACAGAGGGACCCACTTCACTTCGAATGTTATGACGGCACTCTT
CGAGATGCTAGGGGTTGAGGTCAGGTTTCACATCTCATATCATCCCCAGTCATCAGG
ACAGGTGGAGAGGATGAACCGCACGGTGGTTAATATGCTAAGAAAATATGTTGACA
GCCATGCTAAGGACTGGGACGTCAAACTGCCCCTGGTCCTGATGGCTATCCGGTTCA
CACCGCAGACATCCACTGGCGTTACGCCCTTCGAGATGATGACAGGCAGGCAAATG
ACTCTTCCGCTACACCTTATCTATCAGCCAGAATATGTTAGCGTCACGACTGCCTATA

CAGCACATCAATACGTCTCCGATCTGAAAGATCATCTTAGAGCAACATTCGCATGGG
CTCAAAAGAAATTAGAGGCCAGCGTGAAAGGAGCCAAGGCCTACTATGACCGCAAG
GCCACACACACAGAGTATGAAGTGGGAGATAAGATTTTCTACTTCAAATTTGCTCAA
CCACCCGGCCGAGCCAAAAAGTTCCTTCCACGATGGGCTGGACCATTCGAGATCGT
GGGAAAACTCTCACCAGTGGCCTATAGAATAAGGACCTCCAAACCCAGGCAAGCAC
CTGAATACAGGTGGGTCCATGGGAATCAAATAAAACCTTAA
> Loki-Cne consensus sequence Pol protein
LRFLGELHLTKNARRLYCSAVIGGCITADALLDTGSEISLMCAMLFEKVMRAMNTLGKP
FQVETCNVGITSYTQHKSCITKRVWLDVTFRDMTLIHPIYICTLDTEPFLVGQDLLDRLAP
LIDCNQGHLWAQVENPKPFPSTEKQLVPTHLVTCSNEPVKHQEQPLPPAEPLFGVSASKT
QTDNQSASFRSHESFLCTLTPSDNEPFNPRVTEGIHLEDVFQPNAVLALWLEKSAISRGL
YLSRGEACLVEVRRCHHLLSDDRPRRLLKATGVCMLKVNIGSKHISHTFSIVEHPLTPVL
IGADLLVRLGVQLDTCNQVLWARARTPPCSFLEKPEVMRSGQTIPQACQVACETDMLIP
AYTTGVPVRLVILKGQEMENPRVHFQPLPHFHELNLVVRGNPLLELNNRSTYLLVENPT
AVPIPLKTAKPLGVLIDSSFHDFELSIPVIGELPSSLITTQGFTDSASTFPSKMITVRRGEAL
PVISLFGATLEAEGNLVVYSLEARSTKDSQVNLPEPYPGFEDEVNQQLEKADALSTEAEK
VALRDLFYEFQTILSRDSNDCGVTGLHTVRIPTKPNAPPTFVRQYKIPLAAYGSIQEILNK
LLEKNIIRECNSTYNSPIWPVLKPTGKWRLTIDYRALNKQVPLSRWPMIHLDQELAKVK
DARFFSTVDVANGFWTMKVDPADRYKLAFSFGNRQYXNRCPFGYSNSPAEFNIFLHKA
MSDATERGNLIYVDDILMRSRTFKEHLEEIXHVLGQLAAAGAKLALSKGQWCRKKVDY
VGLTVGPNGIEPQADRIRAIQDIAAPTNLSELRSFLGVCNYSRQFIEDYADMARPLTELLR
KNTPFDLGEPQEQAFQWMKQKLVSASCLAYPDKDREFYLEASFSSHCLSAALRQKYDQ
DMRVVAYASKPLSEVEIKFSDCEKALLSTVWAVEHFRSYIGGQKVIIETCHQPVTFLNSQ
RLREGRVSNSRIATWMMALQGYDIEVRYGQNNRMALGQGLAGCNHCECDTPQPQPDIP
TAVGTTLSAHRYYDENICKGLPRVYTDGCAFMHEHQLRAGVGIVWTGIDSILITIQQAA
ALNQKQLVICSDSNYARLSFISHLPVWKTNGMKNARNKEVKHSNIFLACEKLTTENDMT
IYWKKVKGHSRTLGQDKDGNDEADRLARLGAESGTPWEFREDWAKPVENHTVCTITR
QQVRANQNKLGNAKEMLCLGRKPNDQDLCTMQDQDPTLHAIKELVCNGTSPGNQPSP
TGEPKELKSFRHSLPHLRIEKGLLVYSKDGHDPLRWVVPMDHRGIMLAHAHDSPVGGH
RGYKATLHTLQQVAYWPLMANDXRRGVTFPWSHIQIDWVGPVPRSARGNKYFLTVTC
SFTKWVECLPAPNDTATTTAVLLLNHVFSRWGLPLSVDSDRGTHFTSNVMTALFEMLG
VEVRFHISYHPQSSGQVERMNRTVVNMLRKYVDSHAKDWDVKLPLVLMAIRFTPQTST
GVTPFEMMTGRQMTLPLHLIYQPEYVSVTTAYTAHQYVSDLKDHLRATFAWAQKKLE
ASVKGAKAYYDRKATHTEYEVGDKIFYFKFAQPPGRAKKFLPRWAGPFEIVGKLSPVA
YRIRTSKPRQAPEYRWVHGNQIKP*
> Loki-Cne consensus sequence *env* gene
CTGAAGTTGTGCATAACAAATAGGGCCAAAAGACACATAACCTGTACACACTCCGT
CATAATTTGGGTCTCTCTATTGCAGATGGAGATCTTCTGGGTAATCGGCTTCCTCCTG
GGATCAACCATGGGCCAACCATCCAGTGTAATAGAGACAGGACCTCCAACCGGCAC
CGCTCTCCAAGACACACCTGGACTACTCATCACACACTGTTANCTCCATACACAACG
GATGTTCGTCCGACTCGATCCATGGGACGTGTACCGAAAACATATTAAACTACCGCC
TCAAATAGGAGAAGGACGACAACAAGGCAGTCAGATCCGTGACACAGTGAACCATG
CCAAACATACCACAACCCACATCCTTAAACAACTCGAAAAGTATATGGTTACTGAA
CAGGATCTCGTCGGAAGTAAACGCCCGAAAAGATTTTTAGGCGGACTCATAGCCTCT
GTGGCCGCAGTCGGATCACTGTTTCCATAGGCCTCTCAGCGGCAAATGCTGTTAGT
CTGAAAACAATTCAAAAGCATGTACAAGAACTGGACTCTGAAATGCCAGAAATCCA

GCAAAGGCTCCTCATACAACAAAAACTACTGAACATCCAGACAAAAACACTTCAAG
GCACAGTGGTCACGGTTAACCTACACTCTGCCCTTATAAACAACACCATGCATGCAC
TGGAAACACTGACAGAAGTGGTAAAGAATGATCTAATGTACACGCAAGTTGTGAGA
GATTTAATGCAGGATCTCATCCGAGAGGTTGGTGAATCGGTAAACAGCCTAGCTGA
GGGAAGGATCCCTCCTTACTTAGTCCCCGTGGAACTAGTAGAAAATGTGCTGAAATC
CGCTGTCACTGATGCTCTGCAATCACCGCAAACTCACCTGGCTTATAGCTTAGGTAG
TGCAATACCAATCTTTGTAGATCCCGATAACTTAGAAATCGGATTCATGTTAAATCT
CCCAATCATTGAAAGAGCTAATGAGTATAGACTAAAATCAGTACTAAACGTTGGATT
TTGGAAAGATAAAACCCACGTCCACTTAAAAACTCCGTCCATGTTGGCTTTCAACGA
CGCCAACCCCCAGCTCTATTTGGTCCCTAACTTAGACTTATGCACCAAAACTAAAGA
AATCCACTGGGTCTGCCCCAGCAACCCATTCCTTAGAGATGTAACAGACCGCCTCTG
CGGACTTACAGAACAAAACTGCCAAGCTAAAATGACTCTAAAAGATGCAGGAGTAG
AAACAATGATAGAGCGCGCAGGTAATAAATGGCTAGTGAGTACACCAGCTACCGAA
GCACTGCTGTCCTACGAACTACACGACACAGTCACTAGAATTCACCTTCCCAATCAA
ACAATGTTGATTTCAGTTCCTCAAGGCGCCGTGATCCATACTAGAAGTGTGGTAAAT
GTACAACAGATGTCTGTGTCTCAATGTTTAGAAATACTGANNATAGTCTTACACCAT
CTCAATGTTGAACGATACGACTCAGAAATTGAAATTCTAGACGTCTTCAAAGGATAT
AACTTCAGTATTGATGTCGAAATTGGAAAACAACTCCTAGCAGCAGGCACTCAGCTA
GTTAAATTCAGCCTCACACCATCTGAATTACTCTCTACCCCACATTTCCATTATCGAC
AAGACAATGCCTGGCCAAATACCCACACCATGTCTATTCCGATCATTGTTTTTATTTT
TGGTTGGGTCATCACTACGATAATGACGTATAAGGCATACCAAAAGATAAAGGTGT
TACAGAATAAGTTCTATCAACCATCCATTCATTATCTGCAGGACATTATATTTCCACA
ACACTTTGTGCAAAGGACATTAAAGGTTTCCCTTACTGAATGTAGTGAAACCGTTCA
GGGGGTTTTCGGGACCTTTCTGACCTCCTTGTCGCAGTTCCTGTCCATTCATCACATT
CTTTTCATTTGTCATATCCATTTTTTTTGTCATATCCATCTATGA
> Loki-Cne consensus sequence Env protein
LKLCITNRAKRHITCTHSVIIWVSLLQMEIFWVIGFLLGSTMGQPSSVIETGPPTGTALQD
TPGLLITHCXLHTQRMFVRLDPWDVYRKHIKLPPQIGEGRQQGSQIRDTVNHAKHTTTHI
LKQLEKYMVTEQDLVGSKRPKRFLGGLIASVAAVGSLFSIGLSAANAVSLKTIQKHVQE
LDSEMPEIQQRLLIQQKLLNIQTKTLQGTVVTVNLHSALINNTMHALETLTEVVKNDLM
YTQVVRDLMQDLIREVGESVNSLAEGRIPPYLVPVELVENVLKSAVTDALQSPQTHLAY
SLGSAIPIFVDPDNLEIGFMLNLPIIERANEYRLKSVLNVGFWKDKTHVHLKTPSMLAFN
DANPQLYLVPNLDLCTKTKEIHWVCPSNPFLRDVTDRLCGLTEQNCQAKMTLKDAGVE
TMIERAGNKWLVSTPATEALLSYELHDTVTRIHLPNQTMLISVPQGAVIHTRSVVNVQQ
MSVSQCLEILXIVLHHLNVERYDSEIEILDVFKGYNFSIDVEIGKQLLAAGTQLVKFSLTP
SELLSTPHFHYRQDNAWPNTHTMSIPIIVFIFGWVITTIMTYKAYQKIKVLQNKFYQPSIH
YLQDIIFPQHFVQRTLKVSLTECSETVQGVFGTFLTSLSQFLSIHHILFICHIHFFCHIHL*
>Loki-Bpe consensus sequence
ATGGCCGACCCAACCGCAGCTGACGCCCAGCTGTCAGACGACCTGCTAAGCCAGGC
CGAAACTGCCCTTTTCCATGCAACCCAACCAGTAGTACCAGGATACTGGGAAAAGA
GAAAACCTGCAACCATCAAAGAACATGAAATTGAGCTTAAAACTCTAGCAAGTAGT
GTTCCAAAAGCAGATGCGGAGGACGACGAGTGGACCATTTGGCTGAGCCATATGTC
CGTCACCTACCACCTCCTTAGTGTAAGGAACAAAACATTAATCCTCCGCCTTAAAAG
CTCCGCAGATGCCAGAAAAGCTGAGTTAAGTGCCCTACAGGACCAAATTGCCAGAT
TGAGGACAGAACTGGAGCAGGCCGGAGTGACCAACCTCTGTGGACTTTCAACGACT
AAAAATTTGGGCACAGGTAAGGAAACCCCTCCCAATGCCAGCCCCACCAGCACTAG

CCCAGTGCTGCACGCTGGAGATTACATCTCCAGAGCTCGACGCCAAAGGCGACCTTC
GAGGGACAGCTTCTTATGTGAATTTACTTCTTTTGAGGAATCCCCAAGTCCGCGGCT
AGTGGACGGCGTGACTCTTGAGACCACCAGGGTGGATGACGCAAGACTAGCCATAT
GGACGGACAGGTCAGCTATTAGCCTTGAACTGTACAAGATGCTTAACGGTATGAAA
CAAAACATACCCTTCATGCCTAAAACGTCACGGTTCCCCACAGATCACAAGTCTGGG
AATGTCTTTACAGCTGTTGGCTTCTGTTCACTGTCAGTGACTTTAAAAAACAGAAAA
CTAAACCACGTTTTCGTGGTAGTGCCTAACCTATCACACCCGTTGTACGTCGGGGGG
GACATTCTTGTCCGATTGGACGCCCAAATAGACACCATAAACAATGTCATATGGTCT
CTCATAGGGGTGTCTGAAAATGCCTGCACGCCATCCCAGGAAAACCTGAAGTCTGG
GCAGACCATTCCCGAAGTGTGCCAGGTCACCAACCAGTGCCCGGTGGTCATCCCACC
CAATACCGCTACGCCCTTAACCGTGGTTGCCTGCCCTGGACAACAGACCACCAACAA
CCAAGCCTTCTTCCAACCTTCCCACCTGTTTTCAGAACTTGGGCTCTCCGTCGAGGCC
ATGCCCTTGCTCATGGTCAAGAACCGCATGGCATGTCTTGTGGTGGTAAACAACACT
ACCCAAGAAGTCCATATTCCCCGAGCTACTCCTTTAGGCTGGTTGATTGATACAAAG
TTTCATGACTTTGAACTCAAACTTCCCGTGATGGGAAAACTGCCAGCATCTTTCCTTG
AAGACCCCGACAGGAAGGTGTTCTATACAAAACCCTTGAGGGCTATCGCCCTGCAC
TCAATGCTTGTTCCGGACGTAGAGGACATCTGTCGCGTAGACCTGACGGACCAGAA
AGAAATGGTCATTCAAACGGTTTCCGTCATTAGTCTTGAGCCCGGAAAACCTGATGA
AACGCCAGACGCTACAAGCCCGACTCCACAAGCTGACTCCTCGGGTTCTAACCCTAC
CGTTGAGGAGCAATTGAACCAGGTGATAGCTGAAGCAGATGGGCTTAACAGCGACG
ATCAGCGTGAAAAACTTCGAGAAATACTGATGAAATATAAGGCCTTTTTCTCCACAG
GGCCTACGGACTGCGGCCTTACAAACATTCACACGGTCCGTATTCAGACCGTGCCGA
ACGCACCCCCAACTTACGTACGCCAATACAAAATACCCCTGGCGTCATACGAACCG
ATTCAAGAAATAATCGATGATCTGCTGACGAAAGGCGTTATACGACCTTGCAACAG
CACATATTCAGCGCCGATATGGCCCGTGTTGAAGCCCAACGGTAAGTGGAGACTGA
CGGTCGACTACCGACAGCTAAACACTAGAGTGCCCCTGTCTCGATGGCCTATGACCC
AACTTGGACAGGACCTGCCAAAGGTGAGAGACGCCAAGTACTTCTCCACGCTTGAT
ATCGCGTCTGGTTTCTGGACTATACCAGTACACCCTGACGACCAACACAAACTCGCC
TTCAGCTTCGCCAATCGTCAATACACCTTCACACGCTGTCCTTTTGGACTAGCCAACT
CACCGGCTGAGTTCAACATCTTTTTGAACAAAGCATGCCCCAATGCAGCCAGTCGTG
GAACGCTCATCTACGTGGACGATATCTTGATCAGAACTACGACATTGGACGATCACT
TGAAAGAAATCGACCATGTGTTGGACCAACTAACCTCCGCAGGAGCAAAGATTTCT
CTGTCAAAATGCCAGTGGTGCAAGACAAAGGTAAACTATGTAGGCCTCCTAGTNCG
GAGCAGTAGAATTCAAGGAATCACAAACCTCAAAGTGCCTGCAAATGTGTCTGAAC
TACGCAGCTTCCTCGGAGTTTGTAATTACTCAAGACAATTTGTCGAACACTACGCCG
AGTTAGCTAAACCGCTAACCGACCTCCTCAAGAAAGACGAAACTTTCTCATGGGGC
CAAAGCCAAGACAATGCCATGAAAGCCCTGGTAGACCAGCTCTGCCAGGCCCCATG
CTTGGCGTATCCGGACCCCAACAAACCCTTTCACCTCGAGGTAGGCTTCTCAGAGCA
AAGCCTCAGCGCAGGTCTCTACCAAATTCATGATACCGACAAACGAGTCATTGCTTA
TGCCAGCAAGCCGCTGACAGACCCCGAGAGGAAATTCTCGGATTGTGAGAAAGCAC
TCCTAGGCACAGTTTGGGCGGTGAAACATTTCTCAAACTACGTAGGTGGACAAAAA
GTGGTGGTGAATACCAACCACCAGCCAGTAACATTCCTAGCCAGTCAGAGACTCAG
AGACGGGGTGGTCACTAACGCCAGAGTGGCATCTTGGCTCATGGCACTCCAAAGCTT
TGACATAGAAGTGCGCTACGCACAGAACCACAAGTCCCACCTAGGAATGGAACTGG
CCGCATGTCAAAGCTGCGAAAAGGACATACCGGCAGGAACTGCACCATGCGCAGAA
CCCCTACCAGTGTCGAACCAACACCATTACTTTGACCCAAACACATGCCAGGACATG

CCAACCGTCTACGTTGATGGTTGCTCCTTTCACCATGAGGCTACAGTCCGCGCCGGA
GCTGGAACCGTATGGCTCAACAACGACCCATGTGAGCCCGAAAAATTTCAACTAGG
CGCACAAACCTCACAGTACGCTGAAATTTCCGCAGTGCTGATCGCACTGAACCAAG
CAGCTGCTCATGACGTAGACACGTTCGTCATTTGCACTGACTCGAACTATGCCCGTC
TTAGCTTTTCATGCCACCTCCCATTATGGAAAAGCAACGGCTTCCTCACAGCTGCGA
AAAAGTCTGTTAAACATAAGGAACTGTTTATAGCATGTGACCGACTTGTCGAAACCA
AAAACATGAAAGTTTACTGGAAAAAGGTCCGAGGCCACTCACGTGAACCAGGACCT
GACAAAGAACTGAATGACAGGGCAGACGCCCTGGCTAAGCAAGGCGCTTTAGACGG
TACCCCCTGGCGTTTCCAACCAAAACCGCCAGACGTGATACCCCTTCCCACCCAACC
CACCCTTTCTGTCTGTGCCGTCACTCGCTCCCGCGCAGTACCTACTTACGACCGCCGA
GACACAGTCTCCGCCTCCGTAAACCTGGCTCTCACTGACTCTGACTTAACGTCACTT
CAGTCTTCCGACGCCGCTATTAGCCGCATGGTCACCTTCCTGTCCGACCCCTTGCTCC
ACCCAATCTCGGCCGATGATCTGGCCTCCGACCAGGCCCTCAATCTTTTGTGGCGCG
TCCGCGACAAACTCCTCCTACGTCATGGCCTGTTGCTATACGCCTCTGACTCCCTCAC
TTCACCCGCGTTCGTAGTGCCTCAGAGCCTCAGGGGGGTTATGCTCACACACGCACA
TGACACACCAGTCGCAGGCCACAGGGGGGTACAAGCCACATACCAGGCCCTACAGC
AGGTGGCCTATTGGCCACACATGAGAAAAGATGTACAAGAATACACAAAAGGGTGC
CTGGTTTGCTGTCAGTTTCAGCCCTCTAGACCGCTCCACCGTGCACCTCTGCACAAA
AGAGGCATTACCTTCCCCTGGTCAGACTTACAAATAGACTGGGTCGGACTGCTAACG
AAATCCGCAAGAGGAAACAAGTACATCTTGATGGTCACCTGTGCTTTCACCAAATGG
GTAGAGGGCTTACCAGCCCCGAATGACACGGCTCAAACCACGGCCTTCCTGCTCATG
AACCATGTGTTCTCTCGCTTCGGCTTGCCGTGCAAAATCGACTCTGATCAAGGTACG
CACTTCACAGCCGCAGTAATGTCCGAATTGTGGCAGACCCTTGGCGTAAAAGCTACC
TTCCACATCAGTCACCACCCAATGGCCTCTGGGCAGGTGGAAAGAACCAACCGGAC
TGTGATTAACATACTGAAAAAGTATGTGTCAGCTAATCAGAGAGATTGGGACGTGA
AACTTCCCCTAGTCCTCATGGCTATCAGAGCCACCCCCAACAAATCAACCGGGATTT
CCCCCTTTGAGATGATGACAGGAAGACAAATGACTCTGCCCCTGCATCTATTGTACC
AGCCAGGTGAAGCCAGCATTGCCACAGCTTACACCACACACCAGTACATGACAGAC
TTGCAGAGACATTTAAAAGCAACCTTCGCATTTGCCCAAGAGAAACTAGAGAAAAG
TGCAGAAGGACGTAAGGCCTACTATGACCAGAAAGCGTCACACGAGGAACTCCAAG
TAGGCGACAAGGTCTGGTACTATATCTTTGCACAACCTGCAGGACAATCCCAGGAAC
GCACGGCCAGGCTAATCCGGAAGTTCCTACCCAGACGGTCGGGGCCGTACGAGATT
GTGAACAAACTATCCTCGGTGGTATATCAAATAAAGATTACCAAAGGACACAAAGA
GCCAACCCTCAAGTGGGTGCACCGAAACCAGATCAAACCTCACAAAAACCCCATGG
GACAAGTCGGGAACACTGAGGCGCCCCATTAAGACACTTGAGAGTCTAAACCCACA
TTAGGATAGGGACCTTTACTAACACTCGCAATATCCCTAACCCTAAATTAACCAACT
AATGTTAACCTCTAACTAATCAATAAAAATTACTCAACAAAGTGAGACCAAAGTTCT
AAAGATACTCGTTTGTGTGTTGACAGGAACATGGAACCGCGTTTCCTGATCCTACTG
GTCTGCGCGTCCAGCCTAGTAAAAACAGAAGTTGTAGAGCCCGGACCGGCCTCAGG
AATCGTCCTTAAGGACCAACCTGGATTACTCATAACGAACGTCCGACTCCACACTCA
AAAGGTCGTCGTGAGACTCAATCCCCGCGACGTATGCAGGAAAAACGTCCCTATCA
CGGCGAGCCTCACGGGTTGGGCTGGCACTCAGTGGAACCTAGAGGCCATCCAACAC
GCGGAATCAGACATAACACATATGTTAGAGCAACTCGAGAAATTCACAACAACCCA
GGCAGAAATTAGAGGTATGCCGCACCCACGCTCTAAACGCTTTATAGGAGGCATCTT
AGCCGCAGCTTCCACCGTAGGAACCATATTTAACCTCGGGTTAACCACGATAAACAC
GGTCAGCATAGCCTCCCTCAAGCGCCACGTCAGCGAATTACAAGGAGAATTACCGG

TGATCAAAACCCAGCTCGATCAACAGCAGAAAGATCTAGTGACTTTCGGTGAAACC
CTCCGCGGTACAGTGCTGGTAGTAAACACCCATAGCATAGCCCTAAATAAAACAAT
ACAGGCTGTCAACTCCCTCATGGCCGTAGTTCACGTGGATTATGCGCACCTACAGTT
AGTCACCATGCTGTANAACGACATGCTTCGCGACATAGGGTCCTCGGTAGATAGCCT
AGCCATGGGTCGAATCCCACCATACTTGGTGCCCCTGACCCTGGTACAAGACTTACT
ATCCACTGCGACCCGCGACATAGTCACGCCTCTCCAGGCCCACCTAGCATACACGCT
CGGAAGCGCCATCCCTATCCTCGTGAACCCCGCAGGCCGCGAAATCGCTTTCTTAAT
TAACCTGCCAATCGTAGCTTCCGATAACATTTACCGCCTTAAAGACATAAATGTCGG
GTCGTGGCAGGGAGAAATGCACGTTAGAATTCAAACCCCAAGCCTGGTCGCATATC
ACGACAGTAACCCAGACCTGTACTTGACACCAAACCTACGTATGTGCTCTCTTATGA
AAGACATTCATTATCTCTGCCCCAGCAAACCTTTTGTCAGAGATAGCACCGACGGCA
TCTGCGGGCTCAAACCCATGACACCGGGTAACAGATGCCCGACCACAATCACGGCT
CGTGGGAAAGTCCTAGAAACTCAGGCCGAAATCGTGGGGGGGCGCTGGCTAATAAA
CACTCNGGCAACCCTCACGTACGACAAACACGACACCTCGTCCAGAATCATACTGC
CAGAACAGACAATCTGGGTGGATATCCCAACTAACGCCATCCTCCATATTGGGGAC
ATCGCGCTATTTCCGCTCAGTCCGAACCAATATGAGGAAAACTTTGAAGTCCCAGAG
TTCTTTGCGGAAAGAACTCTGAAACTCCAACCTGCAACGATTAATCAAGTGTCCTAC
GAAGGCACCCAAATAATCGATATCACACCCGTCGACAACCTCCTGACAGAAATTAC
CAATCGACCTCTAACGCCGATCCAACCCGTGTCTTACTCTTGGTCAGGCCCTGACTC
ACTCCTTGCCATCGCGATAGGAATTGGGTATATTTGCGCTATAATTGCCACCCTTATC
TTAAACAAAAGATTGAACAGAATTATAGGGCAATTCAGGCAGCACAGAAGAACTAG
GGCAGTCAAACGCCAGTTGACAACAGAGGGCACAAACCGAATCGACGAACGCGTGT
CGTACTCAGTCCGGGACGCAGACTATGAGGATCTCCAAACAACCAATGAAGGGTTA
GTTGTGATCCCGGCCCCAGGCGTCTAGCAGTGCACCACTAACCTCGAGAACCACGC
GCTGTCGGCATCTACCAGTGCACGTTCTCCCGAAACCAAATTTCACACTCTATGCTC
AGATTCTGACTACCGAAGGAACTGAATCATAGGCAGATCACTACATCTTATAATTTT
TTTCACCTATTTCTTTTTCTTGTTATATACACTTATAAGAAGTTTTGTTATATCCTTAA
AACGTCCCTAATTTCAATCCAGGGAATCCATTACCGCAAACAAAAACAACAACCAA
AAAAAAAAACAATTTTTTTTATTATGTTTGCACCAAACTTTGTTTCGTCACTACCTAC
ATATTTGGATAGTGCATAAAA
> Loki-Bpe consensus sequence *pol* gene
ATGGCCGACCCAACCGCAGCTGACGCCCAGCTGTCAGACGACCTGCTAAGCCAGGC
CGAAACTGCCCTTTTCCATGCAACCCAACCAGTAGTACCAGGATACTGGGAAAAGA
GAAAACCTGCAACCATCAAAGAACATGAAATTGAGCTTAAAACTCTAGCAAGTAGT
GTTCCAAAAGCAGATGCGGAGGACGACGAGTGGACCATTTGGCTGAGCCATATGTC
CGTCACCTACCACCTCCTTAGTGTAAGGAACAAAACATTAATCCTCCGCCTTAAAAG
CTCCGCAGATGCCAGAAAAGCTGAGTTAAGTGCCCTACAGGACCAAATTGCCAGAT
TGAGGACAGAACTGGAGCAGGCCGGAGTGACCAACCTCTGTGGACTTTCAACGACT
AAAAATTTGGGCACAGGTAAGGAAACCCCTCCCAATGCCAGCCCCACCAGCACTAG
CCCAGTGCTGCACGCTGGAGATTACATCTCCAGAGCTCGACGCCAAAGGCGACCTTC
GAGGGACAGCTTCTTATGTGAATTTACTTCTTTTGAGGAATCCCCAAGTCCGCGGCT
AGTGGACGGCGTGACTCTTGAGACCACCAGGGTGGATGACGCAAGACTAGCCATAT
GGACGGACAGGTCAGCTATTAGCCTTGAACTGTACAAGATGCTTAACGGTATGAAA
CAAAACATACCCTTCATGCCTAAAACGTCACGGTTCCCCACAGATCACAAGTCTGGG
AATGTCTTTACAGCTGTTGGCTTCTGTTCACTGTCAGTGACTTTAAAAAACAGAAAA
CTAAACCACGTTTTCGTGGTAGTGCCTAACCTATCACACCCGTTGTACGTCGGGGGG

GACATTCTTGTCCGATTGGACGCCCAAATAGACACCATAAACAATGTCATATGGTCT
CTCATAGGGGTGTCTGAAAATGCCTGCACGCCATCCCAGGAAAACCTGAAGTCTGG
GCAGACCATTCCCGAAGTGTGCCAGGTCACCAACCAGTGCCCGGTGGTCATCCCACC
CAATACCGCTACGCCCTTAACCGTGGTTGCCTGCCCTGGACAACAGACCACCAACAA
CCAAGCCTTCTTCCAACCTTCCCACCTGTTTTCAGAACTTGGGCTCTCCGTCGAGGCC
ATGCCCTTGCTCATGGTCAAGAACCGCATGGCATGTCTTGTGGTGGTAAACAACACT
ACCCAAGAAGTCCATATTCCCCGAGCTACTCCTTTAGGCTGGTTGATTGATACAAAG
TTTCATGACTTTGAACTCAAACTTCCCGTGATGGGAAAACTGCCAGCATCTTTCCTTG
AAGACCCCGACAGGAAGGTGTTCTATACAAAACCCTTGAGGGCTATCGCCCTGCAC
TCAATGCTTGTTCCGGACGTAGAGGACATCTGTCGCGTAGACCTGACGGACCAGAA
AGAAATGGTCATTCAAACGGTTTCCGTCATTAGTCTTGAGCCCGGAAAACCTGATGA
AACGCCAGACGCTACAAGCCCGACTCCACAAGCTGACTCCTCGGGTTCTAACCCTAC
CGTTGAGGAGCAATTGAACCAGGTGATAGCTGAAGCAGATGGGCTTAACAGCGACG
ATCAGCGTGAAAAACTTCGAGAAATACTGATGAAATATAAGGCCTTTTTCTCCACAG
GGCCTACGGACTGCGGCCTTACAAACATTCACACGGTCCGTATTCAGACCGTGCCGA
ACGCACCCCCAACTTACGTACGCCAATACAAAATACCCCTGGCGTCATACGAACCG
ATTCAAGAAATAATCGATGATCTGCTGACGAAAGGCGTTATACGACCTTGCAACAG
CACATATTCAGCGCCGATATGGCCCGTGTTGAAGCCCAACGGTAAGTGGAGACTGA
CGGTCGACTACCGACAGCTAAACACTAGAGTGCCCCTGTCTCGATGGCCTATGACCC
AACTTGGACAGGACCTGCCAAAGGTGAGAGACGCCAAGTACTTCTCCACGCTTGAT
ATCGCGTCTGGTTTCTGGACTATACCAGTACACCCTGACGACCAACACAAACTCGCC
TTCAGCTTCGCCAATCGTCAATACACCTTCACACGCTGTCCTTTTGGACTAGCCAACT
CACCGGCTGAGTTCAACATCTTTTTGAACAAAGCATGCCCCAATGCAGCCAGTCGTG
GAACGCTCATCTACGTGGACGATATCTTGATCAGAACTACGACATTGGACGATCACT
TGAAAGAAATCGACCATGTGTTGGACCAACTAACCTCCGCAGGAGCAAAGATTTCT
CTGTCAAAATGCCAGTGGTGCAAGACAAAGGTAAACTATGTAGGCCTCCTAGTNCG
GAGCAGTAGAATTCAAGGAATCACAAACCTCAAAGTGCCTGCAAATGTGTCTGAAC
TACGCAGCTTCCTCGGAGTTTGTAATTACTCAAGACAATTTGTCGAACACTACGCCG
AGTTAGCTAAACCGCTAACCGACCTCCTCAAGAAAGACGAAACTTTCTCATGGGGC
CAAAGCCAAGACAATGCCATGAAAGCCCTGGTAGACCAGCTCTGCCAGGCCCCATG
CTTGGCGTATCCGGACCCCAACAAACCCTTTCACCTCGAGGTAGGCTTCTCAGAGCA
AAGCCTCAGCGCAGGTCTCTACCAAATTCATGATACCGACAAACGAGTCATTGCTTA
TGCCAGCAAGCCGCTGACAGACCCCGAGAGGAAATTCTCGGATTGTGAGAAAGCAC
TCCTAGGCACAGTTTGGGCGGTGAAACATTTCTCAAACTACGTAGGTGGACAAAAA
GTGGTGGTGAATACCAACCACCAGCCAGTAACATTCCTAGCCAGTCAGAGACTCAG
AGACGGGGTGGTCACTAACGCCAGAGTGGCATCTTGGCTCATGGCACTCCAAAGCTT
TGACATAGAAGTGCGCTACGCACAGAACCACAAGTCCCACCTAGGAATGGAACTGG
CCGCATGTCAAAGCTGCGAAAAGGACATACCGGCAGGAACTGCACCATGCGCAGAA
CCCCTACCAGTGTCGAACCAACACCATTACTTTGACCCAAACACATGCCAGGACATG
CCAACCGTCTACGTTGATGGTTGCTCCTTTCACCATGAGGCTACAGTCCGCGCCGGA
GCTGGAACCGTATGGCTCAACAACGACCCATGTGAGCCCGAAAAATTTCAACTAGG
CGCACAAACCTCACAGTACGCTGAAATTTCCGCAGTGCTGATCGCACTGAACCAAG
CAGCTGCTCATGACGTAGACACGTTCGTCATTTGCACTGACTCGAACTATGCCCGTC
TTAGCTTTTCATGCCACCTCCCATTATGGAAAAGCAACGGCTTCCTCACAGCTGCGA
AAAAGTCTGTTAAACATAAGGAACTGTTTATAGCATGTGACCGACTTGTCGAAACCA
AAAACATGAAAGTTTACTGGAAAAAGGTCCGAGGCCACTCACGTGAACCAGGACCT

GACAAAGAACTGAATGACAGGGCAGACGCCCTGGCTAAGCAAGGCGCTTTAGACGG
TACCCCCTGGCGTTTCCAACCAAAACCGCCAGACGTGATACCCCTTCCCACCCAACC
CACCCTTTCTGTCTGTGCCGTCACTCGCTCCCGCGCAGTACCTACTTACGACCGCCGA
GACACAGTCTCCGCCTCCGTAAACCTGGCTCTCACTGACTCTGACTTAACGTCACTT
CAGTCTTCCGACGCCGCTATTAGCCGCATGGTCACCTTCCTGTCCGACCCCTTGCTCC
ACCCAATCTCGGCCGATGATCTGGCCTCCGACCAGGCCCTCAATCTTTTGTGGCGCG
TCCGCGACAAACTCCTCCTACGTCATGGCCTGTTGCTATACGCCTCTGACTCCCTCAC
TTCACCCGCGTTCGTAGTGCCTCAGAGCCTCAGGGGGGTTATGCTCACACACGCACA
TGACACACCAGTCGCAGGCCACAGGGGGGTACAAGCCACATACCAGGCCCTACAGC
AGGTGGCCTATTGGCCACACATGAGAAAAGATGTACAAGAATACACAAAAGGGTGC
CTGGTTTGCTGTCAGTTTCAGCCCTCTAGACCGCTCCACCGTGCACCTCTGCACAAA
AGAGGCATTACCTTCCCCTGGTCAGACTTACAAATAGACTGGGTCGGACTGCTAACG
AAATCCGCAAGAGGAAACAAGTACATCTTGATGGTCACCTGTGCTTTCACCAAATGG
GTAGAGGGCTTACCAGCCCCGAATGACACGGCTCAAACCACGGCCTTCCTGCTCATG
AACCATGTGTTCTCTCGCTTCGGCTTGCCGTGCAAAATCGACTCTGATCAAGGTACG
CACTTCACAGCCGCAGTAATGTCCGAATTGTGGCAGACCCTTGGCGTAAAAGCTACC
TTCCACATCAGTCACCACCCAATGGCCTCTGGGCAGGTGGAAAGAACCAACCGGAC
TGTGATTAACATACTGAAAAAGTATGTGTCAGCTAATCAGAGAGATTGGGACGTGA
AACTTCCCCTAGTCCTCATGGCTATCAGAGCCACCCCCAACAAATCAACCGGGATTT
CCCCCTTTGAGATGATGACAGGAAGACAAATGACTCTGCCCCTGCATCTATTGTACC
AGCCAGGTGAAGCCAGCATTGCCACAGCTTACACCACACACCAGTACATGACAGAC
TTGCAGAGACATTTAAAAGCAACCTTCGCATTTGCCCAAGAGAAACTAGAGAAAAG
TGCAGAAGGACGTAAGGCCTACTATGACCAGAAAGCGTCACACGAGGAACTCCAAG
TAGGCGACAAGGTCTGGTACTATATCTTTGCACAACCTGCAGGACAATCCCAGGAAC
GCACGGCCAGGCTAATCCGGAAGTTCCTACCCAGACGGTCGGGGCCGTACGAGATT
GTGAACAAACTATCCTCGGTGGTATATCAAATAAAGATTACCAAAGGACACAAAGA
GCCAACCCTCAAGTGGGTGCACCGAAACCAGATCAAACCTCACAAAAACCCCATGG
GACAAGTCGGGAACACTGAGGCGCCCCATTAA
> Loki-Bpe consensus sequence Pol protein
MADPTAADAQLSDDLLSQAETALFHATQPVVPGYWEKRKPATIKEHEIELKTLASSVPK
ADAEDDEWTIWLSHMSVTYHLLSVRNKTLILRLKSSADARKAELSALQDQIARLRTELE
QAGVTNLCGLSTTKNLGTGKETPPNASPTSTSPVLHAGDYISRARRQRRPSRDSFLCEFT
SFEESPSPRLVDGVTLETTRVDDARLAIWTDRSAISLELYKMLNGMKQNIPFMPKTSRFP
TDHKSGNVFTAVGFCSLSVTLKNRKLNHVFVVVPNLSHPLYVGGDILVRLDAQIDTINN
VIWSLIGVSENACTPSQENLKSGQTIPEVCQVTNQCPVVIPPNTATPLTVVACPGQQTTN
NQAFFQPSHLFSELGLSVEAMPLLMVKNRMACLVVVNNTTQEVHIPRATPLGWLIDTKF
HDFELKLPVMGKLPASFLEDPDRKVFYTKPLRAIALHSMLVPDVEDICRVDLTDQKEMV
IQTVSVISLEPGKPDETPDATSPTPQADSSGSNPTVEEQLNQVIAEADGLNSDDQREKLRE
ILMKYKAFFSTGPTDCGLTNIHTVRIQTVPNAPPTYVRQYKIPLASYEPIQEIIDDLLTKGV
IRPCNSTYSAPIWPVLKPNGKWRLTVDYRQLNTRVPLSRWPMTQLGQDLPKVRDAKYF
STLDIASGFWTIPVHPDDQHKLAFSFANRQYTFTRCPFGLANSPAEFNIFLNKACPNAAS
RGTLIYVDDILIRTTTLDDHLKEIDHVLDQLTSAGAKISLSKCQWCKTKVNYVGLLXRSS
RIQGITNLKVPANVSELRSFLGVCNYSRQFVEHYAELAKPLTDLLKKDETFSWGQSQDN
AMKALVDQLCQAPCLAYPDPNKPFHLEVGFSEQSLSAGLYQIHDTDKRVIAYASKPLTD
PERKFSDCEKALLGTVWAVKHFSNYVGGQKVVVNTNHQPVTFLASQRLRDGVVTNAR
VASWLMALQSFDIEVRYAQNHKSHLGMELAACQSCEKDIPAGTAPCAEPLPVSNQHHY

FDPNTCQDMPTVYVDGCSFHHEATVRAGAGTVWLNNDPCEPEKFQLGAQTSQYAEISA
VLIALNQAAAHDVDTFVICTDSNYARLSFSCHLPLWKSNGFLTAAKKSVKHKELFIACD
RLVETKNMKVYWKKVRGHSREPGPDKELNDRADALAKQGALDGTPWRFQPKPPDVIP
LPTQPTLSVCAVTRSRAVPTYDRRDTVSASVNLALTDSDLTSLQSSDAAISRMVTFLSDP
LLHPISADDLASDQALNLLWRVRDKLLLRHGLLLYASDSLTSPAFVVPQSLRGVMLTHA
HDTPVAGHRGVQATYQALQQVAYWPHMRKDVQEYTKGCLVCCQFQPSRPLHRAPLH
KRGITFPWSDLQIDWVGLLTKSARGNKYILMVTCAFTKWVEGLPAPNDTAQTTAFLLM
NHVFSRFGLPCKIDSDQGTHFTAAVMSELWQTLGVKATFHISHHPMASGQVERTNRTVI
NILKKYVSANQRDWDVKLPLVLMAIRATPNKSTGISPFEMMTGRQMTLPLHLLYQPGE
ASIATAYTTHQYMTDLQRHLKATFAFAQEKLEKSAEGRKAYYDQKASHEELQVGDKV
WYYIFAQPAGQSQERTARLIRKFLPRRSGPYEIVNKLSSVVYQIKITKGHKEPTLKWVHR
NQIKPHKNPMGQVGNTEAPH*
> Loki-Bpe consensus sequence *env* gene
TTGTGTGTTGACAGGAACATGGAACCGCGTTTCCTGATCCTACTGGTCTGCGCGTCC
AGCCTAGTAAAAACAGAAGTTGTAGAGCCCGGACCGGCCTCAGGAATCGTCCTTAA
GGACCAACCTGGATTACTCATAACGAACGTCCGACTCCACACTCAAAAGGTCGTCGT
GAGACTCAATCCCCGCGACGTATGCAGGAAAAACGTCCCTATCACGGCGAGCCTCA
CGGGTTGGGCTGGCACTCAGTGGAACCTAGAGGCCATCCAACACGCGGAATCAGAC
ATAACACATATGTTAGAGCAACTCGAGAAATTCACAACAACCCAGGCAGAAATTAG
AGGTATGCCGCACCCACGCTCTAAACGCTTTATAGGAGGCATCTTAGCCGCAGCTTC
CACCGTAGGAACCATATTTAACCTCGGGTTAACCACGATAAACACGGTCAGCATAG
CCTCCCTCAAGCGCCACGTCAGCGAATTACAAGGAGAATTACCGGTGATCAAAACC
CAGCTCGATCAACAGCAGAAAGATCTAGTGACTTTCGGTGAAACCCTCCGCGGTAC
AGTGCTGGTAGTAAACACCCATAGCATAGCCCTAAATAAAACAATACAGGCTGTCA
ACTCCCTCATGGCCGTAGTTCACGTGGATTATGCGCACCTACAGTTAGTCACCATGC
TGTANAACGACATGCTTCGCGACATAGGGTCCTCGGTAGATAGCCTAGCCATGGGTC
GAATCCCACCATACTTGGTGCCCCTGACCCTGGTACAAGACTTACTATCCACTGCGA
CCCGCGACATAGTCACGCCTCTCCAGGCCCACCTAGCATACACGCTCGGAAGCGCC
ATCCCTATCCTCGTGAACCCCGCAGGCCGCGAAATCGCTTTCTTAATTAACCTGCCA
ATCGTAGCTTCCGATAACATTTACCGCCTTAAAGACATAAATGTCGGGTCGTGGCAG
GGAGAAATGCACGTTAGAATTCAAACCCCAAGCCTGGTCGCATATCACGACAGTAA
CCCAGACCTGTACTTGACACCAAACCTACGTATGTGCTCTCTTATGAAAGACATTCA
TTATCTCTGCCCCAGCAAACCTTTTGTCAGAGATAGCACCGACGGCATCTGCGGGCT
CAAACCCATGACACCGGGTAACAGATGCCCGACCACAATCACGGCTCGTGGGAAAG
TCCTAGAAACTCAGGCCGAAATCGTGGGGGGGCGCTGGCTAATAAACACTCNGGCA
ACCCTCACGTACGACAAACACGACACCTCGTCCAGAATCATACTGCCAGAACAGAC
AATCTGGGTGGATATCCCAACTAACGCCATCCTCCATATTGGGGACATCGCGCTATT
TCCGCTCAGTCCGAACCAATATGAGGAAAACTTTGAAGTCCCAGAGTTCTTTGCGGA
AAGAACTCTGAAACTCCAACCTGCAACGATTAATCAAGTGTCCTACGAAGGCACCC
AAATAATCGATATCACACCCGTCGACAACCTCCTGACAGAAATTACCAATCGACCTC
TAACGCCGATCCAACCCGTGTCTTACTCTTGGTCAGGCCCTGACTCACTCCTTGCCAT
CGCGATAGGAATTGGGTATATTTGCGCTATAATTGCCACCCTTATCTTAAACAAAAG
ATTGAACAGAATTATAGGGCAATTCAGGCAGCACAGAAGAACTAGGGCAGTCAAAC
GCCAGTTGACAACAGAGGGCACAAACCGAATCGACGAACGCGTGTCGTACTCAGTC
CGGGACGCAGACTATGAGGATCTCCAAACAACCAATGAAGGGTTAGTTGTGATCCC
GGCCCCAGGCGTCTAG

> Loki-Bpe consensus sequence Env protein
LCVDRNMEPRFLILLVCASSLVKTEVVEPGPASGIVLKDQPGLLITNVRLHTQKVVVRLN
PRDVCRKNVPITASLTGWAGTQWNLEAIQHAESDITHMLEQLEKFTTTQAEIRGMPHPR
SKRFIGGILAAASTVGTIFNLGLTTINTVSIASLKRHVSELQGELPVIKTQLDQQQKDLVTF
GETLRGTVLVVNTHSIALNKTIQAVNSLMAVVHVDYAHLQLVTMLXNDMLRDIGSSVD
SLAMGRIPPYLVPLTLVQDLLSTATRDIVTPLQAHLAYTLGSAIPILVNPAGREIAFLINLPI
VASDNIYRLKDINVGSWQGEMHVRIQTPSLVAYHDSNPDLYLTPNLRMCSLMKDIHYL
CPSKPFVRDSTDGICGLKPMTPGNRCPTTITARGKVLETQAEIVGGRWLINTXATLTYDK
HDTSSRIILPEQTIWVDIPTNAILHIGDIALFPLSPNQYEENFEVPEFFAERTLKLQPATINQ
VSYEGTQIIDITPVDNLLTEITNRPLTPIQPVSYSWSGPDSLLAIAIGIGYICAIIATLILNKRL
NRIIGQFRQHRRTRAVKRQLTTEGTNRIDERVSYSVRDADYEDLQTTNEGLVVIPAPGV*
>Loki-Sme consensus sequence
TCATAGACTATGTTACGAATAATTCCTTATCTATCTAAAGTCACAAAAGGGCCTCCA
CTGGAGTCTTCGAGTCTGCTCACACTCACCATCTGTGTTGACTTAAATCAACATTTTT
ATGACTCTCTACCCTGTTTTCCGACACGGCAATCTGTCACAGATGCTCACCACTTTCC
ATCTTAAAGAACCTTCTCACCTTGTGGATTTTAAAGGACAACATCATCAAAGACTCT
AATACACAATCCCCCGTCTGAACCCTGGATGAAACCAGTTCCTCCTTTTTGGGTCAT
AAATTTTAATGACTTTAAGATAACACTTTCACCTGATTTACATGACCCCCTCTAACCC
ACAGGAATTCCTAACTTTTGGTTGAACCCAATAAAAAGAAAGGGACCAGATGTCTTC
ATGATTCAAATGTATATCAATTAACATGTATTTTCCATTTTTAGGTCAAGAATCTTTC
TACACACATGTTTATAAACTTCCTCGTCCAAACCGTATTCAAAACGGGTAATCTACC
GTAAAGTTACAAAAGAATCATGTCAAGGTATAATCCCCTGCTGCAACTCAAGGATTC
TCATGGACCGCTACCGCCATCTAGCGGTTCCTTTAAGTCAATGTTTTCGTAAATAAA
CAGTACCTTAATCTTTGTCGTTTAAATGTTGCATTCGGGATTAATACTATGTACCTTT
TCAGTTTATTGTCTCTATATGTTGTCTCATTGTTTATTGAAACTATCAATGGAAGTTG
GCGAACTCTACATATGTCACAATTCAATGATAGTCCGTGTTTGGATTCATTTTAAAG
GTTATAACGATACATTGTTATTTTTATCCAAAATTAGTAGTAACACTCCAGTATGGTG
TAGTGTTAAAACGGCTAATTCAGATTTTATTAACTTCTAACGAAGTACGAGTTTTTGT
TTTTTTCGACTCGTTCCTCCCCTAAAACTCGCGTCAGCGGAGGGGCGTGTATCATCTT
AGGTGCTTTTTAAACTGAAACGCGCACGTTTCTCGGCCCTTCTTTTGCCCCTTCTTTT
AGCCTCTCTTTATTACCTGTTGGATCCTTCACGTATTTTTCCTTGTTTTATACTCCGTC
GGAACTTTGTAGCCAGGACGGAACTTTGAAAACAGTTTTTCTTCTGTTTTCGAACTTT
AATTTTTCGAGTTATCATTTTTCTTTTAAGGACTTGAACGCTTCCAGTTTTTCCTCAAT
ATTCACCAGCTAAGAAACATTCTATCGATGCTCGGGAAAGAGCAATTTAATAGAATT
GTCTCTATCGATTTTTGGACATTTTTCCTGTTTTTCAAAAGTGAACGATAAGCTGATC
GTCAAACTTTGGACAATATCCTGAGTTCCGTGGGACAGAGCGGACGCGCCGGCACA
TTTCTGAAGACCCGTTAACAACAGAACACGACCAGAATTATTGCGGTAATTCAGGAC
TGCAGTGATCGGAGACGGAACTTCCAATACCCGCTTCCTTACTCCAGTACCAACTCG
GCGAGTTCGTTGGGTCTACAACAAGTAAGACGCTCTACAGCTTGTTGTCAATACAGA
GAAATGCGTGTCATGCGTGCTCACGATTAAGTCCCACTCTCACTTCAGTTTAATATA
ATCCCAACTCCCCTTTTTCCCTCTTTTGCCGTTCAATAATTTTTCCGAATTTAATCAAT
AATCTGGTTAGACAGACGTTAAACACTGTGCCATACTGGATTAGTTACAAATAACAT
TTGTTATCTCTGGGTGCTCTGCATATCTGACGCATTGGGTATCGTTCCGTTATTCTGA
TTTCTTTGTTCATTATTATTCATTAAAGTAATCACGTATAAAATTCAGAAGTGTGTTG
TATTGATTATTGGTGGTTGAACGTTTCGTGTAACTCCGGACCAAAGAATTGGAGACT
GATTTAGAGAAATTAATAAAGAAATACACTTATCGATTTTCATAACTTTCCATAACT

TATTAATCATAATCAATAAGTTTTGTTCCTGGTGAGCCAGGTGGTGCCCTTCTTCTTA
CGAGGTTTGTTAATCCATTTAACGTAATATTAATGATTATTTTCCACAACATATTGGT
GCCCCGTCGCGAGGCGAAACTTGCATCAGAAACCGATAATTTCACGAATACAGCAC
TCTAAAAAGCTCATACCTACCACATTATAAAACCATTAAATCCACTGTCCGCTGATG
ACCAAAACGGCTTTTGTATTTTTGTTTAAAACTCGAAGCAGTGTCCTTAATAAATGTT
TAAAGATTGGGGGATACCCTTCTGGCTCAGCTGAGTGACATCATCAAAGACGATGA
CGCATTAAAATCAGCTGATTGGAATTCCCTGTGTTATTGTTTTGAATTGTGAACTGAT
ATCAGAAGCATTGTTGTTGATCGTGTCGCCGCACACATTTTTCCTGCAGAAGTGATTT
GCCGCTGAGCAATATTGATGTGATATAGGCCTTTCTAAACCGCTTTTGTTTGCGAGA
TTCCGAATAAAAGTCTGAGTTTAGCCGTGGTATACGAAATAAGGCACGAGTTTCTGC
CTTACTGAGTCACTACACGGAAACCCATTATTAGTTTGCAAAGTAGCTAATTGTGAA
CTTTTAAACGCAATTCGTTGTGGGGTTATTAGGACGTTCCTCTCTAAAAGAAGGATT
CGTCGATGTTTTGAACAAGTTGATCGCTTGTGAGTGAGTGTGCACGTTGGTAGCTAA
GGTTAGCCACGGAGCTAATCCTTCCTCCTCGTGGCCTCCGGCTCTTGGAGCAACTTG
GAGTCATTATCGCCCTAGAGGTGGGTGTAGTGATTGATTGCCCGTGATTACGGTCTC
GTAGGAGCTAAAGTGTTAGCCACAGAGCTAACGGAACTCTTCTTCGGGGTTTTTGTG
TTTTTGACTTCTGGCTCTTTGAGTAACTACCGCCCCCTACCGGTGGGGAGTAGTAATT
GATCGCTTGTAACTGTTGCCACGTGGTAACTTAGGTGTTAGCCACAGAGCCAACAGG
AGGCCTCGCCCTCGTCACTCCCGGCTCTTTGGGTCACTACCGTCCCCTAGCGGTGAG
GAGTAGTAGTGGATCGTTCCCTTGATACAGTTGGAGCAAGAGACTCCATCCTGTGCC
TGTTTTGAGTAACTGCGGTTATTTAAAAAGTGACTAAAGACTTTAATTAATTCATGA
CCCTGACGAAACAGACTTGTAAAGCTCCTTTCCCTCGTTGCATGATCTCTTGTTAACT
CTGTTTGAAATTTTAATGATTTATTTTCCCTGCGAAGGCAAACATAACTTTAATTAAA
ATTTTCTGACAGGACAGACAAATAGAGCATTAAGCTAAATCAAAGTAGTTTCCACTT
TTTGGTTAAGATTGATTTGTTAGTTTGAATAAAATTAATTAACATCAATCCTAAATTT
CTCCAAACTTTAGTCTAACATAAATAATTAGGGCTACAGCCTCGTGTTAGATTTAAA
ACGTCCTAGGCAATATCACATTCAAATAACAATCTAATTTGCTGGGTGATGATGCTC
ATTGTTGTGTTCTGAGAATAATTTAATTCCAGTCTATTAGGTGACCATCAACACCCCA
GTGAACTTAGACTTGTATTTTATTCCCATAAATTTTTCAAATCCATGTTTTTCCTTCCT
TATCAAAGGTGCCTAATAATCAGTCAACCATAACTAGCTGTCGTTTAGAACTATTAG
CTCAGACTAATTAACTAAACCAGGACTATTAGTGCACTTTTAAATTGTCAACCAATT
TTGATTCCTTCTATTCAATTTAATTCGGTCAACACTAGGCTTACACCCTAAGTTTGGT
TTAATTCGGTTTAAATCTTTAATTTTCCAGCTTTAATCCAATTAGATTTAGATTGGAT
TTAGGATTTAAAACCTTTAAACTAAACTATTTAAATTGATTTCCAACTTTAATTAATT
TCATAATTAAGACCTTTCGATTTTTCTCTTCATTAATTTAATTTTAATTTTACTTTTTA
ATTTTAATTTCAATTTTAATTGTAAATTCTAATTTCAATTTTAATTGTAAATTTTAATT
TTAATTTAGATTTTTGATTTAGATTTTTTTTTATTTTAATTAAATTTTAATTTTAATTT
TAATTTCGGATTTCGAATTGCGAATAATAATTTTAATTTTAATTTCAATTTTTAATTTT
AATTTCAATTTTAAATTGATTTTAATTTTTAATTTTAAATCTTGATTTTAATTCCGATT
TTAATTTTGTTTAATTTTAATTTTAATTTTAATTTTTAACATTAATTTTAAACCTTGAA
TTTTTTTTTCTTTCTTTTTCTTTCCTAGTTCTTTTGCCTCTCCTTTCACTCCTTTCCTTAA
TCTGGTTGACAGTTGATAAGACCCAAACATGGAGGCCGCCGATAGCCCGCCACCTG
CAGGCTTCGGTGACCCATTAGATCACTTTGATTCGGTTCTGACAGACCTAGTTGGAC
CACCTCTGCGTGAGTTCGTGACTGAAATTGAGACGACTGACCCCTCTCGACTTCAAG
AAAAGCTTTCAGACTTGGTTGCCAAGGCTATGTTTGGGCAGGTTAAGGAGAAAGAC
AAACATAAGATTCTTTCACAAATGGCTGTAATCATGGCAGCCCAAGAGAGACATAC

CCAGAAGTTACTAAACCAGCAGGAGAAGAGCAATAAAGCCCTTCTGCTTATGATTG
AACAGAAAGACGAACTCCTCGTCCAGGCCGATGAGAACATCGCAAAGCTTGAGGAA
GCCAAAGCCCTCCTACAAGCAGAAAAGGAGAATGAGGAAGATAATAGGGAAGATG
AGTTGAGAGCAACTCTTCATGAGAAAGAAAATGAGTTGCAGTCAGTTCATAAACAG
CTCACTGCAGCTGCCAAACGGCTCGACCAGTCTGATGTCCTCCTGATTAAGGCTCAA
GATAAGTTCCACGACTTAAATGCTACCATCAAAGACCTGACCTCACAAATAGAAAG
GCGGGAACAACACGAATGTGACCTCCAAGCCAAGTTGGACGACGTTTCATTGCGAG
TAAGCCTGTTTAAACAGCAGCACGCAATGAAAGAGGATGAATTAGCGAGGGTCAGA
AGACAACTTCAATCTGCTCTGGACCTAGATCCAGGTCGACAACGACCTGGTAATGAA
TCACCCTTCGAGAGAAAGGATGTTCGTAGTCCAAGAGAGAATCCTCCTCGATCTGGT
CTTGTGATTACCCCCGATGATTCCTTGTGGTCTCGAAATCTCGAGCACAGCCCTTACA
GAAGACTCCGATCTCCTATTTCTGAGGAGCGCCAATCTCTGCCAAGCCCCTCATACA
GAGATTTGGACAAAATAGCTCGTAACATCACCCGTTTCGAGCCGGGTAACAAGGGT
GGTTTTGAGATTAGCCTATATCTCAAAGACCTTGAGTACCACCTGCGAAAGTTTTCA
GGTGCTACCACGGACGACAAAATCTATTTGATCAGGATGACGTCCAACAGGACTAT
CGGCAGTTTTATCGAACGTCAGCCTGACAGGGTTAGGAGTGACTACAAGTTATTGTG
CCAAAGGTTGATAGACGAATATTCTGATTACATCCACCCCACAGGGTTGGCGGCTGC
TCTTGTGATCAAACAAGAAAGAGATGAGTCTCCCCAGGTGTATTACAACCGAATACT
TCAAGCTTACTTCGGTCCTAAAAATGAAATTGGGATGGAAGACGATGTGAACTTTAA
GGTCCTTTTCATTCAAAACCTCCATCCAACAACTAGCCATCATCTCGGTGTGGCGGC
TTGTCCACGCACCCTCACCAGTCGACAACTGCGTGATCTGGCCGTGAGAGGCTTCGC
CAAATACAAACAGGTCTCTGTCAAAAAGCCCGACACCAACGTCATATTCAGCATTG
ACACAAACAACATACCCCTCGAGCTTGAAGGAGCCACCGAAAGTATTCGTGAGAAT
CCTTCAAGGCCAAAGGAAAGGCCATTTAGAGACCACAAACCTTTCCGACCTTCACC
GGCCAGTGAAAGCCGGAGGGAACAACCACATTCTGGGTTTCGTGGTAATTATCACC
AACCTGAATCAAGTCGTCACTTCTCTGACATTCGACGACCCAAGTTTCCACCCAGGC
AGTGGAACAATTGGGACACCCGTGACGGTCAGGCCCAAGAGTGGCATCAGAACTAC
CGCCCTCGGCCGAGAGACCAACCACGTAACCCAAGGGTCCGAAACGAGTTTCCCAA
CCAACAGATGCGGGAAAACGAGCGATCAAACCCTCCACAAACATATGAGAAGGGC
AAGGAAGAAGCTATCCATCTTTCAGAGAAAGAGGTAGAAGCTATAAAAAGATTGCT
AAAGCAGCAATCAAACAAAGGTAAGAAGGCCGACCTACTAGCAGTAAGTTTGGTCA
AACCTGAAACATCCCCTCCAGACCAGGAAACGAACGTGGAGGAGGAAGAACTCTAT
CGAAGCGATGATAGCAGCCAGGATATGGTAGAAACTTCAATTGATACTGAACCTCT
AAAATCAGAGGATAAATCAGTTAGCGCAGTCTTGGTAGTCAGTGCTAATTCCACAG
ACACCTCCAACGGTCAAACTCTTGAAGTCATCCAAGGTGAACCACCTTTTCAACAAT
TCTTGTGTCATCTGATCGAAAGGGGGGTAGCACGAAAATTCTACCTGTCGACCATTC
TTGAGGATGAGGTCGAACATGAAGCCCTCCTGGACACAGCAGCCGATGTCACCTTG
ATGTCATTTGCCCTTTTCAATAAGCTACGTCAGGTAGCTCATCAGTCCAACAGGGAC
CTGAAACTCCAGCCTTGCGTAATGAACATTCAACCTTACTCGAATGGTAATACGATT
ATTAAGAAGCTGGCGCTGATAAAACTGACCCTGGGACCAATGACAGTCATTCATCC
AGTTTATGTGTCTCCACTCAGTTCAATCCCTCTTCTTGTCGGCAAGGATCTGCTTAAC
CGCTTCGAACCCCTCATCGACTTTAAACGCTCGAAGATTTGGGCTCAGGTTCGCAGT
CCTCTGCCTATATCGCTCCCTTCTCACGACACAGCTCACTGTCTCGTCCTTGAAGGTA
AAATAGACACACTTGATGGTCAGCCAACACAGTTTACGACCGCCGAAGCGGACGAC
GACTTTGGCACAACCACAGATGCCAATTCGGGAAAAGGATGGTACCCATCCGACAC
TCTCGCTCAGCAAGATAACTTCCTCTGTGCTTTTTCAAACCCCTCGGGGACAGACGA

ATACTGCCCGATGGTAAACCAAGGCATCCATCTCAACGACACACATGTCGAAGACG
TGGCCTTGGCCCTCTGGGCTGATCAAACCGCCATTAGTCAAGAGCTGTTCAACTCTT
TGACAAAGGCCAACCCTGATATCTTGTTCTACTCGGGTGCTGAACGCTTTCCTCTCCA
ACCCCAGTCACGAGAGACCATGTCTGCTCAAGGTGCCTGTACCTTGAAGGTGCGCTG
GAATCACCGAGAGCTACATCATAACTTCTTGGTCGTTCCCAACCTCCCTCATCCAGC
CTACATAGGCAGCGACCTGCTAGTCAGGCTGGACGTGCAAATCGACACCATAAACA
ATCTGCTGTGGTCCCTCAAAACTGTTGGAGACAACACGGCTCCTGATCTGGATAACC
TCAAATCGGGTCAATCCATTCCAGAAGCATGTCAGGTGGCAAACGCACTTCACGTGG
TTGTACCAGCAAACACGACAAACATGTCTATAAGGTTGAACATACGACCGGGGCAA
CGTCGACCCCACTCACAGGTGTTCTTTCAACCATCGTTGCTTTTCTTTCACTTGGGCC
TTACCATTCAGGCCACCCCCTTGATGGATCTAGCCTCCAGATCCACACACCTGCTGG
TACAAAACGAGACCAGGCAGGATATACTGATTCCCAAGTCCACCCCTTTGGGCTGG
CTGATCAGTACAGAATTTCATGATTTTGAGCTGAGAATTCCGGTGATAGGACAAATA
CCTCCACTGTTACTCTCCGGCAGCAACGACGAAAACGTGGTTTACACCAAACCAGCA
GGAGCTATAACTCTCTTCCCTGCTGTAACACTAACCCAAGACACGGTCTGCCGGGTC
GATCTCACTGACAGTACCCAAATGGTTCTCCAAACAGTCCAGGTCATGTCAGTCGAT
TCATCACCCGATCCAGACACGTATCTTGAGGCAGAGCCGTCACGATCAACTTCAGAC
ACACCTGCATCTACCAACTTGGACTTCAAAACCCAAGTGGATCAGGTGGTGATAGA
AGCAGACGCTCTACAGTCAGACGAAGAACGCGAACAACTTCGCAACATTCTGCTCA
AGTATGAAAAGTCGTTTGCACAAGATTCACTTGATTGTGGTCTTACCGACATTCACT
CAGTGCGTATTCCGACGCCACCCGCGGCTCCTCCCACGTTCGTACGCCAATACAAAA
TACCCTTGGCATCGTACGAACCCGTGCAGGAGATTATTGATGACCTTCTTCACAGAG
GCATCATTAGACCATGTAATAGTACCTACTCAGCGCCTGTGTGGCCGGTACTTAAAC
CAAACGGCAAATGGCGACTAACTATCGACTACCGGAAACTGAACCAACAAGTTCCC
TTGTCCAGGTGGCCGATGACACAACTAAACCAGGAACTTCCAAAGGTTAAAGACGC
CAAATACTTCTCTTCCATGGACATTGCGTCCGGATTCTGGACTATTCCTGTCCACATA
GCAGACCAACATAAGCTGGCGTTCACTTTTGCAGGTCGACAGTACACGTTCACCAGA
TGTCCGTTTGGATATGCCAACTCTCCGGCAGAATTCAACATCTTCTTGAACAAATCG
TGCCCTGACGCCAGCGAGCGTGGCACTCTCATCTATGTTGATGATATCCTTATTCGG
ACTCCAACCATGAAGGGTCATATAGACGAGATAGAACATGTGCTCGACCAGCTCAC
CACTGCCGGAGCCAAAATATCTCTGTCCAAATGTCAATGGTGTAAGACAAAGGTAA
ACTATGTGGGACTCCTAGTCGGCTCGAATGGCGTAGAACCACAAACCAATAGGGTG
CAAGGAGTTATGAACATCAAAACTCCAACTAACATCTCTGAGCTACGCAGCTTCTTA
GGTGTTTGCAACTATTCACGCCAATTCATACCCGACTTTGCAGACATTGCGAAACCA
CTCACTGATCTTCTAAAGAAGGATTCCCTTTTCACCTTTGGCGAACCTCAAAAAGAG
GCCATGCAATCTCTGAAACAGCTATTATGCACGGCTCCTTGTCTGGCGTACCCGGAC
TACCAAAAGGAATTCCATCTGGAAGTTGGATTCTCCAGTCACTGTCTAAGTGCAGGA
CTTTATCAAGTTCACGACTCAGACAAACGGGTAGTTGCCTACGCCAGCAAGACCCTA
GCCGCACCTGAGATCAAATACACTGATTGCGAAAAAGCTCTGTTGGCCACAGTCTGG
GCCACAAAGCACTTCTCCAACTATCTTGGAGGCCAAAGGGTGATAGTTGAGACTAAT
CACCAACCAGTTACTTTCCTTAACAGCCAACGGATTAGGGACGGAGTTGTGACGAAC
TCCCGAGTCGCGTCTTGGTTGATGGCCCTGCAAAGTTTTGACATGGAAGTGCGCTAT
GCTCAGAATCGCAAAACGCCTCTCGGCATGGAACTTGCAGCATGCCAACATTGTTCC
GAAGAATCGCCAATTGACTCAATTGTGCTAACCGACAGTGTGCTCCCTGAACCTTCT
AACCACCATTACTTTGACCCCAATGTCTGTGCAGACATGGTCACAGCATATGTGGAT
GGGTGTTCATTCCACCATGAAACTCAAGTAAGGTCGGGAGCAGGAGTTCTGTGGGTC

```
CACGACACTCCATGTGAACCACAATGTTTTCAACTTGGGTCACAATCGTCCCAATAC
GCCGAGATCGCGGCGATCCTGATAACATTGCAATTAGCAAAAGAAAACGATGTCAA
AAATCTGACAATATGCACTGACTCGAACTACGCTCGTTTGAGTTTCACCTCCCACTT
ACCATTGTGGATCAAAAATGGCTTCATGACTTCAAACCGTAAACCGGTCAAACACA
AAGAACTTTTCATAGCTTGCAAACAACTTGTAGACACGTGCGATATGCAAATCTACT
GGAAAAAGGTCAGAGGTCACTCCCGAGTCCCTGGTGCGGACAAGGATCTGAACGAC
CAGGCAGACAAACTGGCCAAACAGGGGGCCATTATCGGCACACCCTGGCAGTTTGA
CGGAAACTCGGTTCCCACCCAAGCCCTTCCCACTGTTTGTGCTGTTACACGTTCAGG
CACGACAGCGACACCAGATCCTCGTCGACCATCGCCTGTTTCGGCCTCGGTGTCACC
TGCCTTCGCCAGCTCGGATCTCTTGACACTACAAACCGAAGACTCATCTATTCGTGC
TATGACACTATTTCTTTCAGATCCTTCTGCTCACCCAATCTCCAAAGATGATCTTGAT
TCAGATGACACCTTGGCGTTTCTTTTTCGGGCACGTGCTTCTTTCAAAGTCGTCAATG
GGCTGCTTGTTCATGTCTCTGACACACTCACCACACCTGCGTTTGTGGTGCCTCAGAG
CCTAAGGGGGGTGATGCTAACCCATGCCCATGACTCCCCTTGTTCAGGCCACAGAGG
TATCAAAGCCACCCTGAAGTCATTACAGGAAGTCGCCTACTGGCCACACATGCAGA
AAGATGTGTTAGATTACATCAAGGGTTGCTTGGTGTGCTGTCAATTCCAACCAGCAA
GCACTCTTCACCGAGCACCACTACAGCGACGAGGTGTCTCATTCCCCTGGTCTGATC
TTCAAATCGACTGGGTGGGACCCCTCACAAAATCCACCAGAGGCAACAAATACTTT
CTCACGGTGGTATGTGCATTCACCAAATGGGTGGAATGCTTACCTGCACCGAATGAC
ACAGCTCAAACCACCGCGTACCTCTTGATGAATCATATCTTCTCGAGGTTCGGTCTG
CCTGGCAGAGTTAACTCTGACAGAGGCACACACTTCACGGCTGAAATGATGCAGCA
GCTATGGCAGCTGCTGGGTGTGAAAGCCAACTTTCACATCAGTTACCGTCCACAATC
GTCAGGACAAGTGGAACGGGCTAACCGAACAGTCGTGAACATCCTCAGAAAGTATG
TGTCCTCAAACCACAAGGATTGGGACGTTAAACTCCCCCTTGTCTTGATGGCCATCA
GGGCGACACCCCACGAGTCAACCGGACTCTCGCCCTTCGAGTTGATGACGGGAAGG
CAGATGACTCTGCCCCTGCACTTGCTCTATCAACCGGGTGAAGCCAGCATCGCCACG
GCTTACACCGCACACCAGTACATGACTGACCTACACCAACACTTGAGTGCCACTTTC
GCTTTCGCCCAGAAAAATCTGGACAAGAGCGCAGAAGGAAGCAAAACCTACTATGA
CAAAAAGGCATCCCACGATGAACTCCAGGTGGGAGATAAAGCGTGGTTTTACATCT
TTGCTCAAACCATAGACAACGACAACCACAATGCTGGAAAACATGGTCGTAAGTTC
CTGCCCCACTGGGCTGGACCTTACGTCATCACACAGAAACTCTCCCCTGTGGTATAT
CAGCTCAAGATTACCAAAGGACAGAAAGAACCTATGCTGAAATGGGTACATCGCAA
CCAAATTAAGCCACACAAAAACCCCATGGGACTAGTCGGGGACTCCAACGTCCCAG
TTAGGTCATAAAAGCTAGAAAGACCAAAGTTCTAGAGGGAAGAAGAGTTAGGCTAC
AGTACTAATTAGTTCAGAGAAACAGTTAATAACTCTTGCTTGCTCTATCTTTTCTCTA
ATCCTCTACTCTCATTTCAAATTACAGACGACCAGACCTTTGACGTTCACCCGGCCG
AAAAGAATGAATCCGTTGCTGCTCCTCGCGACCTTTTTGGTGATGGGTGTCCCGGTC
TCTCCAGAGATGGTACAGACGGACCCCCGTCCGGGATTGCGCTAACGCCCAATCC
CGGACTACTGATCACCAACTGTCGCCTGCACACACAGAAGGTGTTTGTCAGGTTCCG
CCCCCGTGAGGTACTCCAGAAGCATCTCCAAGCGACTGGAAGCCTCACCAGTTGGG
CTGGAATGACCTGGACCAGGAACGCGATTAAGCACGCTGAGGCAGACGTGGCTCAT
ATGCTCCTTCAACTACAGAAATTCACGGTGTCTCAAGATGAACTACAGGGAAAGAA
CAAAAGGCAGAAACGGTTTATCGGAGCCCTACTTACTGCAGCCACAGCAGTAGGCT
CCCTATTCAGCATCGGACTCTCTTCCGTCAATGCCGTCAGCCTAGCAACCGTCAAAC
GTCACATCACAGAGTTAGACGCAGAAATTCCCGAAATACAGTCACGACTACACCTT
CAACAACAGTCGCTCCAAACCATAGGAGAAACTCTTAAGGGAACGGTCCTGATTGT
```

GAATACTCAGAGTCTCGCTCTGAACCGAACCATCAGGGCTGTAGACTCCCTTCTGAC
AGTAATGCAGGTGGACTACGCTCACACACAACTGGTGGACATGCTAATGACAGACA
TGCTGAGAGAGGTCGGGTCATCAGTAGACAGTCTAGCTACAGGCCAGATTCCCCCTT
TCCTTGTCCCCCTCTCCCTAGTACAGAACCTCTTAGCCAAAGCCACTAAGGAAACCA
TCACACCCCTACAAGCCCATCTAGCGTATACCCTTGGTAGTGCTGTTCCCCTCTACGT
CAACCCTGAAAGCGGAGACGTATCCTTCCTCCTCAACCTTCCCATAATAGCTTCAGA
AAACATTTATCGGCTCAAGAACGTGGTGAATGTGGGATGGTGGCAAGGAGATACCC
ATCTGAAGGTGCAAACACCCACGATGGTGGCCTACCATGACGATAGTCCCGAATTGT
ATTTGGCACCCAACCTGCGAATGTGCACCCTTACCAAGGAAATCCATTACCTGTGTC
CAAGCAAACCATTCGTACGTGACAGTACTAATGGAATTTGTGGACTCAAGCCCATGT
TGGCAGAAACTCGCTGTCCCACAACTGCCACTTCTCGATCCCAGGTCACAGACACCC
AAATTGAAATCGTAGGTGATAGATGGCTAGTAAACACTCCGTTCAAAGTAGCCACTC
TCACTTACGGACAACATGACACTGCTACACGGGTTCCCCTCCCAGAGGAGACGTTAT
GGATTACCGTCCCTGAAAACGCTATCCTCCACATTGACGACATGGCCATCTACTATC
TAAACTCCGAACAATACGAGACTGACATTGAACTTCCCGATTTCTTCAAAAACTACC
AAGTAAAAATCGACCCCAACGCTATCGCTCAAATACAATATGTAGGAACCCAAACT
ATTGACCTAACACCACTTAACGACATCTTACAGATCACTACACCATCAGAAAGACCA
CCCATACAACCGTTAAACTACACATGGTCCACCCCCGACACCTTACTCGCCCTTACC
ATAGGACTCGGGTATGTGTTAACGTTTGCCATAGCCTTCATATGTGCTAGACGTACT
CGAGCCCTCCAGAGGCAACTAAACAAGTGTTCCACTAGTATGTCCAAACTGCGTAA
ACACCACGGCAGGAGTAGAACCACTGAAAGGGATGACGAAGACGAAGATGTCGTCT
ACCAAATCTCTCCTACTCTTATAGAGCTAGAATCCGAAACCTCCGTGCCCATTTAGT
CTAGATCACCGCCACATAGTGGTCCTGGATCCATGAGCACCTCAGACCCTTAAGCTC
ATGTTATATTATTGACTTCATACTTCCAACTTATTAGGCATTGTATGGCCATTCATAA
AATTATTAGATTAGTGACATAATTACTTGTACCCCCATATACTTAGATTTTTGTTTTT
GGTGTTTATTGTTTATTCATGACACGATTTCTCTGTGACACCTGCTGTTGATGCTGTT
TTACTGTTTGCTGTCATCACCTTAATTGGCCTGCNGATTCTCTCTCCAAGGATTCACA
TGCCAAAAGGGGGGACTGTTACGAATAATTCCTTATCTATCTAAAGTCACAAAAGG
GCCTCCACTGGAGTCTTCGAGTCTGCTCACACTCACCATCTGTGTTGACTTAAATCAA
CATTTTTATGACTCTCTACCCTGTTTTCCGACACGGCAATCTGTCACAGATGCTCACC
ACTTTCCATCTTAAAGAACCTTCTCACCTTGTGGATTTTAAAGGACAACATCATCAA
AGACTCTAATACACAATCCCCCGTCTGAACCCTGGATGAAACCAGTTCCTCCTTTTT
GGGTCATAAATTTTAATGACTTTAAGATAACACTTTCACCTGATTTACATGACCCCCT
CTAACCCACAGGAATTCCTAACTTTTGGTTGAACCCAATAAAAAGAAAGGGACCAG
ATGTCTTCATGATTCAAATGTATATCAATTAACATGTATTTTCCATTTTTAGGTCAAG
AATCTTTCTACACACATGTTTATAAACTTCCTCGTCCAAACCGTATTCAAAACGGGT
AATCTACCGTAAAGTTACAAAAGAATCATGTCAAGGTATAATCCCCTGCTGCAACTC
AAGGATTCTCATGGACCGCTACCGCCATCTAGCGGTTCCTTTAAGTCAATGTTTTCGT
TAATAAACAGTACCTTAATCTTTGTCGTTTAAATGTTGCATTCGGGATTAATACTATG
TACCTTTTCAGTTTATTGTCTCTATATGTTGTCTCATTGTTTATTGAAACTATCAATGG
AAGTTGGCGAACTCTACATATGTCACAATTCAATGATAGTCCGTGTTTGGATTCATTT
TAAAGGTTATAACGATACATTGTTATTTTTATCCAAAATTAGTAGTAACACTCCAGT
ATGGTGTAGTGTTAAAACGGCTAATTCAGATTTTATTAACTTCTAACGAAGTACGAG
TTTTTGTTTTTTCGACTCGTTCCTCCCCTAAAACTCGCGTCAGCGGAGGGGCGTGTAT
CATCTTAGGTGCTTTTTAAACTGAAACGCGCACGTTTCTCGGCCCTTCTTTTGCCCCT
TCTTTTAGCCTCTCTTTATTACCTGTTGGATCCTTCACGTATTTTTCCTTGTTTTATACT

CCGTCGGAACTTTGTAGCCAGGACGGAACTTTGAAAACAGTTTTTCTTCTGTTTTCGA
ACTTTAATTTTTCGAGTTATCATTTTTCTTTTAAGGACTTGAACGCTTCCAGTTTTTCC
TCAATATTCACCAGCTAAGAAACATTCTATCGATGCTCGGGAAAGAGCAATTTAATA
GAATTGTCTCTATCGATTTTTGGACATTTTTCCTGTTTTTCAAAAGTGAACGATAAGC
TGATCGTCAAACTTTGGACAATATCCTGAGTTCCGTGGGACAGAGCGGACGCGCCG
GCACATTTCTGAAGACCCGTTAACAACAGAACACGACCAGAATTATTGCGGTAATTC
AGGACTGCAGTGATCGGAGACGGAACTTCCAATACCCGCTTCCTTACTCCAGTACCA
ACTCGGCGAGTTCGTTGGGTCTACAACAAGTAAGACGCTCTACAGCTTGTTGTCAAT
ACAGAGAAATGCGTGTCATGCGTGCTCACGATTAAGTCCCACTCTCACTTCAGTTTA
ATATAATCCCAACTCCCCTTTTTCCCTCTTTTGCCGTTCAATAATTTTTCCGAATTTAA
TCAATAATCTGGTTAGACAGACGTTAAACACTGTGCCATACTGGATTAGTTACAAAT
AACATTTGTTATCTCTGGGTGCTCTGCATATCTGACGCATTGGGTATCGTTCCGTTAT
TCTGATTTCTTTGTTCATTATTATTCATTAAAGTAATCACGTATAAAATTCAGAAGTG
TGTTGTATTGATTATTGGTGGTTGAACGTTTCGTGTAACTCCGGACCAAAGAATTGG
AGACTGATTTAGAGAAATTAATAAAGAAATACACTTATCGATTTTCATAACTTTCCA
TAACTTATTAATCATAATCAATAAGTTTTGTTCCTGGTGAGCCAGGTGGTGCCCTTCT
TCTTACGAGGTTTGTTAATCCATTTAACGTAATATTAATGATTATTTTCCACAACAAC
TATG
> Loki-Sme consensus sequence 5’LTR
TGTTACGAATAATTCCTTATCTATCTAAAGTCACAAAAGGGCCTCCACTGGAGTCTT
CGAGTCTGCTCACACTCACCATCTGTGTTGACTTAAATCAACATTTTTATGACTCTCT
ACCCTGTTTTCCGACACGGCAATCTGTCACAGATGCTCACCACTTTCCATCTTAAAG
AACCTTCTCACCTTGTGGATTTTAAAGGACAACATCATCAAAGACTCTAATACACAA
TCCCCCGTCTGAACCCTGGATGAAACCAGTTCCTCCTTTTTGGGTCATAAATTTTAAT
GACTTTAAGATAACACTTTCACCTGATTTACATGACCCCCTCTAACCCACAGGAATT
CCTAACTTTTGGTTGAACCCAATAAAAAGAAAGGGACCAGATGTCTTCATGATTCAA
ATGTATATCAATTAACATGTATTTTCCATTTTTAGGTCAAGAATCTTTCTACACACAT
GTTTATAAACTTCCTCGTCCAAACCGTATTCAAAACGGGTAATCTACCGTAAAGTTA
CAAAAGAATCATGTCAAGGTATAATCCCCTGCTGCAACTCAAGGATTCTCATGGACC
GCTACCGCCATCTAGCGGTTCCTTTAAGTCAATGTTTTCGTAAATAAACAGTACCTTA
ATCTTTGTCGTTTAAATGTTGCATTCGGGATTAATACTATGTACCTTTTCAGTTTATT
GTCTCTATATGTTGTCTCATTGTTTATTGAAACTATCAATGGAAGTTGGCGAACTCTA
CATATGTCACAATTCAATGATAGTCCGTGTTTGGATTCATTTTAAAGGTTATAACGAT
ACATTGTTATTTTTATCCAAAATTAGTAGTAACACTCCAGTATGGTGTAGTGTTAAA
ACGGCTAATTCAGATTTTATTAACTTCTAACGAAGTACGAGTTTTTGTTTTTTTCGAC
TCGTTCCTCCCCTAAAACTCGCGTCAGCGGAGGGGCGTGTATCATCTTAGGTGCTTTT
TAAACTGAAACGCGCACGTTTCTCGGCCCTTCTTTTGCCCCTTCTTTTAGCCTCTCTTT
ATTACCTGTTGGATCCTTCACGTATTTTTCCTTGTTTTATACTCCGTCGGAACTTTGTA
GCCAGGACGGAACTTTGAAAACAGTTTTTCTTCTGTTTTCGAACTTTAATTTTTCGAG
TTATCATTTTTCTTTTAAGGACTTGAACGCTTCCAGTTTTTCCTCAATATTCACCAGCT
AAGAAACATTCTATCGATGCTCGGGAAAGAGCAATTTAATAGAATTGTCTCTATCGA
TTTTTGGACATTTTTCCTGTTTTTCAAAAGTGAACGATAAGCTGATCGTCAAACTTTG
GACAATATCCTGAGTTCCGTGGGACAGAGCGGACGCGCCGGCACATTTCTGAAGAC
CCGTTAACAACAGAACACGACCAGAATTATTGCGGTAATTCAGGACTGCAGTGATC
GGAGACGGAACTTCCAATACCCGCTTCCTTACTCCAGTACCAACTCGGCGAGTTCGT
TGGGTCTACAACAAGTAAGACGCTCTACAGCTTGTTGTCAATACAGAGAAATGCGTG

TCATGCGTGCTCACGATTAAGTCCCACTCTCACTTCAGTTTAATATAATCCCAACTCC
CCTTTTTCCCTCTTTTGCCGTTCAATAATTTTTCCGAATTTAATCAATAATCTGGTTAG
ACAGACGTTAAACACTGTGCCATACTGGATTAGTTACAAATAACATTTGTTATCTCT
GGGTGCTCTGCATATCTGACGCATTGGGTATCGTTCCGTTATTCTGATTTCTTTGTTC
ATTATTATTCATTAAAGTAATCACGTATAAAATTCAGAAGTGTGTTGTATTGATTATT
GGTGGTTGAACGTTTCGTGTAACTCCGGACCAAAGAATTGGAGACTGATTTAGAGA
AATTAATAAAGAAATACACTTATCGATTTTCATAACTTTCCATAACTTATTAATCATA
ATCAATAAGTTTTGTTCCTGGTGAGCCAGGTGGTGCCCTTCTTCTTACGAGGTTTGTT
AATCCATTTAACGTAATATTAATGATTATTTTCCACAACA
> Loki-Sme consensus sequence 3’LTR
TGTTACGAATAATTCCTTATCTATCTAAAGTCACAAAAGGGCCTCCACTGGAGTCTT
CGAGTCTGCTCACACTCACCATCTGTGTTGACTTAAATCAACATTTTTATGACTCTCT
ACCCTGTTTTCCGACACGGCAATCTGTCACAGATGCTCACCACTTTCCATCTTAAAG
AACCTTCTCACCTTGTGGATTTTAAAGGACAACATCATCAAAGACTCTAATACACAA
TCCCCCGTCTGAACCCTGGATGAAACCAGTTCCTCCTTTTTGGGTCATAAATTTTAAT
GACTTTAAGATAACACTTTCACCTGATTTACATGACCCCCTCTAACCCACAGGAATT
CCTAACTTTTGGTTGAACCCAATAAAAAGAAAGGGACCAGATGTCTTCATGATTCAA
ATGTATATCAATTAACATGTATTTTCCATTTTTAGGTCAAGAATCTTTCTACACACAT
GTTTATAAACTTCCTCGTCCAAACCGTATTCAAAACGGGTAATCTACCGTAAAGTTA
CAAAAGAATCATGTCAAGGTATAATCCCCTGCTGCAACTCAAGGATTCTCATGGACC
GCTACCGCCATCTAGCGGTTCCTTTAAGTCAATGTTTTCGTTAATAAACAGTACCTTA
ATCTTTGTCGTTTAAATGTTGCATTCGGGATTAATACTATGTACCTTTTCAGTTTATT
GTCTCTATATGTTGTCTCATTGTTTATTGAAACTATCAATGGAAGTTGGCGAACTCTA
CATATGTCACAATTCAATGATAGTCCGTGTTTGGATTCATTTTAAAGGTTATAACGAT
ACATTGTTATTTTTATCCAAAATTAGTAGTAACACTCCAGTATGGTGTAGTGTTAAA
ACGGCTAATTCAGATTTTATTAACTTCTAACGAAGTACGAGTTTTTGTTTTTTCGACT
CGTTCCTCCCCTAAAACTCGCGTCAGCGGAGGGGCGTGTATCATCTTAGGTGCTTTTT
AAACTGAAACGCGCACGTTTCTCGGCCCTTCTTTTGCCCCTTCTTTTAGCCTCTCTTT
ATTACCTGTTGGATCCTTCACGTATTTTTCCTTGTTTTATACTCCGTCGGAACTTTGTA
GCCAGGACGGAACTTTGAAAACAGTTTTTCTTCTGTTTTCGAACTTTAATTTTTCGAG
TTATCATTTTTCTTTTAAGGACTTGAACGCTTCCAGTTTTTCCTCAATATTCACCAGCT
AAGAAACATTCTATCGATGCTCGGGAAAGAGCAATTTAATAGAATTGTCTCTATCGA
TTTTTGGACATTTTTCCTGTTTTTCAAAAGTGAACGATAAGCTGATCGTCAAACTTTG
GACAATATCCTGAGTTCCGTGGGACAGAGCGGACGCGCCGGCACATTTCTGAAGAC
CCGTTAACAACAGAACACGACCAGAATTATTGCGGTAATTCAGGACTGCAGTGATC
GGAGACGGAACTTCCAATACCCGCTTCCTTACTCCAGTACCAACTCGGCGAGTTCGT
TGGGTCTACAACAAGTAAGACGCTCTACAGCTTGTTGTCAATACAGAGAAATGCGTG
TCATGCGTGCTCACGATTAAGTCCCACTCTCACTTCAGTTTAATATAATCCCAACTCC
CCTTTTTCCCTCTTTTGCCGTTCAATAATTTTTCCGAATTTAATCAATAATCTGGTTAG
ACAGACGTTAAACACTGTGCCATACTGGATTAGTTACAAATAACATTTGTTATCTCT
GGGTGCTCTGCATATCTGACGCATTGGGTATCGTTCCGTTATTCTGATTTCTTTGTTC
ATTATTATTCATTAAAGTAATCACGTATAAAATTCAGAAGTGTGTTGTATTGATTATT
GGTGGTTGAACGTTTCGTGTAACTCCGGACCAAAGAATTGGAGACTGATTTAGAGA
AATTAATAAAGAAATACACTTATCGATTTTCATAACTTTCCATAACTTATTAATCATA
ATCAATAAGTTTTGTTCCTGGTGAGCCAGGTGGTGCCCTTCTTCTTACGAGGTTTGTT
AATCCATTTAACGTAATATTAATGATTATTTTCCACAACA

> Loki-Sme consensus sequence *smc-gag-pol* gene
ATGGAGGCCGCCGATAGCCCGCCACCTGCAGGCTTCGGTGACCCATTAGATCACTTT
GATTCGGTTCTGACAGACCTAGTTGGACCACCTCTGCGTGAGTTCGTGACTGAAATT
GAGACGACTGACCCCTCTCGACTTCAAGAAAAGCTTTCAGACTTGGTTGCCAAGGCT
ATGTTTGGGCAGGTTAAGGAGAAAGACAAACATAAGATTCTTTCACAAATGGCTGT
AATCATGGCAGCCCAAGAGAGACATACCCAGAAGTTACTAAACCAGCAGGAGAAG
AGCAATAAAGCCCTTCTGCTTATGATTGAACAGAAAGACGAACTCCTCGTCCAGGCC
GATGAGAACATCGCAAAGCTTGAGGAAGCCAAAGCCCTCCTACAAGCAGAAAAGG
AGAATGAGGAAGATAATAGGGAAGATGAGTTGAGAGCAACTCTTCATGAGAAAGA
AAATGAGTTGCAGTCAGTTCATAAACAGCTCACTGCAGCTGCCAAACGGCTCGACC
AGTCTGATGTCCTCCTGATTAAGGCTCAAGATAAGTTCCACGACTTAAATGCTACCA
TCAAAGACCTGACCTCACAAATAGAAAGGCGGGAACAACACGAATGTGACCTCCAA
GCCAAGTTGGACGACGTTTCATTGCGAGTAAGCCTGTTTAAACAGCAGCACGCAATG
AAAGAGGATGAATTAGCGAGGGTCAGAAGACAACTTCAATCTGCTCTGGACCTAGA
TCCAGGTCGACAACGACCTGGTAATGAATCACCCTTCGAGAGAAAGGATGTTCGTA
GTCCAAGAGAGAATCCTCCTCGATCTGGTCTTGTGATTACCCCCGATGATTCCTTGTG
GTCTCGAAATCTCGAGCACAGCCCTTACAGAAGACTCCGATCTCCTATTTCTGAGGA
GCGCCAATCTCTGCCAAGCCCCTCATACAGAGATTTGGACAAAATAGCTCGTAACAT
CACCCGTTTCGAGCCGGGTAACAAGGGTGGTTTTGAGATTAGCCTATATCTCAAAGA
CCTTGAGTACCACCTGCGAAAGTTTTCAGGTGCTACCACGGACGACAAAATCTATTT
GATCAGGATGACGTCCAACAGGACTATCGGCAGTTTTATCGAACGTCAGCCTGACA
GGGTTAGGAGTGACTACAAGTTATTGTGCCAAAGGTTGATAGACGAATATTCTGATT
ACATCCACCCCACAGGGTTGGCGGCTGCTCTTGTGATCAAACAAGAAAGAGATGAG
TCTCCCCAGGTGTATTACAACCGAATACTTCAAGCTTACTTCGGTCCTAAAAATGAA
ATTGGGATGGAAGACGATGTGAACTTTAAGGTCCTTTTCATTCAAAACCTCCATCCA
ACAACTAGCCATCATCTCGGTGTGGCGGCTTGTCCACGCACCCTCACCAGTCGACAA
CTGCGTGATCTGGCCGTGAGAGGCTTCGCCAAATACAAACAGGTCTCTGTCAAAAA
GCCCGACACCAACGTCATATTCAGCATTGACACAAACAACATACCCCTCGAGCTTGA
AGGAGCCACCGAAAGTATTCGTGAGAATCCTTCAAGGCCAAAGGAAAGGCCATTTA
GAGACCACAAACCTTTCCGACCTTCACCGGCCAGTGAAAGCCGGAGGGAACAACCA
CATTCTGGGTTTCGTGGTAATTATCACCAACCTGAATCAAGTCGTCACTTCTCTGACA
TTCGACGACCCAAGTTTCCACCCAGGCAGTGGAACAATTGGGACACCCGTGACGGT
CAGGCCCAAGAGTGGCATCAGAACTACCGCCCTCGGCCGAGAGACCAACCACGTAA
CCCAAGGGTCCGAAACGAGTTTCCCAACCAACAGATGCGGGAAAACGAGCGATCAA
ACCCTCCACAAACATATGAGAAGGGCAAGGAAGAAGCTATCCATCTTTCAGAGAAA
GAGGTAGAAGCTATAAAAAGATTGCTAAAGCAGCAATCAAACAAAGGTAAGAAGG
CCGACCTACTAGCAGTAAGTTTGGTCAAACCTGAAACATCCCCTCCAGACCAGGAA
ACGAACGTGGAGGAGGAAGAACTCTATCGAAGCGATGATAGCAGCCAGGATATGGT
AGAAACTTCAATTGATACTGAACCTCTAAAATCAGAGGATAAATCAGTTAGCGCAG
TCTTGGTAGTCAGTGCTAATTCCACAGACACCTCCAACGGTCAAACTCTTGAAGTCA
TCCAAGGTGAACCACCTTTTCAACAATTCTTGTGTCATCTGATCGAAAGGGGGGTAG
CACGAAAATTCTACCTGTCGACCATTCTTGAGGATGAGGTCGAACATGAAGCCCTCC
TGGACACAGCAGCCGATGTCACCTTGATGTCATTTGCCCTTTTCAATAAGCTACGTC
AGGTAGCTCATCAGTCCAACAGGGACCTGAAACTCCAGCCTTGCGTAATGAACATTC
AACCTTACTCGAATGGTAATACGATTATTAAGAAGCTGGCGCTGATAAAACTGACCC
TGGGACCAATGACAGTCATTCATCCAGTTTATGTGTCTCCACTCAGTTCAATCCCTCT

TCTTGTCGGCAAGGATCTGCTTAACCGCTTCGAACCCCTCATCGACTTTAAACGCTC
GAAGATTTGGGCTCAGGTTCGCAGTCCTCTGCCTATATCGCTCCCTTCTCACGACAC
AGCTCACTGTCTCGTCCTTGAAGGTAAAATAGACACACTTGATGGTCAGCCAACACA
GTTTACGACCGCCGAAGCGGACGACGACTTTGGCACAACCACAGATGCCAATTCGG
GAAAAGGATGGTACCCATCCGACACTCTCGCTCAGCAAGATAACTTCCTCTGTGCTT
TTTCAAACCCCTCGGGGACAGACGAATACTGCCCGATGGTAAACCAAGGCATCCAT
CTCAACGACACACATGTCGAAGACGTGGCCTTGGCCCTCTGGGCTGATCAAACCGCC
ATTAGTCAAGAGCTGTTCAACTCTTTGACAAAGGCCAACCCTGATATCTTGTTCTACT
CGGGTGCTGAACGCTTTCCTCTCCAACCCCAGTCACGAGAGACCATGTCTGCTCAAG
GTGCCTGTACCTTGAAGGTGCGCTGGAATCACCGAGAGCTACATCATAACTTCTTGG
TCGTTCCCAACCTCCCTCATCCAGCCTACATAGGCAGCGACCTGCTAGTCAGGCTGG
ACGTGCAAATCGACACCATAAACAATCTGCTGTGGTCCCTCAAAACTGTTGGAGACA
ACACGGCTCCTGATCTGGATAACCTCAAATCGGGTCAATCCATTCCAGAAGCATGTC
AGGTGGCAAACGCACTTCACGTGGTTGTACCAGCAAACACGACAAACATGTCTATA
AGGTTGAACATACGACCGGGGCAACGTCGACCCCACTCACAGGTGTTCTTTCAACCA
TCGTTGCTTTTCTTTCACTTGGGCCTTACCATTCAGGCCACCCCCTTGATGGATCTAG
CCTCCAGATCCACACACCTGCTGGTACAAAACGAGACCAGGCAGGATATACTGATT
CCCAAGTCCACCCCTTTGGGCTGGCTGATCAGTACAGAATTTCATGATTTTGAGCTG
AGAATTCCGGTGATAGGACAAATACCTCCACTGTTACTCTCCGGCAGCAACGACGA
AAACGTGGTTTACACCAAACCAGCAGGAGCTATAACTCTCTTCCCTGCTGTAACACT
AACCCAAGACACGGTCTGCCGGGTCGATCTCACTGACAGTACCCAAATGGTTCTCCA
AACAGTCCAGGTCATGTCAGTCGATTCATCACCCGATCCAGACACGTATCTTGAGGC
AGAGCCGTCACGATCAACTTCAGACACACCTGCATCTACCAACTTGGACTTCAAAAC
CCAAGTGGATCAGGTGGTGATAGAAGCAGACGCTCTACAGTCAGACGAAGAACGCG
AACAACTTCGCAACATTCTGCTCAAGTATGAAAAGTCGTTTGCACAAGATTCACTTG
ATTGTGGTCTTACCGACATTCACTCAGTGCGTATTCCGACGCCACCCGCGGCTCCTC
CCACGTTCGTACGCCAATACAAAATACCCTTGGCATCGTACGAACCCGTGCAGGAG
ATTATTGATGACCTTCTTCACAGAGGCATCATTAGACCATGTAATAGTACCTACTCA
GCGCCTGTGTGGCCGGTACTTAAACCAAACGGCAAATGGCGACTAACTATCGACTA
CCGGAAACTGAACCAACAAGTTCCCTTGTCCAGGTGGCCGATGACACAACTAAACC
AGGAACTTCCAAAGGTTAAAGACGCCAAATACTTCTCTTCCATGGACATTGCGTCCG
GATTCTGGACTATTCCTGTCCACATAGCAGACCAACATAAGCTGGCGTTCACTTTTG
CAGGTCGACAGTACACGTTCACCAGATGTCCGTTTGGATATGCCAACTCTCCGGCAG
AATTCAACATCTTCTTGAACAAATCGTGCCCTGACGCCAGCGAGCGTGGCACTCTCA
TCTATGTTGATGATATCCTTATTCGGACTCCAACCATGAAGGGTCATATAGACGAGA
TAGAACATGTGCTCGACCAGCTCACCACTGCCGGAGCCAAAATATCTCTGTCCAAAT
GTCAATGGTGTAAGACAAAGGTAAACTATGTGGGACTCCTAGTCGGCTCGAATGGC
GTAGAACCACAAACCAATAGGGTGCAAGGAGTTATGAACATCAAAACTCCAACTAA
CATCTCTGAGCTACGCAGCTTCTTAGGTGTTTGCAACTATTCACGCCAATTCATACCC
GACTTTGCAGACATTGCGAAACCACTCACTGATCTTCTAAAGAAGGATTCCCTTTTC
ACCTTTGGCGAACCTCAAAAAGAGGCCATGCAATCTCTGAAACAGCTATTATGCACG
GCTCCTTGTCTGGCGTACCCGGACTACCAAAAGGAATTCCATCTGGAAGTTGGATTC
TCCAGTCACTGTCTAAGTGCAGGACTTTATCAAGTTCACGACTCAGACAAACGGGTA
GTTGCCTACGCCAGCAAGACCCTAGCCGCACCTGAGATCAAATACACTGATTGCGA
AAAAGCTCTGTTGGCCACAGTCTGGGCCACAAAGCACTTCTCCAACTATCTTGGAGG
CCAAAGGGTGATAGTTGAGACTAATCACCAACCAGTTACTTTCCTTAACAGCCAACG

GATTAGGGACGGAGTTGTGACGAACTCCCGAGTCGCGTCTTGGTTGATGGCCCTGCA
AAGTTTTGACATGGAAGTGCGCTATGCTCAGAATCGCAAAACGCCTCTCGGCATGGA
ACTTGCAGCATGCCAACATTGTTCCGAAGAATCGCCAATTGACTCAATTGTGCTAAC
CGACAGTGTGCTCCCTGAACCTTCTAACCACCATTACTTTGACCCCAATGTCTGTGC
AGACATGGTCACAGCATATGTGGATGGGTGTTCATTCCACCATGAAACTCAAGTAAG
GTCGGGAGCAGGAGTTCTGTGGGTCCACGACACTCCATGTGAACCACAATGTTTTCA
ACTTGGGTCACAATCGTCCCAATACGCCGAGATCGCGGCGATCCTGATAACATTGCA
ATTAGCAAAAGAAAACGATGTCAAAAATCTGACAATATGCACTGACTCGAACTACG
CTCGTTTGAGTTTCACCTCCCACTTACCATTGTGGATCAAAAATGGCTTCATGACTTC
AAACCGTAAACCGGTCAAACACAAAGAACTTTTCATAGCTTGCAAACAACTTGTAG
ACACGTGCGATATGCAAATCTACTGGAAAAAGGTCAGAGGTCACTCCCGAGTCCCT
GGTGCGGACAAGGATCTGAACGACCAGGCAGACAAACTGGCCAAACAGGGGGCCA
TTATCGGCACACCCTGGCAGTTTGACGGAAACTCGGTTCCCACCCAAGCCCTTCCCA
CTGTTTGTGCTGTTACACGTTCAGGCACGACAGCGACACCAGATCCTCGTCGACCAT
CGCCTGTTTCGGCCTCGGTGTCACCTGCCTTCGCCAGCTCGGATCTCTTGACACTACA
AACCGAAGACTCATCTATTCGTGCTATGACACTATTTCTTTCAGATCCTTCTGCTCAC
CCAATCTCCAAAGATGATCTTGATTCAGATGACACCTTGGCGTTTCTTTTTCGGGCAC
GTGCTTCTTTCAAAGTCGTCAATGGGCTGCTTGTTCATGTCTCTGACACACTCACCAC
ACCTGCGTTTGTGGTGCCTCAGAGCCTAAGGGGGGTGATGCTAACCCATGCCCATGA
CTCCCCTTGTTCAGGCCACAGAGGTATCAAAGCCACCCTGAAGTCATTACAGGAAGT
CGCCTACTGGCCACACATGCAGAAAGATGTGTTAGATTACATCAAGGGTTGCTTGGT
GTGCTGTCAATTCCAACCAGCAAGCACTCTTCACCGAGCACCACTACAGCGACGAG
GTGTCTCATTCCCCTGGTCTGATCTTCAAATCGACTGGGTGGGACCCCTCACAAAAT
CCACCAGAGGCAACAAATACTTTCTCACGGTGGTATGTGCATTCACCAAATGGGTGG
AATGCTTACCTGCACCGAATGACACAGCTCAAACCACCGCGTACCTCTTGATGAATC
ATATCTTCTCGAGGTTCGGTCTGCCTGGCAGAGTTAACTCTGACAGAGGCACACACT
TCACGGCTGAAATGATGCAGCAGCTATGGCAGCTGCTGGGTGTGAAAGCCAACTTTC
ACATCAGTTACCGTCCACAATCGTCAGGACAAGTGGAACGGGCTAACCGAACAGTC
GTGAACATCCTCAGAAAGTATGTGTCCTCAAACCACAAGGATTGGGACGTTAAACTC
CCCCTTGTCTTGATGGCCATCAGGGCGACACCCCACGAGTCAACCGGACTCTCGCCC
TTCGAGTTGATGACGGGAAGGCAGATGACTCTGCCCCTGCACTTGCTCTATCAACCG
GGTGAAGCCAGCATCGCCACGGCTTACACCGCACACCAGTACATGACTGACCTACA
CCAACACTTGAGTGCCACTTTCGCTTTCGCCCAGAAAAATCTGGACAAGAGCGCAG
AAGGAAGCAAAACCTACTATGACAAAAAGGCATCCCACGATGAACTCCAGGTGGGA
GATAAAGCGTGGTTTTACATCTTTGCTCAAACCATAGACAACGACAACCACAATGCT
GGAAAACATGGTCGTAAGTTCCTGCCCCACTGGGCTGGACCTTACGTCATCACACAG
AAACTCTCCCCTGTGGTATATCAGCTCAAGATTACCAAAGGACAGAAAGAACCTAT
GCTGAAATGGGTACATCGCAACCAAATTAAGCCACACAAAAACCCCATGGACTAG
TCGGGGACTCCAACGTCCCAGTTAGGTCATAA
> Loki-Sme consensus sequence Smc-Gag-Pol polyprotein
MEAADSPPPAGFGDPLDHFDSVLTDLVGPPLREFVTEIETTDPSRLQEKLSDLVAKAMFG
QVKEKDKHKILSQMAVIMAAQERHTQKLLNQQEKSNKALLLMIEQKDELLVQADENIA
KLEEAKALLQAEKENEEDNREDELRATLHEKENELQSVHKQLTAAAKRLDQSDVLLIK
AQDKFHDLNATIKDLTSQIERREQHECDLQAKLDDVSLRVSLFKQQHAMKEDELARVR
RQLQSALDLDPGRQRPGNESPFERKDVRSPRENPPRSGLVITPDDSLWSRNLEHSPYRRL
RSPISEERQSLPSPSYRDLDKIARNITRFEPGNKGGFEISLYLKDLEYHLRKFSGATTDDKI

YLIRMTSNRTIGSFIERQPDRVRSDYKLLCQRLIDEYSDYIHPTGLAAALVIKQERDESPQ
VYYNRILQAYFGPKNEIGMEDDVNFKVLFIQNLHPTTSHHLGVAACPRTLTSRQLRDLA
VRGFAKYKQVSVKKPDTNVIFSIDTNNIPLELEGATESIRENPSRPKERPFRDHKPFRPSPA
SESRREQPHSGFRGNYHQPESSRHFSDIRRPKFPPRQWNNWDTRDGQAQEWHQNYRPR
PRDQPRNPRVRNEFPNQQMRENERSNPPQTYEKGKEEAIHLSEKEVEAIKRLLKQQSNK
GKKADLLAVSLVKPETSPPDQETNVEEEELYRSDDSSQDMVETSIDTEPLKSEDKSVSAV
LVVSANSTDTSNGQTLEVIQGEPPFQQFLCHLIERGVARKFYLSTILEDEVEHEALLDTAA
DVTLMSFALFNKLRQVAHQSNRDLKLQPCVMNIQPYSNGNTIIKKLALIKLTLGPMTVIH
PVYVSPLSSIPLLVGKDLLNRFEPLIDFKRSKIWAQVRSPLPISLPSHDTAHCLVLEGKIDT
LDGQPTQFTTAEADDDFGTTTDANSGKGWYPSDTLAQQDNFLCAFSNPSGTDEYCPMV
NQGIHLNDTHVEDVALALWADQTAISQELFNSLTKANPDILFYSGAERFPLQPQSRETMS
AQGACTLKVRWNHRELHHNFLVVPNLPHPAYIGSDLLVRLDVQIDTINNLLWSLKTVG
DNTAPDLDNLKSGQSIPEACQVANALHVVVPANTTNMSIRLNIRPGQRRPHSQVFFQPSL
LFFHLGLTIQATPLMDLASRSTHLLVQNETRQDILIPKSTPLGWLISTEFHDFELRIPVIGQI
PPLLLSGSNDENVVYTKPAGAITLFPAVTLTQDTVCRVDLTDSTQMVLQTVQVMSVDSS
PDPDTYLEAEPSRSTSDTPASTNLDFKTQVDQVVIEADALQSDEEREQLRNILLKYEKSF
AQDSLDCGLTDIHSVRIPTPPAAPPTFVRQYKIPLASYEPVQEIIDDLLHRGIIRPCNSTYSA
PVWPVLKPNGKWRLTIDYRKLNQQVPLSRWPMTQLNQELPKVKDAKYFSSMDIASGF
WTIPVHIADQHKLAFTFAGRQYTFTRCPFGYANSPAEFNIFLNKSCPDASERGTLIYVDDI
LIRTPTMKGHIDEIEHVLDQLTTAGAKISLSKCQWCKTKVNYVGLLVGSNGVEPQTNRV
QGVMNIKTPTNISELRSFLGVCNYSRQFIPDFADIAKPLTDLLKKDSLFTFGEPQKEAMQS
LKQLLCTAPCLAYPDYQKEFHLEVGFSSHCLSAGLYQVHDSDKRVVAYASKTLAAPEIK
YTDCEKALLATVWATKHFSNYLGGQRVIVETNHQPVTFLNSQRIRDGVVTNSRVASWL
MALQSFDMEVRYAQNRKTPLGMELAACQHCSEESPIDSIVLTDSVLPEPSNHHYFDPNV
CADMVTAYVDGCSFHHETQVRSGAGVLWVHDTPCEPQCFQLGSQSSQYAEIAAILITLQ
LAKENDVKNLTICTDSNYARLSFTSHLPLWIKNGFMTSNRKPVKHKELFIACKQLVDTC
DMQIYWKKVRGHSRVPGADKDLNDQADKLAKQGAIIGTPWQFDGNSVPTQALPTVCA
VTRSGTTATPDPRRPSPVSASVSPAFASSDLLTLQTEDSSIRAMTLFLSDPSAHPISKDDLD
SDDTLAFLFRARASFKVVNGLLVHVSDTLTTPAFVVPQSLRGVMLTHAHDSPCSGHRGI
KATLKSLQEVAYWPHMQKDVLDYIKGCLVCCQFQPASTLHRAPLQRRGVSFPWSDLQI
DWVGPLTKSTRGNKYFLTVVCAFTKWVECLPAPNDTAQTTAYLLMNHIFSRFGLPGRV
NSDRGTHFTAEMMQQLWQLLGVKANFHISYRPQSSGQVERANRTVVNILRKYVSSNHK
DWDVKLPLVLMAIRATPHESTGLSPFELMTGRQMTLPLHLLYQPGEASIATAYTAHQY
MTDLHQHLSATFAFAQKNLDKSAEGSKTYYDKKASHDELQVGDKAWFYIFAQTIDND
NHNAGKHGRKFLPHWAGPYVITQKLSPVVYQLKITKGQKEPMLKWVHRNQIKPHKNP
MGLVGDSNVPVRS*
> Loki-Sme consensus sequence *env* gene
TTGACGTTCACCCGGCCGAAAAGAATGAATCCGTTGCTGCTCCTCGCGACCTTTTTG
GTGATGGGTGTCCCGGTCTCTCCAGAGATGGTACAGACGGGACCCCCGTCCGGGATT
GCGCTAACGCCCAATCCCGGACTACTGATCACCAACTGTCGCCTGCACACACAGAA
GGTGTTTGTCAGGTTCCGCCCCCGTGAGGTACTCCAGAAGCATCTCCAAGCGACTGG
AAGCCTCACCAGTTGGGCTGGAATGACCTGGACCAGGAACGCGATTAAGCACGCTG
AGGCAGACGTGGCTCATATGCTCCTTCAACTACAGAAATTCACGGTGTCTCAAGATG
AACTACAGGGAAAGAACAAAAGGCAGAAACGGTTTATCGGAGCCCTACTTACTGCA
GCCACAGCAGTAGGCTCCCTATTCAGCATCGGACTCTCTTCCGTCAATGCCGTCAGC
CTAGCAACCGTCAAACGTCACATCACAGAGTTAGACGCAGAAATTCCCGAAATACA

GTCACGACTACACCTTCAACAACAGTCGCTCCAAACCATAGGAGAAACTCTTAAGG
GAACGGTCCTGATTGTGAATACTCAGAGTCTCGCTCTGAACCGAACCATCAGGGCTG
TAGACTCCCTTCTGACAGTAATGCAGGTGGACTACGCTCACACACAACTGGTGGACA
TGCTAATGACAGACATGCTGAGAGAGGTCGGGTCATCAGTAGACAGTCTAGCTACA
GGCCAGATTCCCCCTTTCCTTGTCCCCCTCTCCCTAGTACAGAACCTCTTAGCCAAAG
CCACTAAGGAAACCATCACACCCCTACAAGCCCATCTAGCGTATACCCTTGGTAGTG
CTGTTCCCCTCTACGTCAACCCTGAAAGCGGAGACGTATCCTTCCTCCTCAACCTTCC
CATAATAGCTTCAGAAAACATTTATCGGCTCAAGAACGTGGTGAATGTGGGATGGT
GGCAAGGAGATACCCATCTGAAGGTGCAAACACCCACGATGGTGGCCTACCATGAC
GATAGTCCCGAATTGTATTTGGCACCCAACCTGCGAATGTGCACCCTTACCAAGGAA
ATCCATTACCTGTGTCCAAGCAAACCATTCGTACGTGACAGTACTAATGGAATTTGT
GGACTCAAGCCCATGTTGGCAGAAACTCGCTGTCCCACAACTGCCACTTCTCGATCC
CAGGTCACAGACACCCAAATTGAAATCGTAGGTGATAGATGGCTAGTAAACACTCC
GTTCAAAGTAGCCACTCTCACTTACGGACAACATGACACTGCTACACGGGTTCCCCT
CCCAGAGGAGACGTTATGGATTACCGTCCCTGAAAACGCTATCCTCCACATTGACGA
CATGGCCATCTACTATCTAAACTCCGAACAATACGAGACTGACATTGAACTTCCCGA
TTTCTTCAAAAACTACCAAGTAAAAATCGACCCCAACGCTATCGCTCAAATACAATA
TGTAGGAACCCAAACTATTGACCTAACACCACTTAACGACATCTTACAGATCACTAC
ACCATCAGAAAGACCACCCATACAACCGTTAAACTACACATGGTCCACCCCCGACA
CCTTACTCGCCCTTACCATAGGACTCGGGTATGTGTTAACGTTTGCCATAGCCTTCAT
ATGTGCTAGACGTACTCGAGCCCTCCAGAGGCAACTAAACAAGTGTTCCACTAGTAT
GTCCAAACTGCGTAAACACCACGGCAGGAGTAGAACCACTGAAAGGGATGACGAA
GACGAAGATGTCGTCTACCAAATCTCTCCTACTCTTATAGAGCTAGAATCCGAAACC
TCCGTGCCCATTTAG
> Loki-Sme consensus sequence Env protein
LTFTRPKRMNPLLLLATFLVMGVPVSPEMVQTGPPSGIALTPNPGLLITNCRLHTQKVFV
RFRPREVLQKHLQATGSLTSWAGMTWTRNAIKHAEADVAHMLLQLQKFTVSQDELQG
KNKRQKRFIGALLTAATAVGSLFSIGLSSVNAVSLATVKRHITELDAEIPEIQSRLHLQQQ
SLQTIGETLKGTVLIVNTQSLALNRTIRAVDSLLTVMQVDYAHTQLVDMLMTDMLREV
GSSVDSLATGQIPPFLVPLSLVQNLLAKATKETITPLQAHLAYTLGSAVPLYVNPESGDV
SFLLNLPIIASENIYRLKNVVNVGWWQGDTHLKVQTPTMVAYHDDSPELYLAPNLRMC
TLTKEIHYLCPSKPFVRDSTNGICGLKPMLAETRCPTTATSRSQVTDTQIEIVGDRWLVNT
PFKVATLTYGQHDTATRVPLPEETLWITVPENAILHIDDMAIYYLNSEQYETDIELPDFFK
NYQVKIDPNAIAQIQYVGTQTIDLTPLNDILQITTPSERPPIQPLNYTWSTPDTLLALTIGL
GYVLTFAIAFICARRTRALQRQLNKCSTSMSKLRKHHGRSRTTERDDEDEDVVYQISPTL
IELESETSVPI*
>Loki-Lca consensus sequence
TTTACATAACATTTCGATACTTTGCTGGGTTATTTAATTTTAGTGTGTTTCTTTCTCTC
CCTAAGCCTGCTCACTTCTTTTTGTACTAATTGACATGTACATTTTGTCTCTAACCCCT
CCCTCTCGCTGCCTCCGCCCCCCTCTGTCTATTTAAGGCCACCCCACAGAGACGATC
ACCCCTCTCTGTACCTGTGTTGCCCGCTCTGTGTATCTGTCTGCTTATATGTCTGTCTA
TACATTTTTCTGGTTAGATTGTATGTCGCTATTTCATTATACCTGATGATGATTGCTT
GTGTTGCAATCGAAACGTCGTATGCTATTATTTAATTTATTATTTTTTCTTTATACGTG
ACTTTACCGAAAGAACCAAAGAATAATTCCTGCAGTCACGAAAGCTCCAAAGTTAA
GATTGACCCACCGGCCAACCACAACAATGGATCTAGGGACGAGGCAAGAGTAAGCT
GGCACCCGCGACAACCACCACCAGGGAATGCCGACACAAGCGAATCCTGGAACAAC

```
CACGGGGGGTGGAGAGCGAACGATGGACCACCCACACCCCGGGGAGGCCGACGCC
CACCAAACCAATCACCATACCGCCTCCCACTGCCGTCTCCACTAGCCATAGAGCATA
GACCATCACCAACCAACTGGAATCGATATGCTCACCACATCCGGATGGAAAGGGAA
CCCCTAGAGAATAGGAAGACTCCCATGCAAACAGCAAAGGCAAGGACATCAACAG
AACACACGCAGGAAGAGGACGCACCCAACGCAAAGATAGCCCAAGTGAACATCGG
AAGCTCAAGAAACCCCGAAGCAGCCACTGAGGTGATACCCGAGGGTAGAGGCTACC
ATCTGGGAGAACTAATTCCTGTAGGACCGGTCGGACGAGTCCACCTGGAAGTAAGA
CTGCAAGACACTCTGACCAGCCAGGGCCTGTGTGATACCGGAGCAGAAATATCAAT
GATTTCCGATCACCTATTTGGTGAACTCCAGACTGCCCTCCAGCGGAAAAATGAAGC
ACTACAGATGATGCCAAGTAAGCTTACCTACTCAGGATGCGAAGGGACCCACATGC
GACCTTTGGGCAAAGTCATGATGAAGCTCACGTTCAACGGGATCAACATGAACCAT
CCACTATATGTCGTACAATACCCAGGAAAAGAGCTGATCATTGGGATCGAACTACTG
AATCGGCTAAAACCCATTGTGAACCTTGGATCCCGACAAATCTGGGCACAATCAATA
CCGGAATGGCATAGAGATCCAGAGGACGAAGAAGAAGAACGAAGCCCCACGGGAA
CCCAAGTGCTAGCACAAGAACTCAACGACCTGTCACCGACAATCAAAGACAAAAAC
AAGATTCCGGCTGAAATACGAACCACTCAACTAATGACCAACGGACGGGGAAAGGA
CCTTGACGAAGCCGACCTACGGCCGGTTTTTCCGAAGGACCATCACTCCAGAGAGTA
TCGACAGCCCACTAGTGGCCAAACAGGTGAACAGGAAACAAAGGGAGAGCCCAAG
GGCAAGAGCTCATCCCCTTCGAGTGAAGCAACGGCGATCCAACCACCACTAACGCA
AGGAAAACACCAAACTGAGCCGCCCGAAACCACGAAGGCCCTAGATCCAAACCAA
GAGAAAAGAGTCCTCACAATAGAAGAACAAATAGAACAACAAATTGAAGCAGCAG
ACGCCCTAGAGAATGACACTCAACGAGACGAGCTACGACAACTGCTCCAAAAATAC
AACTCCCTGTTCTCCGAAGACTCCTATGATTGTGGCAAAACCGACCTACATGTGGTA
TCCATTCCCACACTGGAGGGGGCACCAGCGGTATACGTGAAACAATACAAGATTCC
ACTGGCCGCCTATGAATCGATCCAAGAAACCCTGGAGAACCTACTCAAAAAGAACA
TAATCCGGGAATGCAACTCGACCTACAACTCGCCCATTTGGCCAGTACTAAAACCCA
CGGGAAAATGGAGACTCACCATCGACTACCGGCAACTCAATAAAACAGTTCCCCTA
TCTCGATGGCCAATGGCAGAAATCGACCACGGTCTCAACCAAATAAAAGGCGCAAA
GGTACTCACCACAATGGACCTAGCAAATGGCTTCTGGACCATGCCAGTAAAAGAAA
CCGATCAATACAAGTTGGCCTTCACCTTCGACAGAATACAGTATACTTGGAATCGCT
GCCCGTTCGGATACTCGAACTCACCTGCAGACTTCAACATATTCCTACACAAGGCGA
TGGGCGACGCAAAAGAGCGAGGCACCATCGTCTACGTCGACGACATACTAGTCAAG
GACTCCTCATGGCAAGAACATCTCAAGAGCCTCCAACATACCCTTGAACAATTAAA
GGAAGCAGGAGCAAAAATCTCAATACAAAAAGGGCAGTGGGCACGAAAAAGGGTT
GACTATCTTGGGTTTCAGATTGGCACTGAGGGATTGCTACCCCAAACAAACAGAATT
GAAGCCCTCCTCGCACTCAAGTCGCCCAACACAGTCCCGACCCTCCGGAGTTTCTTG
GGAATCTGCAACTACCTCCGACAGTTCGTGGACGACTATGCAGGGATCACTCGACCC
TTAGTCAAACTACTACAAAAGGACGAGCCATGGGAGTGGGGACCAGAACAGGAAG
CTGCAGAAAAAGAATTGAAACAACAGATCACAAGAGCATCCTGCCTAGCATATCCC
GAGAAAGGAAAGGAATACTACCTAGAGACTGCCTACTCAGACCACAGTATCAGCTC
GGTACTGTACCAGAGACAAGAGGCAGAGAAACAAATTATTGCCTATGCCAGTAAAG
CACTCCGGGGAGTAGAGACAAAGTTCTCCGAATGTGAGAAGGCAATATTCGGAAAC
ATCTGGGCTATACAACATTTTCGGAACTTGTTGAATGGGGAAAAGATCATCTTGGAA
ACAAATCATGAATCCTTGGCGTATCTCAACAGCAAGAAGATTCGGGAGGGCCGCGT
GACCAGCAGCCGAATAGCCAATTGGGCACTGACCCTCCAAGGTTTACCAATCACAG
TAAAATACGCGAAAAATAAGAAAAGCCCAGTAGCCCAAGGACTTGCCGATCTACAT
```

GATTGCACCCTAGACAGCGTAGGAGCACATGACGTAGGGGTACCTCCCGAGGACAG
CCGGCACCTTCCCTTTGACAAAGAAACATGTGCCAACATACCAAGGGTGTACATCG
ATGGATGCGCCAAAATGAAAGAAAACAAGCTCCACGCTGGAGCAGGAATCATGTGG
GAATCCGGACCACTCACAGGCATCCAAGAGGGGTTCCAGCTGGGACCGAAGTCAAA
CCAGTATGCAGAGCTGGCTGGGGTACACATTGCCCTCCAAATGGCAGCCGATAACA
AGATACAGACCATGGTGGTGTGCACCGATTCTGACTACGTACAATACAGCTTCCTAA
ACCACCTACCCATATGGAGGAAAAATGGCATGAAGAACCATAGGAATAGACCGATC
CACCATCGAGAACTGTTCGAGGCCATAGACACTATGATCCAGCAGAACGACATGAA
GATATTTTGGAGGAAAGTAAAAGGGCACTCCAAGATACCAGGCCAGGAGAAGGCG
GGGAACGACCGAGCAGACGCCCTAGCAAAACTAGGAAGCCTGGAAGGCACCCCAT
GGAAGCTAGAAGACCAACCCAAACCTAACAGACTAGAAGAGCCGACGAACTCGGA
CTCTGTCGTTATAGGGTTGGTCACAAGAAGCGGAAAGAACTCAGCAGAAACTGGGA
CCAGAGCTCACTTGAGCAATACACAGCCTCCCGAAGACCTGATCAGAATGCAACAA
GAGGACGAAACAATAAAGACACTAGCCGACCATATCCAGGATCCGACAGGAAAAC
CAATCACGACAGACAACTTGGACAACAACAAGGAACTCGAGACCATGCATAACCAA
GTCAGAAGATTTTCAATGCGGAAAGGGCTACTGGTATACACAAACGAAAAAGACGG
TAAAACCAGATGGGTAGTTCCTTCGCAATATCGAGACACAATGCTACAACACGCCC
ACAATGAACCATGTAGCGGACATCGGTGCGAGATGAACACAGGACAGATGTTGCGA
CAGGTGGCCTACTGGCCCAAAATGGGTGAAGACGTACACAAGCATGTGTCCAACTG
CCTCGTGTGTTGCAAATACCAGCCAACGAACTCACAACACCGAGCACCCCTACAGA
AAAGGGGAACAGACTTCCCTTGGTCTAACCTACAAATCGACTGGATAGGACCTGTC
AACAGATCAACCAAAGGAAACAAATACATGCTGACCGTAACCTGTGCCTTTTCCAA
GTGGGTGGAGTGCTTACCAGCCCCCAACAACACCGCGGAGACGACAGCTACCCTAC
TGGTCAACCAGATATTCAGTAGATGGGGCCTACCGTCAACCATCGATTCAGACAAA
GGAAGCCACTTCACAGCCGATGTAATTGCAGAAGTGTGGAAGACATTGGGAGTAAA
AAGACAACTACATGTCGCATACCACCCACAATCCTCAGGACAGGTGGAAAGAGCCA
ACAGGACCATTGTGAGCCTCCTAAAAAAATATGTGTCCACTACCGGATGCGACTGG
GACACAAAACTACCCCTAGTATTGATGGCCATTAGAGCCACCCCCAACCGGGCGAC
AGGCAGGACACCCTTCGAGATCATGACCGGAAGACAAATGACCCTCCCAGTACACC
TGCTATACGTTCCAGCGGAGGACGAGATCACGACGGCAATGACACCGACTGAATAC
ATAACCGGATTGAACAGCCACCTACAGAACACGTTCGCGATGGTCAGAGACAAATT
GCAAGCCGAGGCTCAAGGGAGCAAAGCATACTACGACCAAAAAGCTAGCCATGCA
GAATATCAGCTAGGAGACCAGGTGTTGTACCTCAATTACGAGAAAAGAAAACAAAA
ATGCAGGAAATTCCTCCCTTCCTGGACAGGACCATATGAGATTGTTGACAAAATATC
CCCAGTCGTGTACCAATTAAAAATTACAAAGACAGGAAGAGAGCAGCAATACAAGT
GGGTACACATCAACCAAATCAAGTTGTACCCTGTGCAGATGGAACCAGGGAAGTAA
ATGTAACGGTACAATAGAGGGGGGGAAGAAACTAAGGAGTGGGGATATCACATTA
AAACCACGTATAAATGTTTTCCAGGAAATGGCAGCCCTCCTGCTCACGCTCCTGCTA
ACGTGCATCGCCAAGCTAGAAGGCAGCCAAGACCTACTCAAGCCAGGACCAGAATC
CGGGGTGATGCTGAGAGATTCTTCCGGTTTCCTAGTAACTGAAACAAAAATAGTAAC
CCAACAAGTCTATGTTTCATTAGAACCTGCCATGACAGTAGAGAGACATTTCGGGAA
GTTACCATCCTCGACCACCGCAGAAACCCGAATTTGGTACCATGGCTTCCTAAAATA
CAGCGAACTGAATATCAGAACGATCCTGGAGCAATTACAGAGAACCATGATACTTC
CCCCCGCCACGTTCGACACCACACGGACCAAGAGATCGATCGGCCTGGGAGTGGCC
TTCGGAGTCCTGTCCGGCTTTAGCTCCTTAATCACATTAGGAGTATCTATATCCAACG
CCGTGAAAATTGGACACCTCACGCAAGGCATCGAAAACCTACAGCTCGATATGCAT

CAAATCCGCCTACAAATCCATCACCAACAAAAAAACTTGATCGAAATGGGTTCCAC
CCTTAGAGACACCATAGTTCTGGTCAACCTACATTCAGCCATGTTGAACTTAACCAT
CCAAGAACTGGGGCGTACACTTCGAGAACTATACGACGACGAAAATCTTCGACAAC
CAATCCAATGGGTAATCCAAGACTTGTTGCGTAGCATCAGCGGAAGCATCAGTGATC
TAATCCAAGGTCAAATTCCCTCGTATCTGGTAACTCCGACGCTGATACAAAACGTGC
TACAAACAGCCTCCCTAGAAAGAATCAGTCCAACACAAGTACAAATTGCATTTAAC
ATGGGTACAGCCACACCTATATATGTGAACCCCCAAACCAGACAACTGGCGTTCCTG
TTAAACTTGCCACTAGTCAAACCCGAACAGGTATACCGATTACGAACAGTATTAAAC
GTGGGTACATGGAGAAATGACATTTACTCCAAGATGGAGACACCCTCCGTAATAGC
CTACCAAGAAGAAAGCCCCCTTAAATACCTAGTGCCCGACCTAACCATGTGCCAAA
TAATGAAAGACGTCCATTGGCTCTGCCCAGGCAAACCCTTCCTAATCAATACAGCTG
AAATAATGTGCGGCCTGGCGACCCAAAGCTCCAGAGAGACCTGCAAGTTCAAACTA
AGTAAACCTAAAGAAAGAACCGAAACAGTGGCCATGATAGCTGATGCCCAGTGGCT
AGTCAGCACGCCTCTGGATACAGCAACCATTTCCACCGATCGACATTCCGTTACCAT
GAAAGTTCCGATACCAAGCCCTGTCGTGCTAGTCACGGTACCAAAAGGAACGTTCGT
ACACATAGGGGATAAAGTATTATATAATTTAGACAAAGGCCTACATCAAGCAGATG
TAGAAACAGTAAACATATTCAGGAGCCACAAGTTTGAAATAAGCAACGCCCTGTCG
AACCTATTAAACAGCAAAGAATTCCACACCGTGAGACTAGCCATCAGTGATTCAGA
AATTAACGTAACTGACCTCACCTGGGCAACCGATGAATCCGGACACCAGCGGAATC
AACAGATGGGACTGGGAATCACCGGAGCAACTGGAATGATCATACTAGCCATGGCT
ATCCTAACCCTAGCATACCACGTGTCCAAGATGAATGCCAAGCTTAAAAGACTAATC
GCAGGCACAAAGCACATACCGTTCGAGCCCAGGAACAAAGGAGGAACCAAGGCCT
AA
> Loki-Lca consensus sequence *pol* gene
CTGCCGTCTCCACTAGCCATAGAGCATAGACCATCACCAACCAACTGGAATCGATAT
GCTCACCACATCCGGATGGAAAGGGAACCCCTAGAGAATAGGAAGACTCCCATGCA
AACAGCAAAGGCAAGGACATCAACAGAACACACGCAGGAAGAGGACGCACCCAAC
GCAAAGATAGCCCAAGTGAACATCGGAAGCTCAAGAAACCCCGAAGCAGCCACTG
AGGTGATACCCGAGGGTAGAGGCTACCATCTGGGAGAACTAATTCCTGTAGGACCG
GTCGGACGAGTCCACCTGGAAGTAAGACTGCAAGACACTCTGACCAGCCAGGGCCT
GTGTGATACCGGAGCAGAAATATCAATGATTTCCGATCACCTATTTGGTGAACTCCA
GACTGCCCTCCAGCGGAAAAATGAAGCACTACAGATGATGCCAAGTAAGCTTACCT
ACTCAGGATGCGAAGGGACCCACATGCGACCTTTGGGCAAAGTCATGATGAAGCTC
ACGTTCAACGGGATCAACATGAACCATCCACTATATGTCGTACAATACCCAGGAAA
AGAGCTGATCATTGGGATCGAACTACTGAATCGGCTAAAACCCATTGTGAACCTTGG
ATCCCGACAAATCTGGGCACAATCAATACCGGAATGGCATAGAGATCCAGAGGACG
AAGAAGAAGAACGAAGCCCCACGGGAACCCAAGTGCTAGCACAAGAACTCAACGA
CCTGTCACCGACAATCAAAGACAAAAACAAGATTCCGGCTGAAATACGAACCACTC
AACTAATGACCAACGGACGGGGAAAGGACCTTGACGAAGCCGACCTACGGCCGGTT
TTTCCGAAGGACCATCACTCCAGAGAGTATCGACAGCCCACTAGTGGCCAAACAGG
TGAACAGGAAACAAAGGGAGAGCCCAAGGGCAAGAGCTCATCCCCTTCGAGTGAA
GCAACGGCGATCCAACCACCACTAACGCAAGGAAAACACCAAACTGAGCCGCCCGA
AACCACGAAGGCCCTAGATCCAAACCAAGAGAAAAGAGTCCTCACAATAGAAGAA
CAAATAGAACAACAAATTGAAGCAGCAGACGCCCTAGAGAATGACACTCAACGAG
ACGAGCTACGACAACTGCTCCAAAAATACAACTCCCTGTTCTCCGAAGACTCCTATG
ATTGTGGCAAAACCGACCTACATGTGGTATCCATTCCCACACTGGAGGGGGCACCA

GCGGTATACGTGAAACAATACAAGATTCCACTGGCCGCCTATGAATCGATCCAAGA
AACCCTGGAGAACCTACTCAAAAAGAACATAATCCGGGAATGCAACTCGACCTACA
ACTCGCCCATTTGGCCAGTACTAAAACCCACGGGAAAATGGAGACTCACCATCGAC
TACCGGCAACTCAATAAAACAGTTCCCCTATCTCGATGGCCAATGGCAGAAATCGA
CCACGGTCTCAACCAAATAAAAGGCGCAAAGGTACTCACCACAATGGACCTAGCAA
ATGGCTTCTGGACCATGCCAGTAAAAGAAACCGATCAATACAAGTTGGCCTTCACCT
TCGACAGAATACAGTATACTTGGAATCGCTGCCCGTTCGGATACTCGAACTCACCTG
CAGACTTCAACATATTCCTACACAAGGCGATGGGCGACGCAAAAGAGCGAGGCACC
ATCGTCTACGTCGACGACATACTAGTCAAGGACTCCTCATGGCAAGAACATCTCAAG
AGCCTCCAACATACCCTTGAACAATTAAAGGAAGCAGGAGCAAAAATCTCAATACA
AAAAGGGCAGTGGGCACGAAAAAGGGTTGACTATCTTGGGTTTCAGATTGGCACTG
AGGGATTGCTACCCCAAACAAACAGAATTGAAGCCCTCCTCGCACTCAAGTCGCCC
AACACAGTCCCGACCCTCCGGAGTTTCTTGGGAATCTGCAACTACCTCCGACAGTTC
GTGGACGACTATGCAGGGATCACTCGACCCTTAGTCAAACTACTACAAAAGGACGA
GCCATGGGAGTGGGGACCAGAACAGGAAGCTGCAGAAAAAGAATTGAAACAACAG
ATCACAAGAGCATCCTGCCTAGCATATCCCGAGAAAGGAAAGGAATACTACCTAGA
GACTGCCTACTCAGACCACAGTATCAGCTCGGTACTGTACCAGAGACAAGAGGCAG
AGAAACAAATTATTGCCTATGCCAGTAAAGCACTCCGGGGAGTAGAGACAAAGTTC
TCCGAATGTGAGAAGGCAATATTCGGAAACATCTGGGCTATACAACATTTTCGGAAC
TTGTTGAATGGGGAAAAGATCATCTTGGAAACAAATCATGAATCCTTGGCGTATCTC
AACAGCAAGAAGATTCGGGAGGGCCGCGTGACCAGCAGCCGAATAGCCAATTGGG
CACTGACCCTCCAAGGTTTACCAATCACAGTAAAATACGCGAAAAATAAGAAAAGC
CCAGTAGCCCAAGGACTTGCCGATCTACATGATTGCACCCTAGACAGCGTAGGAGC
ACATGACGTAGGGGTACCTCCCGAGGACAGCCGGCACCTTCCCTTTGACAAAGAAA
CATGTGCCAACATACCAAGGGTGTACATCGATGGATGCGCCAAAATGAAAGAAAAC
AAGCTCCACGCTGGAGCAGGAATCATGTGGGAATCCGGACCACTCACAGGCATCCA
AGAGGGGTTCCAGCTGGGACCGAAGTCAAACCAGTATGCAGAGCTGGCTGGGGTAC
ACATTGCCCTCCAAATGGCAGCCGATAACAAGATACAGACCATGGTGGTGTGCACC
GATTCTGACTACGTACAATACAGCTTCCTAAACCACCTACCCATATGGAGGAAAAAT
GGCATGAAGAACCATAGGAATAGACCGATCCACCATCGAGAACTGTTCGAGGCCAT
AGACACTATGATCCAGCAGAACGACATGAAGATATTTTGGAGGAAAGTAAAAGGGC
ACTCCAAGATACCAGGCCAGGAGAAGGCGGGGAACGACCGAGCAGACGCCCTAGC
AAAACTAGGAAGCCTGGAAGGCACCCCATGGAAGCTAGAAGACCAACCCAAACCT
AACAGACTAGAAGAGCCGACGAACTCGGACTCTGTCGTTATAGGGTTGGTCACAAG
AAGCGGAAAGAACTCAGCAGAAACTGGGACCAGAGCTCACTTGAGCAATACACAG
CCTCCCGAAGACCTGATCAGAATGCAACAAGAGGACGAAACAATAAAGACACTAGC
CGACCATATCCAGGATCCGACAGGAAAACCAATCACGACAGACAACTTGGACAACA
ACAAGGAACTCGAGACCATGCATAACCAAGTCAGAAGATTTTCAATGCGGAAAGGG
CTACTGGTATACACAAACGAAAAAGACGGTAAAACCAGATGGGTAGTTCCTTCGCA
ATATCGAGACACAATGCTACAACACGCCCACAATGAACCATGTAGCGGACATCGGT
GCGAGATGAACACAGGACAGATGTTGCGACAGGTGGCCTACTGGCCCAAAATGGGT
GAAGACGTACACAAGCATGTGTCCAACTGCCTCGTGTGTTGCAAATACCAGCCAAC
GAACTCACAACACCGAGCACCCCTACAGAAAAGGGGAACAGACTTCCCTTGGTCTA
ACCTACAAATCGACTGGATAGGACCTGTCAACAGATCAACCAAAGGAAACAAATAC
ATGCTGACCGTAACCTGTGCCTTTTCCAAGTGGGTGGAGTGCTTACCAGCCCCCAAC
AACACCGCGGAGACGACAGCTACCCTACTGGTCAACCAGATATTCAGTAGATGGGG

CCTACCGTCAACCATCGATTCAGACAAAGGAAGCCACTTCACAGCCGATGTAATTGC
AGAAGTGTGGAAGACATTGGGAGTAAAAAGACAACTACATGTCGCATACCACCCAC
AATCCTCAGGACAGGTGGAAAGAGCCAACAGGACCATTGTGAGCCTCCTAAAAAAA
TATGTGTCCACTACCGGATGCGACTGGGACACAAAACTACCCCTAGTATTGATGGCC
ATTAGAGCCACCCCCAACCGGGCGACAGGCAGGACACCCTTCGAGATCATGACCGG
AAGACAAATGACCCTCCCAGTACACCTGCTATACGTTCCAGCGGAGGACGAGATCA
CGACGGCAATGACACCGACTGAATACATAACCGGATTGAACAGCCACCTACAGAAC
ACGTTCGCGATGGTCAGAGACAAATTGCAAGCCGAGGCTCAAGGGAGCAAAGCATA
CTACGACCAAAAAGCTAGCCATGCAGAATATCAGCTAGGAGACCAGGTGTTGTACC
TCAATTACGAGAAAAGAAAACAAAAATGCAGGAAATTCCTCCCTTCCTGGACAGGA
CCATATGAGATTGTTGACAAAATATCCCCAGTCGTGTACCAATTAAAAATTACAAAG
ACAGGAAGAGAGCAGCAATACAAGTGGGTACACATCAACCAAATCAAGTTGTACCC
TGTGCAGATGGAACCAGGGAAGTAA

> Loki-Lca consensus sequence Pol protein

LPSPLAIEHRPSPTNWNRYAHHIRMEREPLENRKTPMQTAKARTSTEHTQEEDAPNAKIA
QVNIGSSRNPEAATEVIPEGRGYHLGELIPVGPVGRVHLEVRLQDTLTSQGLCDTGAEIS
MISDHLFGELQTALQRKNEALQMMPSKLTYSGCEGTHMRPLGKVMMKLTFNGINMNH
PLYVVQYPGKELIIGIELLNRLKPIVNLGSRQIWAQSIPEWHRDPEDEEEERSPTGTQVLA
QELNDLSPTIKDKNKIPAEIRTTQLMTNGRGKDLDEADLRPVFPKDHHSREYRQPTSGQT
GEQETKGEPKGKSSSPSSEATAIQPPLTQGKHQTEPPETTKALDPNQEKRVLTIEEQIEQQI
EAADALENDTQRDELRQLLQKYNSLFSEDSYDCGKTDLHVVSIPTLEGAPAVYVKQYKI
PLAAYESIQETLENLLKKNIIRECNSTYNSPIWPVLKPTGKWRLTIDYRQLNKTVPLSRWP
MAEIDHGLNQIKGAKVLTTMDLANGFWTMPVKETDQYKLAFTFDRIQYTWNRCPFGYS
NSPADFNIFLHKAMGDAKERGTIVYVDDILVKDSSWQEHLKSLQHTLEQLKEAGAKISI
QKGQWARKRVDYLGFQIGTEGLLPQTNRIEALLALKSPNTVPTLRSFLGICNYLRQFVD
DYAGITRPLVKLLQKDEPWEWGPEQEAAEKELKQQITRASCLAYPEKGKEYYLETAYS
DHSISSVLYQRQEAEKQIIAYASKALRGVETKFSECEKAIFGNIWAIQHFRNLLNGEKIILE
TNHESLAYLNSKKIREGRVTSSRIANWALTLQGLPITVKYAKNKKSPVAQGLADLHDCT
LDSVGAHDVGVPPEDSRHLPFDKETCANIPRVYIDGCAKMKENKLHAGAGIMWESGPL
TGIQEGFQLGPKSNQYAELAGVHIALQMAADNKIQTMVVCTDSDYVQYSFLNHLPIWR
KNGMKNHRNRPIHHRELFEAIDTMIQQNDMKIFWRKVKGHSKIPGQEKAGNDRADALA
KLGSLEGTPWKLEDQPKPNRLEEPTNSDSVVIGLVTRSGKNSAETGTRAHLSNTQPPEDL
IRMQQEDETIKTLADHIQDPTGKPITTDNLDNNKELETMHNQVRRFSMRKGLLVYTNEK
DGKTRWVVPSQYRDTMLQHAHNEPCSGHRCEMNTGQMLRQVAYWPKMGEDVHKHV
SNCLVCCKYQPTNSQHRAPLQKRGTDFPWSNLQIDWIGPVNRSTKGNKYMLTVTCAFS
KWVECLPAPNNTAETTATLLVNQIFSRWGLPSTIDSDKGSHFTADVIAEVWKTLGVKRQ
LHVAYHPQSSGQVERANRTIVSLLKKYVSTTGCDWDTKLPLVLMAIRATPNRATGRTPF
EIMTGRQMTLPVHLLYVPAEDEITTAMTPTEYITGLNSHLQNTFAMVRDKLQAEAQGSK
AYYDQKASHAEYQLGDQVLYLNYEKRKQKCRKFLPSWTGPYEIVDKISPVVYQLKITK
TGREQQYKWVHINQIKLYPVQMEPGK*

> Loki-Lca consensus sequence *env* gene

CTGTGCAGATGGAACCAGGGAAGTAAATGTAACGGTACAATAGAGGGGGGGAAGA
AACTAAGGAGTGGGGATATCACATTAAAACCACGTATAAATGTTTTCCAGGAAATG
GCAGCCCTCCTGCTCACGCTCCTGCTAACGTGCATCGCCAAGCTAGAAGGCAGCCAA
GACCTACTCAAGCCAGGACCAGAATCCGGGGTGATGCTGAGAGATTCTTCCGGTTTC
CTAGTAACTGAAACAAAAATAGTAACCCAACAAGTCTATGTTTCATTAGAACCTGCC

ATGACAGTAGAGAGACATTTCGGGAAGTTACCATCCTCGACCACCGCAGAAACCCG
AATTTGGTACCATGGCTTCCTAAAATACAGCGAACTGAATATCAGAACGATCCTGGA
GCAATTACAGAGAACCATGATACTTCCCCCCGCCACGTTCGACACCACACGGACCA
AGAGATCGATCGGCCTGGGAGTGGCCTTCGGAGTCCTGTCCGGCTTTAGCTCCTTAA
TCACATTAGGAGTATCTATATCCAACGCCGTGAAAATTGGACACCTCACGCAAGGCA
TCGAAAACCTACAGCTCGATATGCATCAAATCCGCCTACAAATCCATCACCAACAA
AAAAACTTGATCGAAATGGGTTCCACCCTTAGAGACACCATAGTTCTGGTCAACCTA
CATTCAGCCATGTTGAACTTAACCATCCAAGAACTGGGGCGTACACTTCGAGAACTA
TACGACGACGAAAATCTTCGACAACCAATCCAATGGGTAATCCAAGACTTGTTGCGT
AGCATCAGCGGAAGCATCAGTGATCTAATCCAAGGTCAAATTCCCTCGTATCTGGTA
ACTCCGACGCTGATACAAAACGTGCTACAAACAGCCTCCCTAGAAAGAATCAGTCC
AACACAAGTACAAATTGCATTTAACATGGGTACAGCCACACCTATATATGTGAACCC
CCAAACCAGACAACTGGCGTTCCTGTTAAACTTGCCACTAGTCAAACCCGAACAGGT
ATACCGATTACGAACAGTATTAAACGTGGGTACATGGAGAAATGACATTTACTCCA
AGATGGAGACACCCTCCGTAATAGCCTACCAAGAAGAAAGCCCCCTTAAATACCTA
GTGCCCGACCTAACCATGTGCCAAATAATGAAAGACGTCCATTGGCTCTGCCCAGGC
AAACCCTTCCTAATCAATACAGCTGAAATAATGTGCGGCCTGGCGACCCAAAGCTCC
AGAGAGACCTGCAAGTTCAAACTAAGTAAACCTAAAGAAAGAACCGAAACAGTGG
CCATGATAGCTGATGCCCAGTGGCTAGTCAGCACGCCTCTGGATACAGCAACCATTT
CCACCGATCGACATTCCGTTACCATGAAAGTTCCGATACCAAGCCCTGTCGTGCTAG
TCACGGTACCAAAAGGAACGTTCGTACACATAGGGGATAAAGTATTATATAATTTA
GACAAAGGCCTACATCAAGCAGATGTAGAAACAGTAAACATATTCAGGAGCCACAA
GTTTGAAATAAGCAACGCCCTGTCGAACCTATTAAACAGCAAAGAATTCCACACCGT
GAGACTAGCCATCAGTGATTCAGAAATTAACGTAACTGACCTCACCTGGGCAACCG
ATGAATCCGGACACCAGCGGAATCAACAGATGGGACTGGGAATCACCGGAGCAACT
GGAATGATCATACTAGCCATGGCTATCCTAACCCTAGCATACCACGTGTCCAAGATG
AATGCCAAGCTTAAAAGACTAATCGCAGGCACAAAGCACATACCGTTCGAGCCCAG
GAACAAAGGAGGAACCAAGGCCTAA
> Loki-Lca consensus sequence Env protein
LCRWNQGSKCNGTIEGGKKLRSGDITLKPRINVFQEMAALLLTLLLTCIAKLEGSQDLLK
PGPESGVMLRDSSGFLVTETKIVTQQVYVSLEPAMTVERHFGKLPSSTTAETRIWYHGFL
KYSELNIRTILEQLQRTMILPPATFDTTRTKRSIGLGVAFGVLSGFSSLITLGVSISNAVKIG
HLTQGIENLQLDMHQIRLQIHHQQKNLIEMGSTLRDTIVLVNLHSAMLNLTIQELGRTLR
ELYDDENLRQPIQWVIQDLLRSISGSISDLIQGQIPSYLVTPTLIQNVLQTASLERISPTQVQ
IAFNMGTATPIYVNPQTRQLAFLLNLPLVKPEQVYRLRTVLNVGTWRNDIYSKMETPSVI
AYQEESPLKYLVPDLTMCQIMKDVHWLCPGKPFLINTAEIMCGLATQSSRETCKFKLSK
PKERTETVAMIADAQWLVSTPLDTATISTDRHSVTMKVPIPSPVVLVTVPKGTFVHIGDK
VLYNLDKGLHQADVETVNIFRSHKFEISNALSNLLNSKEFHTVRLAISDSEINVTDLTWA
TDESGHQRNQQMGLGITGATGMIILAMAILTLAYHVSKMNAKLKRLIAGTKHIPFEPRN
KGGTKA*
>Loki-Pma consensus sequence
ATGGACAAAGCGCTCGTGAGAATTCCCTATTTCGATCCACACAACAAGAACAGGGA
TATACACACTCACATAGCTACAGTAGGACAGGAGGCTAAGATCCGGGGTCTGGGTA
AATCAGCAGAAAAAGCATGTTTATTATCAAGGAGCTTACATGAATCCGCCAGACCA
TGGGTAACCTCCCTGCCAAATAGGGTCAAGAGCGATTTCAAGCAATTAAGCCGAGC
AGTAGTTAAGAGATTTGGGGAATACTCCACACTATCTGAAGGATTGAACGCGGCAA

AATTGATGACTCAGCGTGCAGACGAAGACCCACGAGATTATCTCCTACGCCTCAAG
AGAGCATATTACGCAGGTGACGTGGACAACGCCCAGGAAGAAGGGAAAATATTAA
AAGTTATGTTCTTCCAATCACTAAGACCTGAAATAATGAGGATAGTAGGAATGATGC
TGGACCCCGAGACTGCCTCCCTGCGCCAGATTGAAAGCAAGGCCCACACGGCCTGG
GTGAACGGACGGAGAACCACTAGCATAGCCAAAACACACATCATCACCAAAGATAA
AGAACGCCTCGATCTGGAGGGAAGCGAATGTCCAAGAAATAATAACAGCACCACAC
GGGACGGACATGAAACCCCCCAGGAGAGGGCGCCAAATCCACTAGAATCAACACCC
ATCAACAACACGCTGAATAACACTCCACGTGGGGAGAGAAGCGAGGCAAGAGGGA
CCTGGCACCCACGGCCGCCACAGCCCTGGAATACAGAAGAGGGCGGATCTTGGAAC
TATGGAGGAGGGAGAGCGAGTGACGGATCATACCTACCCCGCGGAGGCCGACGCCC
ACCGAACCAATCACCATATCGTCTCCCACTGCCGCCTCCACTAGCAATAGGGCATAG
ACCACAGTCAGCCAACAGGAGTCGATACGCCCACCACATTCGGATGGAAAGAGAAC
ACCAAGAGAGTATGAAGACTCCCATGCAAACAGCAAAGGCAAGAATACCAACAGA
ACACACACAGGAAGAGGACGCACCCAACGCAAAGATAGCCCAAGTGAACATAAGA
AACTCAAGAAACCCCAAAGCAGTCACCGAAGTGGTACCAGAGGGTAAAGGGTACC
ATCTGGGGGAATTAATTCCCGTAGGACCGGTCGGACGAGTCCACGTGGAAGTAATA
CTGCAAGACACCCTGACCAGCCAGGGCCTTTGTGATACTGGAGCAGAAATATCCAT
GATATCCAGAGACCTATTTGATGAACTCCAGACTGCCCTCCAGAGGAAGAATGAAA
CACTACAGATGATGCCAAGTAGGCTTACCTACTCAGGATGCGAAGGGACCCACATG
CGACCATTGGGCAAGGTCATGATGAAGCTCACGTTCAACGAAATCAACATGAGCCA
TCCACTATATGTCGTACAGAACCCAGGAAAAGAGCTGATCATTGGGATCGAACTACT
GAATCGCCTTAAACCCATTATGAACCTTGAATCCCGACAAATATGGGCACAGTCAAT
GCCGGAGCAGCATAAGGATCCAGAAGATGAAGAAGTCGGAAAAGACTCTACGGGA
ATTCAAATGATAACACAAGAACTCAACGACCTGTCACAGACAATCAAGAACGAAAC
CAAGCGTCCAGCAGAAATACGGACCACTCAACTACTGACCAACAGACTGGGAAAGG
ACCTTGAGGAAGCCGACCTACGACCAGTAGTTCCCAAGGACTGTCCCCTCATGGAGT
ATATACCGCCCACTAGTGGCCAAACAGGTGAACAGGGGACTGGTGAAGGAGAACAC
AAGGCCAAGAGCATATTCCCCTCGGGTGAAGCAAAGGCAATTCAACCGATAATAAC
GCAAAGGAAACAACAAACTGAGCAACCAGGAATCACAGAAGCCTTAGACCCAGAC
CAAGGAAAAAGAGTCCCCACAATAGAAGAACAAATTGAACAACAAATTGAAGCAG
CAGACGCCCTTGAGAATGACACTCAACGAAACGAGCTCCGACAACTGCTCCAGAAA
TACCACTTCCTGTTCTCAGAAGACTCCTATGATTGCGGCAAAACCGACCTACATGTG
GTATCTATTCCCACCCTGGAGGGAGCACCAGCGGTATACGTGAAACAATACAAGAT
TCCACTGGCCGCCTACGAATCAATCCAAGAAACCTTGGAAAACCTACTCAAGAAGA
ATATAATTCGGAAATGCAACTCAACCTACAACTCACCAATTTGGCCTGTGCTAAAAC
CCACGGGCAAGTGGCGACTCACCATCGACTATAGGCAACTCAACAAAACGGTTCCC
CTATCTCGATGGCCAATGGCAGAAATCGACCAAAATCTCAACCAAATAAAAGGCGC
AAAAATACTCACCACAATGGACCTATCAAATGGTTTCTGGACCATGCCCGTGAAAG
AAGCAGATCAATACAAGTTGGCCTTCACATTCAACGGGATACAGTATACTTGGAACC
GATGCCCATTTGGATACTCAAACTCACCTGCAGACTTCAACATATTCCTACACAAGG
CAATGGGTGACGCAAAAGAGCGAGGCACCATCGTCTATGTCGACGACATACTGGTC
AAGGACTCCTCATGGGAAGAACACCTCGAGAGCCTCCAGCATACCCTCGAGCAGTT
AAAGGAAGCAGGTGCAAAGATATCAATACAAAAGGGACAGTGGGCACGCAAAAGG
GTGGACTACCTCGGGTTCCAGATCGGCACTGAGGGAATGTTACCCCAAACGAAAAG
ACTTGAAGCCCTCCTCGCGCTCAAACCGCCCACCACAGTCACGACCCTTCGGAGCTT
CTTAGGGATCTGCAACTACCTTCGACAGTTCGTGGACGACTATGCAGGGATCGCTCG

ACCCTTAGTCAAACTACTACAAAAAGACGAACCATGGGAATGGGGACAAGAACAGG
AATCCGCAGTGGAAGAATTAAAACGACAGATCACAGGGGCATCCTGCCTAGCATAC
CCAGAGAAAGGAAAAGAATACTACCTAGAGACCGCATACTCAGATCATAGTATCAG
CTCGGTACTGTACCAGATACAGGAAGCAGGGAAACGAATTATTGCCTATGCCAGCA
AAGCCCTCCGGGGGGTAGAGACAAAGTTCTCCGAGTGTGAGAAGGCAATATTCGGA
AACATCTGGGCTATACAACATTTTCGGAACTTGTTGAACGGAGAGAAAATCATCCTA
GAAACGAATCATGAATCCTTGGCGTACCTCAACAGTAAGAAGATTCGAGAGGGCCG
CGTAACCAGCAGCCGAATAGCCAATTGGGCACTGACACTCCAAGGCCTACCAATCA
CAGTAAAATACGCAAAAAACAAGAAAAGTCCAATGGCCCAAGGTCTTGCTGACTTA
CATGATTGCACCCTAAACAGCGTAGAAGCATGTGCCGCAAAGGTACCTCCGGAGGA
CAACAGGCACCTGCCCTTCAAAAAGGAAACATGTGCCGACCTACCAATGGTGTACA
TCGATGGATGCTCCAAAATGAAGGAAAACAAACTCCACGCCGGATCAGGAATCGTT
TGGGAAACTGGCCCACTCACAGGGATTCAAGAGGGGTTTCAACTGGGACCAAAGTC
CAACCAGTATGCAGAATTGGCCGGGGTACACATCGCCGTCCAAATGGCATCCAATA
ACCAAATACAGACCATGGTGGTGTGCACTGATTCCAACTACGTACAGCACAGCTTCC
TGCACCACTTACCCATATGGAAAAAGAACGGTATGAAGAACCATAGAAATAAACCA
ATTCACCATCAAGAACTGTTCGAGGCCATAGATACGATGGTTCAGAACAAAGATAT
GAAGATATTTTGGAAGAAGGTCAAAGGGCACTCCAAGATACCTGGCCAAGAGAAAA
CTGGGAATGACCAGGCAGACGCCCTAGCAAAAGTGGGAAGCCTGGAAGGCACACC
GTGGAAGCTAGAAAACCAAACCGAAAGAGACGGACAACAAGAGTCGACGAACCAA
GACCCCGCCGTGATAAGACTGGTCACAAGAAGTGGGAAGAACCCTGCAGAAGCTGG
AACCAGAGCCCAATTGAACAGTGCACAGCCTTCCGAAGACCTAGTCAGAATGCAAC
AAGAGGACGAAACACTAAAGACACTAGCTGATCATATCCAGGATCCAACAGGAAAA
CCAATCACTACCGACAACCTGAACGACAACAAAGAGCTCGAGACCATGCATAACCA
AATCAAAAGGTTCTCAATGCGAAAAGGGCTACTGGTATATACGAACGAAAAAGACG
GTAAAACTAGATGGGTAGTCCCCTCACAATATCGAGAGATAATGTTACAACACGCC
CACGACGAACCATGCAGCGGCCACCGATGCGAGATGAACACAGGACAGATGCTGCG
GCAGGTGGCCTACTGGCCTAAAATGGGTGAAGACATACACAAACATGTGTCCAACT
GCCTCGTGTGTTGTAAGTACCAGCCAACGACCTCACAGCATAGGGCACCCCTACAG
AAGAGGGGAACAGACTATCCTTGGTCCAACCTACAAATCGACTGGATAGGACCTGT
CAACCGATCAACCAAAGGAAATAAGTACATGCTGACTGTGACCTGTGCATTCTCCAA
ATGGATAGAGTGTTTACCAGCACCCAACAACACCGCGGAGACGACAGCTATCCTAC
TGATCAACCAAATATTCAGTAGGTGGGGCCTACCGTCAACAATCGATTCAGACAAA
GGAAGCCACTTCACAGCCGAAGTCATCACAGAAGTGTGGAAGATGTTGGGAGTAAA
AAGACAACTACATGTCGCATACCATCCACAATCCTCAGGACAGGTGGAAAGAGCCA
ACAGGACCATCGTGAGCCTCCTAAAGAAATATGTGTCCACCACGGGACGCGACTGG
GATACGAAACTACCCCTAGTCTTGATGGCCATGCGAGCCACACCCAACCGAGCTAC
AGGCAGGACACCCTTCGAGATCATGACAGGGAGACTAATGACCCTCCCAGTGCACC
TGCTGTATGTCCCAGCGGAGGACGAGACCCCGACGGCAATGACGCCCACTGAATAC
CTAACGGGACTGAACAGTCACCTACAGAGCACATTCGCTATGGTCAGAGACCAGTT
GCAAGCAGAGGCTCAGGGGAGCAAAGCATACTATGACCAAAAAGCCAGCCACGCA
GAATACCAGCTAGGAGACCAGGTGTTGTACCTCAATTACGAGAAAAGAAAGCAGAA
GTGCAGGAAATTCCTCCCTTCCTGGACAGGACCACACGAGATTGTTGACAAAATGTC
CCCAGTCGTATATCAGCTAAAAATCACAAAGACAGGAAGAGAGCAGCAGTACAAGT
GGGTACACATCAATCAAATCAAACCGTACCCTGTGCTGGTGGAACAAGAGAAGGAA
TAAAGAAAAAAAAGGGGGGGGGGACCTAAACTAAATTGGGCATGTCACACTAAA

ATAATGTATAAAATGTTTTCCAGGAAATGGCAGCCCTCCTGCTCACGCTCCTGCTAA
CTTGCTTTGCCAAGCTAGAGGGCAGCCAAGACCTGCTCAAGCCAGGACCAGAATCC
GGGGTGATGCTGAGAGACTCTTCCGGCTTCCTAGTGACCGAAACGAAAATAGTAAC
CCAACAAGTCTATGTTTCCTTGGAACCTGCCGTGGTAATAGAGAAACACTTCGGGAA
GATACCACCCTCAACTACCAGAGCAACCCGAAACTGGTACCATGACTTCCTAAGAT
ACAGTGAACAGAATGTCGGAACGATCCTGAAGCAATTACAAAGAACCATGATACTT
CCCCCCACCACTGTCAACACAACACGTTCCAAGAGATCGATCGGCCTGGGAGGGGC
CTTCGGAATACTATCCGGCTTTAGCTCCTTAATCACACTAGGAGTATCCATATCCAA
CGCCGTGAAGATTGGACACCTCACGCAAGGCGTCGAGAATTTACAGCTCGATATGC
ATCAAATCCGTCAGCAAATCCGACACCAACAAAAGAATCTGATCGAAATGGGATCC
ACCCTCAGAGACACCGTGGTTCTCGTTAACCTGCATTCCGCCATGATTAACTTAACA
ATCCAAGAATTGGGTTTCGCAATCCGAGATCTATACGAAGAAGAGAATACTCGACA
GCCAATTCAATGGTTGATCCAAGACTTGTTGCGTAGCACCAGCGGCAGCATCAGTGA
TCTTATCCAAGGACAAATTCCTTCATATCTGGTAACTCCAACACTGATAAAAAATGT
GCTACAAGCAGCATCCCCAGAAACAATCAGTCCAACACAAGTACAAATTGCCTTTA
ACATGGGTACAGCCACACCCATACATGTGGACCCCCAAACCAGACAGCTGGCATTT
TTGTTAAATTTACCACTGGTCAAACCCGAAAATGTATTTCGATTGCGGACAGTATTG
AACGTGGGTACATGGAATAATGACATTTACTCCAAGATGGAAACCCCCTTAGTGGTG
GCCTACCAGGAAGAGGGCCCCCTTAAGTACCTAGTGCCTGACCTAAACATGTGCCA
AATAATGAAAGAGGTTCATTGGCTCTGCCCAGGGGAACCATTCCTAATCAATACAGC
AGAAATAATGTGCGGTCTGTCAACCCTAAGCTCCAGGGAGACTTGCAAGTTCAAGCT
AAGTAAACCCAAAGAAAGAACCGAAACAGTGGCCATGATAGCCGATGCTCAGTGGT
TAATAAGCACACCTCTAGATACGGCAACCGTTTCCACCGATCAACACTCCGTAACCA
TGAAAGTTCCGATTCCAAGCCACGTCGCACTGGTCACAGTACCAAGAGGAACATTC
GTACACATAGGGGACAAGGTATTATATAATTTAGACAAAAGCCTACATCAAGCAGA
GGTAGAAATAGTAAACACATTCCGAAGCCACAAGTTTGAAATAAGCAGCACTCTGT
CTAACCTGTTGAACAGTAAAGAATTTCATACGGTGAAATTAGCCATCAGCGAGTCAG
AGATTAACGTAACAGATCTCACTTGGGCGACCGAAGAGGCCGGATATCAGCGGAAT
CAACAGATGGGACTGGGAATCACAGGGGCAACAGGAATGATCATACTAGCCATGGC
TATCCTAACCCTAGCATACCATATGGCCAAAATGAATGCCGAACTTAAAAGACTAAT
CGCAGGCACAAGCCACATACCATTCGAGCCCAGGAACAAAGGTAGAGTCAAGGCCT
GA
> Loki-Pma consensus sequence *gag-pol* gene
ATGGACAAAGCGCTCGTGAGAATTCCCTATTTCGATCCACACAACAAGAACAGGGA
TATACACACTCACATAGCTACAGTAGGACAGGAGGCTAAGATCCGGGGTCTGGGTA
AATCAGCAGAAAAAGCATGTTTATTATCAAGGAGCTTACATGAATCCGCCAGACCA
TGGGTAACCTCCCTGCCAAATAGGGTCAAGAGCGATTTCAAGCAATTAAGCCGAGC
AGTAGTTAAGAGATTTGGGGAATACTCCACACTATCTGAAGGATTGAACGCGGCAA
AATTGATGACTCAGCGTGCAGACGAAGACCCACGAGATTATCTCCTACGCCTCAAG
AGAGCATATTACGCAGGTGACGTGGACAACGCCCAGGAAGAAGGGAAAATATTAA
AAGTTATGTTCTTCCAATCACTAAGACCTGAAATAATGAGGATAGTAGGAATGATGC
TGGACCCCGAGACTGCCTCCCTGCGCCAGATTGAAAGCAAGGCCCACACGGCCTGG
GTGAACGGACGGAGAACCACTAGCATAGCCAAAACACACATCATCACCAAAGATAA
AGAACGCCTCGATCTGGAGGGAAGCGAATGTCCAAGAAATAATAACAGCACCACAC
GGGACGGACATGAAACCCCCCAGGAGAGGGCGCCAAATCCACTAGAATCAACACCC
ATCAACAACACGCTGAATAACACTCCACGTGGGGAGAGAAGCGAGGCAAGAGGGA

CCTGGCACCCACGGCCGCCACAGCCCTGGAATACAGAAGAGGGCGGATCTTGGAAC
TATGGAGGAGGGAGAGCGAGTGACGGATCATACCTACCCCGCGGAGGCCGACGCCC
ACCGAACCAATCACCATATCGTCTCCCACTGCCGCCTCCACTAGCAATAGGGCATAG
ACCACAGTCAGCCAACAGGAGTCGATACGCCCACCACATTCGGATGGAAAGAGAAC
ACCAAGAGAGTATGAAGACTCCCATGCAAACAGCAAAGGCAAGAATACCAACAGA
ACACACACAGGAAGAGGACGCACCCAACGCAAAGATAGCCCAAGTGAACATAAGA
AACTCAAGAAACCCCAAAGCAGTCACCGAAGTGGTACCAGAGGGTAAAGGGTACC
ATCTGGGGGAATTAATTCCCGTAGGACCGGTCGGACGAGTCCACGTGGAAGTAATA
CTGCAAGACACCCTGACCAGCCAGGGCCTTTGTGATACTGGAGCAGAAATATCCAT
GATATCCAGAGACCTATTTGATGAACTCCAGACTGCCCTCCAGAGGAAGAATGAAA
CACTACAGATGATGCCAAGTAGGCTTACCTACTCAGGATGCGAAGGGACCCACATG
CGACCATTGGGCAAGGTCATGATGAAGCTCACGTTCAACGAAATCAACATGAGCCA
TCCACTATATGTCGTACAGAACCCAGGAAAAGAGCTGATCATTGGGATCGAACTACT
GAATCGCCTTAAACCCATTATGAACCTTGAATCCCGACAAATATGGGCACAGTCAAT
GCCGGAGCAGCATAAGGATCCAGAAGATGAAGAAGTCGGAAAAGACTCTACGGGA
ATTCAAATGATAACACAAGAACTCAACGACCTGTCACAGACAATCAAGAACGAAAC
CAAGCGTCCAGCAGAAATACGGACCACTCAACTACTGACCAACAGACTGGGAAAGG
ACCTTGAGGAAGCCGACCTACGACCAGTAGTTCCCAAGGACTGTCCCCTCATGGAGT
ATATACCGCCCACTAGTGGCCAAACAGGTGAACAGGGGACTGGTGAAGGAGAACAC
AAGGCCAAGAGCATATTCCCCTCGGGTGAAGCAAAGGCAATTCAACCGATAATAAC
GCAAAGGAAACAACAAACTGAGCAACCAGGAATCACAGAAGCCTTAGACCCAGAC
CAAGGAAAAAGAGTCCCCACAATAGAAGAACAAATTGAACAACAAATTGAAGCAG
CAGACGCCCTTGAGAATGACACTCAACGAAACGAGCTCCGACAACTGCTCCAGAAA
TACCACTTCCTGTTCTCAGAAGACTCCTATGATTGCGGCAAAACCGACCTACATGTG
GTATCTATTCCCACCCTGGAGGGAGCACCAGCGGTATACGTGAAACAATACAAGAT
TCCACTGGCCGCCTACGAATCAATCCAAGAAACCTTGGAAAACCTACTCAAGAAGA
ATATAATTCGGAAATGCAACTCAACCTACAACTCACCAATTTGGCCTGTGCTAAAAC
CCACGGGCAAGTGGCGACTCACCATCGACTATAGGCAACTCAACAAAACGGTTCCC
CTATCTCGATGGCCAATGGCAGAAATCGACCAAAATCTCAACCAAATAAAAGGCGC
AAAAATACTCACCACAATGGACCTATCAAATGGTTTCTGGACCATGCCCGTGAAAG
AAGCAGATCAATACAAGTTGGCCTTCACATTCAACGGGATACAGTATACTTGGAACC
GATGCCCATTTGGATACTCAAACTCACCTGCAGACTTCAACATATTCCTACACAAGG
CAATGGGTGACGCAAAAGAGCGAGGCACCATCGTCTATGTCGACGACATACTGGTC
AAGGACTCCTCATGGGAAGAACACCTCGAGAGCCTCCAGCATACCCTCGAGCAGTT
AAAGGAAGCAGGTGCAAAGATATCAATACAAAAGGGACAGTGGGCACGCAAAAGG
GTGGACTACCTCGGGTTCCAGATCGGCACTGAGGGAATGTTACCCCAAACGAAAAG
ACTTGAAGCCCTCCTCGCGCTCAAACCGCCCACCACAGTCACGACCCTTCGGAGCTT
CTTAGGGATCTGCAACTACCTTCGACAGTTCGTGGACGACTATGCAGGGATCGCTCG
ACCCTTAGTCAAACTACTACAAAAAGACGAACCATGGGAATGGGGACAAGAACAGG
AATCCGCAGTGGAAGAATTAAAACGACAGATCACAGGGGCATCCTGCCTAGCATAC
CCAGAGAAAGGAAAAGAATACTACCTAGAGACCGCATACTCAGATCATAGTATCAG
CTCGGTACTGTACCAGATACAGGAAGCAGGGAAACGAATTATTGCCTATGCCAGCA
AAGCCCTCCGGGGGGTAGAGACAAAGTTCTCCGAGTGTGAGAAGGCAATATTCGGA
AACATCTGGGCTATACAACATTTTCGGAACTTGTTGAACGGAGAGAAAATCATCCTA
GAAACGAATCATGAATCCTTGGCGTACCTCAACAGTAAGAAGATTCGAGAGGGCCG
CGTAACCAGCAGCCGAATAGCCAATTGGGCACTGACACTCCAAGGCCTACCAATCA

CAGTAAAATACGCAAAAAACAAGAAAAGTCCAATGGCCCAAGGTCTTGCTGACTTA
CATGATTGCACCCTAAACAGCGTAGAAGCATGTGCCGCAAAGGTACCTCCGGAGGA
CAACAGGCACCTGCCCTTCAAAAAGGAAACATGTGCCGACCTACCAATGGTGTACA
TCGATGGATGCTCCAAAATGAAGGAAAACAAACTCCACGCCGGATCAGGAATCGTT
TGGGAAACTGGCCCACTCACAGGGATTCAAGAGGGGTTTCAACTGGGACCAAAGTC
CAACCAGTATGCAGAATTGGCCGGGGTACACATCGCCGTCCAAATGGCATCCAATA
ACCAAATACAGACCATGGTGGTGTGCACTGATTCCAACTACGTACAGCACAGCTTCC
TGCACCACTTACCCATATGGAAAAAGAACGGTATGAAGAACCATAGAAATAAACCA
ATTCACCATCAAGAACTGTTCGAGGCCATAGATACGATGGTTCAGAACAAAGATAT
GAAGATATTTTGGAAGAAGGTCAAAGGGCACTCCAAGATACCTGGCCAAGAGAAAA
CTGGGAATGACCAGGCAGACGCCCTAGCAAAAGTGGGAAGCCTGGAAGGCACACC
GTGGAAGCTAGAAAACCAAACCGAAAGAGACGGACAACAAGAGTCGACGAACCAA
GACCCCGCCGTGATAAGACTGGTCACAAGAAGTGGGAAGAACCCTGCAGAAGCTGG
AACCAGAGCCCAATTGAACAGTGCACAGCCTTCCGAAGACCTAGTCAGAATGCAAC
AAGAGGACGAAACACTAAAGACACTAGCTGATCATATCCAGGATCCAACAGGAAAA
CCAATCACTACCGACAACCTGAACGACAACAAAGAGCTCGAGACCATGCATAACCA
AATCAAAAGGTTCTCAATGCGAAAAGGGCTACTGGTATATACGAACGAAAAAGACG
GTAAAACTAGATGGGTAGTCCCCTCACAATATCGAGAGATAATGTTACAACACGCC
CACGACGAACCATGCAGCGGCCACCGATGCGAGATGAACACAGGACAGATGCTGCG
GCAGGTGGCCTACTGGCCTAAAATGGGTGAAGACATACACAAACATGTGTCCAACT
GCCTCGTGTGTTGTAAGTACCAGCCAACGACCTCACAGCATAGGGCACCCCTACAG
AAGAGGGGAACAGACTATCCTTGGTCCAACCTACAAATCGACTGGATAGGACCTGT
CAACCGATCAACCAAAGGAAATAAGTACATGCTGACTGTGACCTGTGCATTCTCCAA
ATGGATAGAGTGTTTACCAGCACCCAACAACACCGCGGAGACGACAGCTATCCTAC
TGATCAACCAAATATTCAGTAGGTGGGGCCTACCGTCAACAATCGATTCAGACAAA
GGAAGCCACTTCACAGCCGAAGTCATCACAGAAGTGTGGAAGATGTTGGGAGTAAA
AAGACAACTACATGTCGCATACCATCCACAATCCTCAGGACAGGTGGAAAGAGCCA
ACAGGACCATCGTGAGCCTCCTAAAGAAATATGTGTCCACCACGGGACGCGACTGG
GATACGAAACTACCCCTAGTCTTGATGGCCATGCGAGCCACACCCAACCGAGCTAC
AGGCAGGACACCCTTCGAGATCATGACAGGGAGACTAATGACCCTCCCAGTGCACC
TGCTGTATGTCCCAGCGGAGGACGAGACCCCGACGGCAATGACGCCCACTGAATAC
CTAACGGGACTGAACAGTCACCTACAGAGCACATTCGCTATGGTCAGAGACCAGTT
GCAAGCAGAGGCTCAGGGGAGCAAAGCATACTATGACCAAAAAGCCAGCCACGCA
GAATACCAGCTAGGAGACCAGGTGTTGTACCTCAATTACGAGAAAAGAAAGCAGAA
GTGCAGGAAATTCCTCCCTTCCTGGACAGGACCACACGAGATTGTTGACAAAATGTC
CCCAGTCGTATATCAGCTAAAAATCACAAAGACAGGAAGAGAGCAGCAGTACAAGT
GGGTACACATCAATCAAATCAAACCGTACCCTGTGCTGGTGGAACAAGAGAAGGAA
TAA
> Loki-Pma consensus sequence Gag-Pol polyprotein
MDKALVRIPYFDPHNKNRDIHTHIATVGQEAKIRGLGKSAEKACLLSRSLHESARPWVT
SLPNRVKSDFKQLSRAVVKRFGEYSTLSEGLNAAKLMTQRADEDPRDYLLRLKRAYYA
GDVDNAQEEGKILKVMFFQSLRPEIMRIVGMMLDPETASLRQIESKAHTAWVNGRRTTS
IAKTHIITKDKERLDLEGSECPRNNNSTTRDGHETPQERAPNPLESTPINNTLNNTPRGER
SEARGTWHPRPPQPWNTEEGGSWNYGGGRASDGSYLPRGGRRPPNQSPYRLPLPPPLAI
GHRPQSANRSRYAHHIRMEREHQESMKTPMQTAKARIPTEHTQEEDAPNAKIAQVNIRN
SRNPKAVTEVVPEGKGYHLGELIPVGPVGRVHVEVILQDTLTSQGLCDTGAEISMISRDL

FDELQTALQRKNETLQMMPSRLTYSGCEGTHMRPLGKVMMKLTFNEINMSHPLYVVQ
NPGKELIIGIELLNRLKPIMNLESRQIWAQSMPEQHKDPEDEEVGKDSTGIQMITQELNDL
SQTIKNETKRPAEIRTTQLLTNRLGKDLEEADLRPVVPKDCPLMEYIPPTSGQTGEQGTG
EGEHKAKSIFPSGEAKAIQPIITQRKQQTEQPGITEALDPDQGKRVPTIEEQIEQQIEAADA
LENDTQRNELRQLLQKYHFLFSEDSYDCGKTDLHVVSIPTLEGAPAVYVKQYKIPLAAY
ESIQETLENLLKKNIIRKCNSTYNSPIWPVLKPTGKWRLTIDYRQLNKTVPLSRWPMAEID
QNLNQIKGAKILTTMDLSNGFWTMPVKEADQYKLAFTFNGIQYTWNRCPFGYSNSPAD
FNIFLHKAMGDAKERGTIVYVDDILVKDSSWEEHLESLQHTLEQLKEAGAKISIQKGQW
ARKRVDYLGFQIGTEGMLPQTKRLEALLALKPPTTVTTLRSFLGICNYLRQFVDDYAGIA
RPLVKLLQKDEPWEWGQEQESAVEELKRQITGASCLAYPEKGKEYYLETAYSDHSISSV
LYQIQEAGKRIIAYASKALRGVETKFSECEKAIFGNIWAIQHFRNLLNGEKIILETNHESLA
YLNSKKIREGRVTSSRIANWALTLQGLPITVKYAKNKKSPMAQGLADLHDCTLNSVEAC
AAKVPPEDNRHLPFKKETCADLPMVYIDGCSKMKENKLHAGSGIVWETGPLTGIQEGF
QLGPKSNQYAELAGVHIAVQMASNNQIQTMVVCTDSNYVQHSFLHHLPIWKKNGMKN
HRNKPIHHQELFEAIDTMVQNKDMKIFWKKVKGHSKIPGQEKTGNDQADALAKVGSLE
GTPWKLENQTERDGQQESTNQDPAVIRLVTRSGKNPAEAGTRAQLNSAQPSEDLVRMQ
QEDETLKTLADHIQDPTGKPITTDNLNDNKELETMHNQIKRFSMRKGLLVYTNEKDGKT
RWVVPSQYREIMLQHAHDEPCSGHRCEMNTGQMLRQVAYWPKMGEDIHKHVSNCLV
CCKYQPTTSQHRAPLQKRGTDYPWSNLQIDWIGPVNRSTKGNKYMLTVTCAFSKWIEC
LPAPNNTAETTAILLINQIFSRWGLPSTIDSDKGSHFTAEVITEVWKMLGVKRQLHVAYH
PQSSGQVERANRTIVSLLKKYVSTTGRDWDTKLPLVLMAMRATPNRATGRTPFEIMTG
RLMTLPVHLLYVPAEDETPTAMTPTEYLTGLNSHLQSTFAMVRDQLQAEAQGSKAYYD
QKASHAEYQLGDQVLYLNYEKRKQKCRKFLPSWTGPHEIVDKMSPVVYQLKITKTGRE
QQYKWVHINQIKPYPVLVEQEKE*
> Loki-Pma consensus sequence *env* gene
ATGGCAGCCCTCCTGCTCACGCTCCTGCTAACTTGCTTTGCCAAGCTAGAGGGCAGC
CAAGACCTGCTCAAGCCAGGACCAGAATCCGGGGTGATGCTGAGAGACTCTTCCGG
CTTCCTAGTGACCGAAACGAAAATAGTAACCCAACAAGTCTATGTTTCCTTGGAACC
TGCCGTGGTAATAGAGAAACACTTCGGGAAGATACCACCCTCAACTACCAGAGCAA
CCCGAAACTGGTACCATGACTTCCTAAGATACAGTGAACAGAATGTCGGAACGATC
CTGAAGCAATTACAAAGAACCATGATACTTCCCCCCACCACTGTCAACACAACACGT
TCCAAGAGATCGATCGGCCTGGGAGGGGCCTTCGGAATACTATCCGGCTTTAGCTCC
TTAATCACACTAGGAGTATCCATATCCAACGCCGTGAAGATTGGACACCTCACGCAA
GGCGTCGAGAATTTACAGCTCGATATGCATCAAATCCGTCAGCAAATCCGACACCA
ACAAAAGAATCTGATCGAAATGGGATCCACCCTCAGAGACACCGTGGTTCTCGTTA
ACCTGCATTCCGCCATGATTAACTTAACAATCCAAGAATTGGGTTTCGCAATCCGAG
ATCTATACGAAGAAGAGAATACTCGACAGCCAATTCAATGGTTGATCCAAGACTTGT
TGCGTAGCACCAGCGGCAGCATCAGTGATCTTATCCAAGGACAAATTCCTTCATATC
TGGTAACTCCAACACTGATAAAAAATGTGCTACAAGCAGCATCCCCAGAAACAATC
AGTCCAACACAAGTACAAATTGCCTTTAACATGGGTACAGCCACACCCATACATGTG
GACCCCCAAACCAGACAGCTGGCATTTTTGTTAAATTTACCACTGGTCAAACCCGAA
AATGTATTTCGATTGCGGACAGTATTGAACGTGGGTACATGGAATAATGACATTTAC
TCCAAGATGGAAACCCCCTTAGTGGTGGCCTACCAGGAAGAGGGCCCCCTTAAGTA
CCTAGTGCCTGACCTAAACATGTGCCAAATAATGAAAGAGGTTCATTGGCTCTGCCC
AGGGGAACCATTCCTAATCAATACAGCAGAAATAATGTGCGGTCTGTCAACCCTAA
GCTCCAGGGAGACTTGCAAGTTCAAGCTAAGTAAACCCAAAGAAAGAACCGAAACA

```
GTGGCCATGATAGCCGATGCTCAGTGGTTAATAAGCACACCTCTAGATACGGCAACC
GTTTCCACCGATCAACACTCCGTAACCATGAAAGTTCCGATTCCAAGCCACGTCGCA
CTGGTCACAGTACCAAGAGGAACATTCGTACACATAGGGGACAAGGTATTATATAA
TTTAGACAAAAGCCTACATCAAGCAGAGGTAGAAATAGTAAACACATTCCGAAGCC
ACAAGTTTGAAATAAGCAGCACTCTGTCTAACCTGTTGAACAGTAAAGAATTTCATA
CGGTGAAATTAGCCATCAGCGAGTCAGAGATTAACGTAACAGATCTCACTTGGGCG
ACCGAAGAGGCCGGATATCAGCGGAATCAACAGATGGGACTGGGAATCACAGGGG
CAACAGGAATGATCATACTAGCCATGGCTATCCTAACCCTAGCATACCATATGGCCA
AAATGAATGCCGAACTTAAAAGACTAATCGCAGGCACAAGCCACATACCATTCGAG
CCCAGGAACAAAGGTAGAGTCAAGGCCTGA
> Loki-Pma consensus sequence Env protein
MAALLLTLLLTCFAKLEGSQDLLKPGPESGVMLRDSSGFLVTETKIVTQQVYVSLEPAV
VIEKHFGKIPPSTTRATRNWYHDFLRYSEQNVGTILKQLQRTMILPPTTVNTTRSKRSIGL
GGAFGILSGFSSLITLGVSISNAVKIGHLTQGVENLQLDMHQIRQQIRHQQKNLIEMGSTL
RDTVVLVNLHSAMINLTIQELGFAIRDLYEEENTRQPIQWLIQDLLRSTSGSISDLIQGQIP
SYLVTPTLIKNVLQAASPETISPTQVQIAFNMGTATPIHVDPQTRQLAFLLNLPLVKPENV
FRLRTVLNVGTWNNDIYSKMETPLVVAYQEEGPLKYLVPDLNMCQIMKEVHWLCPGEP
FLINTAEIMCGLSTLSSRETCKFKLSKPKERTETVAMIADAQWLISTPLDTATVSTDQHSV
TMKVPIPSHVALVTVPRGTFVHIGDKVLYNLDKSLHQAEVEIVNTFRSHKFEISSTLSNLL
NSKEFHTVKLAISESEINVTDLTWATEEAGYQRNQQMGLGITGATGMIILAMAILTLAYH
MAKMNAELKRLIAGTSHIPFEPRNKGRVKA*
>Loki-Ape consensus sequence
TTGGGCGACCTGGTCCGTCAGGGCAACGCGCGCCGCATCCACACGTCAGTGATGCTT
GGTAGCTGTGTCAGCATTCTTGCATTGTTGGACTCTGGCTCTGAAATAAGCCTGATG
AGTTCCGACACTTTCTCAAAGGTGGCGAAAGCGATGCACCTCATTGGCAAACCCCTT
CAAATCGAGCCCTGCGACGTTAGCATCACCAGCTACACGCAGGACAGGTCACCCAT
CACGAAGCGCGCTTGGGTTGACATAACTTTCAAAGAAATGTCTCTTGTACACCCCAT
TTACATCTGCCTTCTTGACACCGAACCTCTGCTTATTGGCCAAGATCTCCTAGATCGA
CTGGCACCTCTCATCGACTGCCATCAGGGTCATATCTGGGCTCAGGTGAAAACCCCA
AAGCCACTTTGCCCCTCTAACAATGGGCCTGACGGACATGTCGCCGTCATTCAAGAG
ACTCACGGGGATCTGCCTACGCTCGTGGCCTCTCCCAGGCCAACCGGCGACCCCGCC
TCTTTCACGACCTCCTCCCCTCTGTGGGAGGATAGTTCAAACCCTCTCCGCAGTCATG
AGTCTTTCCTCTGTTCCTTAGCAAACTCCAACCAGGAGCTGTACAGTCGCAGAGTGG
TCGGGGAAGTCCAGCTCAACGGTCTTCTCATATCGGACGCCCGTCTGGCGCTCTGGT
TGGAGAGAACCGCAGTCAGCCAGGACCTGTACCACGAACTTCAAAAGAAAGACCCA
CCCCCAGTTTTCTCCCAACGTAGCCACCGTGTCCTGTCGGCAACACACCCACAAACC
CTCCTCAAGGCCGTCGGTGTATGTGCCCTCTTGGTACTGATTGGGAGAAAAAGGGTG
TTTCATTTTGTCAGTGTGGTACCCGACCTCCATGCCCCGTTCTTCATGGGGGCGGACC
TTCTGGTCAGGCTCGGGGCCCAGTTAGACACGGTCAATCAGGTTTTGTGGTCACGAG
CCAACGCAGAAGCTCATTCCCTAACGACCGAACCACAGCAAATGACTTCGGGGCAG
ACCATTCCCCAAGCCTGCCGGGTAATGAGTGAAACGGTGGTGACCATCCCAGCACG
GACAGCCGGTGTTCCCATCCGTCTTGCAATCTGTCAGGGCCAGAAGATACCTGGCCC
TCGAGCTTTCTTCCAGCCGTTGCCCGATTTTGGGAAAATTAACTTGGCCGTCTGTGGC
ACGCCCCTGCTTGAGTTGGACAATCGTTCAGCCTACCTCCTTGTTCAGAATCTTACGC
AAGCCCCGATCCAGGTTGCAGCACGGCAACTTCTGGGGATGCTGATTGACCACTCTT
TTCATGATTTTGAACTGACCATTCCGGTCATAGGGGAGATGCCCCCGTCTTTGGACA
```

CCAATCTTGACGGTGTTCTCTTCACTTACCCGACTAAAATGATCACCGTTGCCAGTG
ACGAAACGCTACAGAACGAAACAATCTGCAGCACAAGTTTAGGCAAAAGAGGGGA
CATGGTTGTGTACGCCCTGGTGGCACAACCAGGAGGGTCCCCAGATACTGGTGACA
AAAGACAGGAGTCACCAGAGGAACCGTACCCTGGGTTCGAGTCTGAGATTGAACAG
CAGCTGGTTAAAGCAGATGCCCTGACAACAGAGCAGCAGCGAAACGAGCTCCGGAA
ACTATTTCATGAGTTCCAGCCAATCTTCTCCAAAGACTCCCAAGATTGCGGAGTCAC
GGACATCCACACTGTTCGCATCCCCACTGATCCAGGCGCAGCACCCACCTTTGTACG
CCAGTACAAAATCCCGTTGGCCGCCTATGAGTCGATCCAAGAGATTCTCGACAAGCT
GTTGGAAAAACGGATTGTCAGGGAATGCAACTCCACATACAATTCACCCATTTGGCT
GATGTTAAAGTGCCCATATTATGAAAAAAACACTTTTTCTGGGATTTGGGGTGTTCT
TTTGTGTCTCTGGTGCTTCCACACACATACAAACTTTGAAAAAAATCCATCCATGCT
GTTTTGNAGTGAGATATGGTTTCTGAATGTGTCCTGCCTTCAGTCTCCTAGTGAGCTG
TTCAAAATCGGCTCGGACTGTGACGTCACAGTCCGAAATGAGCTGGCTAACCACAA
CCGTTANGCTCGTAGCGTTAGCATGCTAACGCTAATGCTAACGCTAGCATGCTACGT
CGTTCTCAATAGCAAAGCACTGCTACAACACACACAAGTTCACCATAATCTACAAA
AGAACTACTTACATGTGCGCCCTCATTTAGAAGTCTCCCAGCTAATCCTGCCTTGTA
ACTGNCCAAAGTTGGAGAAACAGCCTTTCTTTTACTGTCTCGAGTTAGCTAGCTGNC
ATGATCTACATCTGAGCTACTGNGCATGTGCGAGTGCAATCAAAGATAGTTANGAA
GAAGAAGATGAAAAGAGGTCTCACTCTGTAGCTAAAACAGAAACCAGGTGNAAAG
AGGATCTGCAGCAGTGNGAGAGAGCTGTGCAGTACAACAAAAATATGGTGTTTTTT
GAAAATCAAACCATGTANACCTATTCTGGTACAACCTTAAAATACAGTTATGAACCT
GAAAATGAGCATAATATGGGCGCTTTAAAACCGACCGGGAAATGGCGTCTCACTAT
TGACTATCGGCCATTAAACAAACAGGTCCCACTCTCCCGTTGGCCCATGATCCACTT
GGATCAGGAACTGGCCAAGGTCAAAGAGGCCCGTTTCTTTTCAACGGTAGATGTGG
CAAACGGTTTCTGGACCATGACGGTTGACCCGGTTGACCAGTACAAGCTGGCTTTCT
CTTTCTGCAATCGCCAGTACACCTGGAACCGCTGCCCGTTCGGATATTCAAACTCCC
CAGCTGAGTTCAACATTTTCCTCCACAAGGCAATGAGCGATGCCTCAGCCCGTGGAA
ACCTCATCTACGTGGACGACATCCTCATGAGGAGTCGGTCCTTCGATGACCACCTGG
CCGAGATCCGTCACGTTCTTGGCCAGCTTGCGGCCGCAGGAGCAAAACTGGCAATC
GCCAAGGGCCAGTGGTGCCGCACCAAGGTAGAGTACGTGGGACTGACCGTGGGTGC
CAATGGTATTGAGCCTAAAGCAGAAAGGATTCGAGCTATCCAAGACATCATGGCCC
CTGTGAACGTGTCAAAACTCCAGAGCTTTCTGGGAGTCTGTAACTATTCCCGACAGT
TCATCGCAGACTACGCAGAAATAGCCAGACCCCTCACAGAACTCCTCGGCAAAGAC
AAACCGTTCCTGTGGGGTGAACCCCAAGAGCAGGCCTTCCGCCAAATGAAGGCAGA
ACTCTGCTCCGCTCCCTGTCTGGCCTACCCGGACAAGGACAAGGAGTTCAACCTTGA
GGCGAGTTTCTCCACCCAGTGCCTTAGTGCCGCCTTGGCCCAGAAATATGATGCAGA
CAAACGGGTCGTGGCTTACGCCAGCCGACCCCTTAGCAGCGTTGATATGAAGTTTTC
CGACTGCGAGAAAGCCCTTCTGGCCACAGTGTGGGCGGTGGAACACTTCCACAGCT
ATGTAGGGGGGCAAAAGGTGGTCATTGAGACCTGTCACCAACTTGTCACCTTCCTGA
ACAGCCAGCGGCTTCGGGAAGGGCGGGTGTCCAATAGCCGCATCGCCTCCTGGATG
ATGGCGTTGCAAGGCTATGACATCGAGGTAAAGTACGCTCAGAACCACAAGATGGC
ACTGGGTCAAGGCCTGGCTGAGTGTCAACACTGCGACTGTGATGAGCAGCCAAGTC
TCCAGTCCCTACTGGTCACCACCCCGTCACTCCCGTCCAATCATCATTATTATGATGA
AATGTTTNGTCAAGGCCTCCTTGAAGTCTATGTGGATGGATGCTCATTCCACCACGA
AAGTCAGACTCAAGCTGGTGTGGGCATTTACTGGGTCAACTACCGCAGTGAGGAAC
CCAGTCACTACCAACTTGGTCCCAAGACAAGCCAGTACGCAGAAATCGCTGCAGGG

TTCATTGCGCTTCAACAAGCGGCCAGACAAAAGATCAGCCAACTGGTGATCTGCTCT
GACTCCAACTATGCCCGCTACAGTTTCGTTGCACACTTTCCTGCGTGGAAAGAGAAA
GGTATGCTGAATGCCCGCAACAAAGAGATCAAGCACGCAAAACTTTTCCTTGCATGC
GACCGACTGGTCTCACAGCAAGGAATGACCGTGTACTGGAGAAAGGTCAAGGGCCA
TTCCCAAATCTCTGGTGGGGATAAGGAAGGAAATGATGAAGCTGACCGCCTGGCCA
AACTGGGCGCAAGTCAAGGTACGCCCTGGGAGTTTCAGGAGGAGTGGCTTCCAGTC
CCCCAAACTCACGCTGTGAACGCAATCACTCGCCGTCAGGCCAGAGAGAGACGAGA
AAACTCTCAAGCCTGTGCGCAAACCCTAAATCTGTGTCGCCAGCCTGAGGACACAG
ACCTAATCACCATGCAGGAGAGGGACCCAGCTATTTGGGCCGTACGGCAGCTGCTA
GCAAACCCCTCACCTCAAAACACCACTCAATCCTCACCAGGTGAGCCAGAGGAACT
ACAAGCCTTCCGCCTCAACCCGCAGTCTTTCAGGATTGAGAAGGGTCTGTTGGTCTA
CTCTGATGGTGATCAGGGACCGACACGGTGGGTCGTCCCCACTGACCACAGGGGGG
TGATGTTAGCACACGCTCACGACTCTCCAGTGGGGGGCCATCGAGGCCCCCAACCC
ACACTGGAAACCCTCCGACAAGTTGCCTATTGGCCGTCCATGAGCCGCGATGTCAAG
GCTTACGTTAAGGGATGTTTGGTCTGCTGCCAGTTCCAGCCATCCCGGCCATTGGGT
AGAGCCCCACTCCAGAAAAGGGGGGTCACGTTCCCCTGGTCCCACCTCCAAGTGGA
CTGGGTGGGCCCCGTACCCAGATCGTCAAGAGGAAATAAATACCTCCTCACGGTCA
CATGCGCTTTTACAAAATGGGTGGAATGCCTTCCAGCACCCAACGACACGGCAACG
ACCACAGCAGTCTTGCTACTAAACCATGTTTTCAGTCGTTGGGGACTACCACTCTCC
GTCAATTCGGATCAAGGCACCCATTTCACGTCCGGGGTCATGACTGCAATGTATGAC
ATTCTGGGAGTAAAAGCAAGATTCCACATCCCATATCACCCCCAATCATCAGGCCAA
GTGGAACGCGCTAACCGGACCATTGTCAACATGCTCAAGAAGTATGTCAATGGCAG
CGCTAGGGATTGGGACGTGAAACTTCCCCTGGTACTGATGGCCATCAGGTCCACTCC
ACACCGCACCACGGGGATCACTCCATTCGAGATGATGACCGGGAGAGAGATGACGC
TGCCACTGCATCTCTTGTACCACCCTGAAGATGTCAGTGTCGCAACGGCCTACACTG
CCACTCAATACGTCTCAAAATTGCGTGAACATCTTCAGTCCACGTTTGCCTGGGCCC
AAGAAAACCTGGAGACAAGGGTAAAGGGGGCCAAGGCGTACTACGACCGAAAGAC
AACTCCCCACGAGTATGAAATAGGTGACAAAGTCTTCTATTTCAGGTTCACCAAACC
GGTGAGAATCTCTCGAAAGTTCCTACCGAGCTGGTCAGGCCAAGTTGAGATTGTGGG
AAAACGTTCCCCGGTGGCGTACCGGATCAGATTGGCTAAGCCGAAACAGCCACCAA
CCTACAAATGGGTCCATTCAAACCAAATCAAGCCATACGAGACGCCACCCCTCGTTT
CAGAAAGGGAAAACTCCCCATCAGTAGAAAGGGGGAACAGCCCGACACAGACATA
GGATATACTACAGGGCCCATAGGAATAACTGCATGACCAATCACTTCCGCGAGCAT
GCACTAACCCCTAAAAGAAATCCCCTGTCCACCTAGTAATCAAAACTCTGTTCTGTT
GCAGGATGGAACTACTGTGGGTCCTTGGCATTCTTCTCGGGCTATCAACCCCTCTCC
CAACTGAGGTGGTTGAATCAGGCCCGTCCACCGGGATCGTCCTCCAGGGGGCGCCC
GGACTGCTAATCACCCGTGGTCTCTTCACCCAGAGAGTCTATGTACGACTAGACCCC
TGGGATGTCTACCGCAAACACATCGTTCTACCACCCCAATTGACCGAGGGACGGTTA
AGTGGGATCCAGACCCATGACACAGTAGAACGCACTAAACAGACCACTGCCCACAT
TCTGGAACAACTGCAAAAATTCATAGTCACAGAAGAAGACCTCAGTGGCCAGAAAC
GACCTAAACGTTTCTTAGGCGGCCTGCTCGTCGCAGCCTCAGCCATCGGGTCATTAT
TTTCTATTGGCCTGTCCGCGGCAAACTCAGTTAGCCTCGGCGCACTCCAGAGACACA
TGGGGGAACTCGACCAAAAAATGCCAGAGATCCAACAGAGACTCCTCCTCCAACAA
GAACAGTTACAAAACCTGGGAAAAACTTTCCAGGGGACCATAGTTACTGTACATTTG
CACTCTGCCCTCCTGAATAACACCATGCATGCCTTAGAGAATTTGTCCAAAGTGGTC
CAAGAAGACCTAACATACGTACGAATAGTCAGAGACCTAATGCAGGACCTTGTCAG

GGAGGTTAGCTCCTCAGTCAATAGCCTGAGCGCGGGAAAAATTCCAGCCTATCTAGT
CTCATTAGGACTCGTTGAACAAATCTTGAAATCCACCACTAGTACGGTAGTACACCC
TTCACAGGTACACCTGGCGTATAGCCTTGGCAGTGCAATCCCCATCTCAGTAGACTC
TAAGAACTTAGAGATAGGTTTCATCCTCAATCTTCCGATTCTAGAGAGACAAAACAT
ATACCGACTAAAGTCTGTACTGAACGTGGGCTTCTGGAAAGGTAACACACATATAC
ACCTTAAGACCCCACCCACAATTGCTTTCCATGAGGACAACCCCTCACTCTACCTCA
TTCCTAACCTTAACATGTGCACTAAAACAACAAAAGACATTCACTGGGTCTGCCCGA
GCAACCCATTCGTTAGAGATGTCACTAACTACCTCTGCGGCCTGAGGGCCGACACAC
CTGAACAAAAATGTCAGGGCAGTATGTCCATAAGGGACGAGGGAACAGAAACAAG
GGTGGAACGGGCGGGTAACCGATGGCTCGTTAGTACACCCGTCACTGAAGCGTTAA
TGTCCTACGACCGTCATGACACAGCCACTAAGTTAACCTTACCTAACCAAACGGTAT
TTTTAACGGTTCCCCAAGGAGCTACCATTCATATCAATGACATTGTCCTCCATCATCT
CAATCCCGACAGACACGATGCCGAGATTGAGATCATGGACGCATTTCGAGGCCACG
ACCTCACTATTGAATACCCCCTCCAACAGCAACTTCTGGCTGAGGGGACTAAACTGG
TTCAGTTCAGCCTCAAACCTACTGGGCTTACTACCACGTTTTCCAACCATGTAGGGA
GTTCGGCATCTTACCAAGACCACCCTGTTAGCATGGTTGCCTTGGGCCTCCTTCTTAG
TGGCTGGATCATCACCGCGATAGTCGCACATTTCATGTACAGACACATACAGAATTT
GCAAACTAAACTTGACTCGATTCTGTTCGTATCACCACGTTTCACACAGCAGCCCAC
ATCGTCCCACTAG
> Loki-Ape consensus sequence *pol* gene
TTGGGCGACCTGGTCCGTCAGGGCAACGCGCGCCGCATCCACACGTCAGTGATGCTT
GGTAGCTGTGTCAGCATTCTTGCATTGTTGGACTCTGGCTCTGAAATAAGCCTGATG
AGTTCCGACACTTTCTCAAAGGTGGCGAAAGCGATGCACCTCATTGGCAAACCCCTT
CAAATCGAGCCCTGCGACGTTAGCATCACCAGCTACACGCAGGACAGGTCACCCAT
CACGAAGCGCGCTTGGGTTGACATAACTTTCAAAGAAATGTCTCTTGTACACCCCAT
TTACATCTGCCTTCTTGACACCGAACCTCTGCTTATTGGCCAAGATCTCCTAGATCGA
CTGGCACCTCTCATCGACTGCCATCAGGGTCATATCTGGGCTCAGGTGAAAACCCCA
AAGCCACTTTGCCCCTCTAACAATGGGCCTGACGGACATGTCGCCGTCATTCAAGAG
ACTCACGGGGATCTGCCTACGCTCGTGGCCTCTCCCAGGCCAACCGGCGACCCCGCC
TCTTTCACGACCTCCTCCCCTCTGTGGGAGGATAGTTCAAACCCTCTCCGCAGTCATG
AGTCTTTCCTCTGTTCCTTAGCAAACTCCAACCAGGAGCTGTACAGTCGCAGAGTGG
TCGGGGAAGTCCAGCTCAACGGTCTTCTCATATCGGACGCCCGTCTGGCGCTCTGGT
TGGAGAGAACCGCAGTCAGCCAGGACCTGTACCACGAACTTCAAAAGAAAGACCCA
CCCCCAGTTTTCTCCCAACGTAGCCACCGTGTCCTGTCGGCAACACACCCACAAACC
CTCCTCAAGGCCGTCGGTGTATGTGCCCTCTTGGTACTGATTGGGAGAAAAAGGGTG
TTTCATTTTGTCAGTGTGGTACCCGACCTCCATGCCCCGTTCTTCATGGGGGCGGACC
TTCTGGTCAGGCTCGGGGCCCAGTTAGACACGGTCAATCAGGTTTTGTGGTCACGAG
CCAACGCAGAAGCTCATTCCCTAACGACCGAACCACAGCAAATGACTTCGGGGCAG
ACCATTCCCCAAGCCTGCCGGGTAATGAGTGAAACGGTGGTGACCATCCCAGCACG
GACAGCCGGTGTTCCCATCCGTCTTGCAATCTGTCAGGGCCAGAAGATACCTGGCCC
TCGAGCTTTCTTCCAGCCGTTGCCCGATTTTGGGAAAATTAACTTGGCCGTCTGTGGC
ACGCCCCTGCTTGAGTTGGACAATCGTTCAGCCTACCTCCTTGTTCAGAATCTTACGC
AAGCCCCGATCCAGGTTGCAGCACGGCAACTTCTGGGGATGCTGATTGACCACTCTT
TTCATGATTTTGAACTGACCATTCCGGTCATAGGGGAGATGCCCCCGTCTTTGGACA
CCAATCTTGACGGTGTTCTCTTCACTTACCCGACTAAAATGATCACCGTTGCCAGTG
ACGAAACGCTACAGAACGAAACAATCTGCAGCACAAGTTTAGGCAAAAGAGGGGA

```
CATGGTTGTGTACGCCCTGGTGGCACAACCAGGAGGGTCCCCAGATACTGGTGACA
AAAGACAGGAGTCACCAGAGGAACCGTACCCTGGGTTCGAGTCTGAGATTGAACAG
CAGCTGGTTAAAGCAGATGCCCTGACAACAGAGCAGCAGCGAAACGAGCTCCGGAA
ACTATTTCATGAGTTCCAGCCAATCTTCTCCAAAGACTCCCAAGATTGCGGAGTCAC
GGACATCCACACTGTTCGCATCCCCACTGATCCAGGCGCAGCACCCACCTTTGTACG
CCAGTACAAAATCCCGTTGGCCGCCTATGAGTCGATCCAAGAGATTCTCGACAAGCT
GTTGGAAAAACGGATTGTCAGGGAATGCAACTCCACATACAATTCACCCATTTGGCT
GATGTTAAAGTGCCCATATTATGAAAAAAACACTTTTTCTGGGATTTGGGGTGTTCT
TTTGTGTCTCTGGTGCTTCCACACACATACAAACTTTGAAAAAAATCCATCCATGCT
GTTTTGNAGTGAGATATGGTTTCTGAATGTGTCCTGCCTTCAGTCTCCTAGTGAGCTG
TTCAAAATCGGCTCGGACTGTGACGTCACAGTCCGAAATGAGCTGGCTAACCACAA
CCGTTANGCTCGTAGCGTTAGCATGCTAACGCTAATGCTAACGCTAGCATGCTACGT
CGTTCTCAATAGCAAAGCACTGCTACAACACACACAAGTTCACCATAATCTACAAA
AGAACTACTTACATGTGCGCCCTCATTTAGAAGTCTCCCAGCTAATCCTGCCTTGTA
ACTGNCCAAAGTTGGAGAAACAGCCTTTCTTTTACTGTCTCGAGTTAGCTAGCTGNC
ATGATCTACATCTGAGCTACTGNGCATGTGCGAGTGCAATCAAAGATAGTTANGAA
GAAGAAGATGAAAAGAGGTCTCACTCTGTAGCTAAAACAGAAACCAGGTGNAAAG
AGGATCTGCAGCAGTGNGAGAGAGCTGTGCAGTACAACAAAAATATGGTGTTTTT
GAAAATCAAACCATGTANACCTATTCTGGTACAACCTTAAAATACAGTTATGAACCT
GAAAATGAGCATAATATGGGCGCTTTAAAACCGACCGGGAAATGGCGTCTCACTAT
TGACTATCGGCCATTAAACAAACAGGTCCCACTCTCCCGTTGGCCCATGATCCACTT
GGATCAGGAACTGGCCAAGGTCAAAGAGGCCCGTTTCTTTTCAACGGTAGATGTGG
CAAACGGTTTCTGGACCATGACGGTTGACCCGGTTGACCAGTACAAGCTGGCTTTCT
CTTTCTGCAATCGCCAGTACACCTGGAACCGCTGCCCGTTCGGATATTCAAACTCCC
CAGCTGAGTTCAACATTTTCCTCCACAAGGCAATGAGCGATGCCTCAGCCCGTGGAA
ACCTCATCTACGTGGACGACATCCTCATGAGGAGTCGGTCCTTCGATGACCACCTGG
CCGAGATCCGTCACGTTCTTGGCCAGCTTGCGGCCGCAGGAGCAAAACTGGCAATC
GCCAAGGGCCAGTGGTGCCGCACCAAGGTAGAGTACGTGGGACTGACCGTGGGTGC
CAATGGTATTGAGCCTAAAGCAGAAAGGATTCGAGCTATCCAAGACATCATGGCCC
CTGTGAACGTGTCAAAACTCCAGAGCTTTCTGGGAGTCTGTAACTATTCCCGACAGT
TCATCGCAGACTACGCAGAAATAGCCAGACCCCTCACAGAACTCCTCGGCAAAGAC
AAACCGTTCCTGTGGGGTGAACCCCAAGAGCAGGCCTTCCGCCAAATGAAGGCAGA
ACTCTGCTCCGCTCCCTGTCTGGCCTACCCGGACAAGGACAAGGAGTTCAACCTTGA
GGCGAGTTTCTCCACCCAGTGCCTTAGTGCCGCCTTGGCCCAGAAATATGATGCAGA
CAAACGGGTCGTGGCTTACGCCAGCCGACCCCTTAGCAGCGTTGATATGAAGTTTTC
CGACTGCGAGAAAGCCCTTCTGGCCACAGTGTGGGCGGTGGAACACTTCCACAGCT
ATGTAGGGGGGCAAAAGGTGGTCATTGAGACCTGTCACCAACTTGTCACCTTCCTGA
ACAGCCAGCGGCTTCGGGAAGGGCGGGTGTCCAATAGCCGCATCGCCTCCTGGATG
ATGGCGTTGCAAGGCTATGACATCGAGGTAAAGTACGCTCAGAACCACAAGATGGC
ACTGGGTCAAGGCCTGGCTGAGTGTCAACACTGCGACTGTGATGAGCAGCCAAGTC
TCCAGTCCCTACTGGTCACCACCCCGTCACTCCCGTCCAATCATCATTATTATGATGA
AATGTTTNGTCAAGGCCTCCTTGAAGTCTATGTGGATGGATGCTCATTCCACCACGA
AAGTCAGACTCAAGCTGGTGTGGGCATTTACTGGGTCAACTACCGCAGTGAGGAAC
CCAGTCACTACCAACTTGGTCCCAAGACAAGCCAGTACGCAGAAATCGCTGCAGGG
TTCATTGCGCTTCAACAAGCGGCCAGACAAAAGATCAGCCAACTGGTGATCTGCTCT
GACTCCAACTATGCCCGCTACAGTTTCGTTGCACACTTTCCTGCGTGGAAAGAGAAA
```

GGTATGCTGAATGCCCGCAACAAAGAGATCAAGCACGCAAAACTTTTCCTTGCATGC
GACCGACTGGTCTCACAGCAAGGAATGACCGTGTACTGGAGAAAGGTCAAGGGCCA
TTCCCAAATCTCTGGTGGGGATAAGGAAGGAAATGATGAAGCTGACCGCCTGGCCA
AACTGGGCGCAAGTCAAGGTACGCCCTGGGAGTTTCAGGAGGAGTGGCTTCCAGTC
CCCCAAACTCACGCTGTGAACGCAATCACTCGCCGTCAGGCCAGAGAGAGACGAGA
AAACTCTCAAGCCTGTGCGCAAACCCTAAATCTGTGTCGCCAGCCTGAGGACACAG
ACCTAATCACCATGCAGGAGAGGGACCCAGCTATTTGGGCCGTACGGCAGCTGCTA
GCAAACCCCTCACCTCAAAACACCACTCAATCCTCACCAGGTGAGCCAGAGGAACT
ACAAGCCTTCCGCCTCAACCCGCAGTCTTTCAGGATTGAGAAGGGTCTGTTGGTCTA
CTCTGATGGTGATCAGGGACCGACACGGTGGGTCGTCCCCACTGACCACAGGGGGG
TGATGTTAGCACACGCTCACGACTCTCCAGTGGGGGGCCATCGAGGCCCCCAACCC
ACACTGGAAACCCTCCGACAAGTTGCCTATTGGCCGTCCATGAGCCGCGATGTCAAG
GCTTACGTTAAGGGATGTTTGGTCTGCTGCCAGTTCCAGCCATCCCGGCCATTGGGT
AGAGCCCCACTCCAGAAAAGGGGGGTCACGTTCCCCTGGTCCCACCTCCAAGTGGA
CTGGGTGGGCCCCGTACCCAGATCGTCAAGAGGAAATAAATACCTCCTCACGGTCA
CATGCGCTTTTACAAAATGGGTGGAATGCCTTCCAGCACCCAACGACACGGCAACG
ACCACAGCAGTCTTGCTACTAAACCATGTTTTCAGTCGTTGGGGACTACCACTCTCC
GTCAATTCGGATCAAGGCACCCATTTCACGTCCGGGGTCATGACTGCAATGTATGAC
ATTCTGGGAGTAAAAGCAAGATTCCACATCCCATATCACCCCCAATCATCAGGCCAA
GTGGAACGCGCTAACCGGACCATTGTCAACATGCTCAAGAAGTATGTCAATGGCAG
CGCTAGGGATTGGGACGTGAAACTTCCCCTGGTACTGATGGCCATCAGGTCCACTCC
ACACCGCACCACGGGGATCACTCCATTCGAGATGATGACCGGGAGAGAGATGACGC
TGCCACTGCATCTCTTGTACCACCCTGAAGATGTCAGTGTCGCAACGGCCTACACTG
CCACTCAATACGTCTCAAAATTGCGTGAACATCTTCAGTCCACGTTTGCCTGGGCCC
AAGAAAACCTGGAGACAAGGGTAAAGGGGGCCAAGGCGTACTACGACCGAAAGAC
AACTCCCCACGAGTATGAAATAGGTGACAAAGTCTTCTATTTCAGGTTCACCAAACC
GGTGAGAATCTCTCGAAAGTTCCTACCGAGCTGGTCAGGCCAAGTTGAGATTGTGGG
AAAACGTTCCCCGGTGGCGTACCGGATCAGATTGGCTAAGCCGAAACAGCCACCAA
CCTACAAATGGGTCCATTCAAACCAAATCAAGCCATACGAGACGCCACCCCTCGTTT
CAGAAAGGGAAAACTCCCCATCAGTAGAAAGGGGGAACAGCCCGACACAGACATA
G
> Loki-Ape consensus sequence Pol protein
LGDLVRQGNARRIHTSVMLGSCVSILALLDSGSEISLMSSDTFSKVAKAMHLIGKPLQIEP
CDVSITSYTQDRSPITKRAWVDITFKEMSLVHPIYICLLDTEPLLIGQDLLDRLAPLIDCHQ
GHIWAQVKTPKPLCPSNNGPDGHVAVIQETHGDLPTLVASPRPTGDPASFTTSSPLWEDS
SNPLRSHESFLCSLANSNQELYSRRVVGEVQLNGLLISDARLALWLERTAVSQDLYHEL
QKKDPPPVFSQRSHRVLSATHPQTLLKAVGVCALLVLIGRKRVFHFVSVVPDLHAPFFM
GADLLVRLGAQLDTVNQVLWSRANAEAHSLTTEPQQMTSGQTIPQACRVMSETVVTIP
ARTAGVPIRLAICQGQKIPGPRAFFQPLPDFGKINLAVCGTPLLELDNRSAYLLVQNLTQ
APIQVAARQLLGMLIDHSFHDFELTIPVIGEMPPSLDTNLDGVLFTYPTKMITVASDETLQ
NETICSTSLGKRGDMVVYALVAQPGGSPDTGDKRQESPEEPYPGFESEIEQQLVKADAL
TTEQQRNELRKLFHEFQPIFSKDSQDCGVTDIHTVRIPTDPGAAPTFVRQYKIPLAAYESI
QEILDKLLEKRIVRECNSTYNSPIWLMLKCPYYEKNTFSGIWGVLLCLWCFHTHTNFEK
NPSMLFXSEIWFLNVSCLQSPSELFKIGSDCDVTVRNELANHNRXARSVSMLTLMLTLA
CYVVLNSKALLQHTQVHHNLQKNYLHVRPHLEVSQLILPCNXPKLEKQPFFYCLELASX
HDLHLSYXACASAIKDSXEEEDEKRSHSVAKTETRXKEDLQQXERAVQYNKNMVFFEN

QTMXTYSGTTLKYSYEPENEHNMGALKPTGKWRLTIDYRPLNKQVPLSRWPMIHLDQE
LAKVKEARFFSTVDVANGFWTMTVDPVDQYKLAFSFCNRQYTWNRCPFGYSNSPAEF
NIFLHKAMSDASARGNLIYVDDILMRSRSFDDHLAEIRHVLGQLAAAGAKLAIAKGQW
CRTKVEYVGLTVGANGIEPKAERIRAIQDIMAPVNVSKLQSFLGVCNYSRQFIADYAEIA
RPLTELLGKDKPFLWGEPQEQAFRQMKAELCSAPCLAYPDKDKEFNLEASFSTQCLSAA
LAQKYDADKRVVAYASRPLSSVDMKFSDCEKALLATVWAVEHFHSYVGGQKVVIETC
HQLVTFLNSQRLREGRVSNSRIASWMMALQGYDIEVKYAQNHKMALGQGLAECQHCD
CDEQPSLQSLLVTTPSLPSNHHYYDEMFXQGLLEVYVDGCSFHHESQTQAGVGIYWVN
YRSEEPSHYQLGPKTSQYAEIAAGFIALQQAARQKISQLVICSDSNYARYSFVAHFPAWK
EKGMLNARNKEIKHAKLFLACDRLVSQQGMTVYWRKVKGHSQISGGDKEGNDEADRL
AKLGASQGTPWEFQEEWLPVPQTHAVNAITRRQARERRENSQACAQTLNLCRQPEDTD
LITMQERDPAIWAVRQLLANPSPQNTTQSSPGEPEELQAFRLNPQSFRIEKGLLVYSDGD
QGPTRWVVPTDHRGVMLAHAHDSPVGGHRGPQPTLETLRQVAYWPSMSRDVKAYVK
GCLVCCQFQPSRPLGRAPLQKRGVTFPWSHLQVDWVGPVPRSSRGNKYLLTVTCAFTK
WVECLPAPNDTATTTAVLLLNHVFSRWGLPLSVNSDQGTHFTSGVMTAMYDILGVKAR
FHIPYHPQSSGQVERANRTIVNMLKKYVNGSARDWDVKLPLVLMAIRSTPHRTTGITPF
EMMTGREMTLPLHLLYHPEDVSVATAYTATQYVSKLREHLQSTFAWAQENLETRVKG
AKAYYDRKTTPHEYEIGDKVFYFRFTKPVRISRKFLPSWSGQVEIVGKRSPVAYRIRLAK
PKQPPTYKWVHSNQIKPYETPPLVSERENSPSVERGNSPTQT*
> Loki-Ape consensus sequence *env* gene
CTGTTCTGTTGCAGGATGGAACTACTGTGGGTCCTTGGCATTCTTCTCGGGCTATCAA
CCCCTCTCCCAACTGAGGTGGTTGAATCAGGCCCGTCCACCGGGATCGTCCTCCAGG
GGGCGCCCGGACTGCTAATCACCCGTGGTCTCTTCACCCAGAGAGTCTATGTACGAC
TAGACCCCTGGGATGTCTACCGCAAACACATCGTTCTACCACCCCAATTGACCGAGG
GACGGTTAAGTGGGATCCAGACCCATGACACAGTAGAACGCACTAAACAGACCACT
GCCCACATTCTGGAACAACTGCAAAAATTCATAGTCACAGAAGAAGACCTCAGTGG
CCAGAAACGACCTAAACGTTTCTTAGGCGGCCTGCTCGTCGCAGCCTCAGCCATCGG
GTCATTATTTTCTATTGGCCTGTCCGCGGCAAACTCAGTTAGCCTCGGCGCACTCCAG
AGACACATGGGGGAACTCGACCAAAAAATGCCAGAGATCCAACAGAGACTCCTCCT
CCAACAAGAACAGTTACAAAACCTGGGAAAAACTTTCCAGGGGACCATAGTTACTG
TACATTTGCACTCTGCCCTCCTGAATAACACCATGCATGCCTTAGAGAATTTGTCCA
AAGTGGTCCAAGAAGACCTAACATACGTACGAATAGTCAGAGACCTAATGCAGGAC
CTTGTCAGGGAGGTTAGCTCCTCAGTCAATAGCCTGAGCGCGGGAAAAATTCCAGCC
TATCTAGTCTCATTAGGACTCGTTGAACAAATCTTGAAATCCACCACTAGTACGGTA
GTACACCCTTCACAGGTACACCTGGCGTATAGCCTTGGCAGTGCAATCCCCATCTCA
GTAGACTCTAAGAACTTAGAGATAGGTTTCATCCTCAATCTTCCGATTCTAGAGAGA
CAAAACATATACCGACTAAAGTCTGTACTGAACGTGGGCTTCTGGAAAGGTAACAC
ACATATACACCTTAAGACCCCACCCACAATTGCTTTCCATGAGGACAACCCCTCACT
CTACCTCATTCCTAACCTTAACATGTGCACTAAAACAACAAAAGACATTCACTGGGT
CTGCCCGAGCAACCCATTCGTTAGAGATGTCACTAACTACCTCTGCGGCCTGAGGGC
CGACACACCTGAACAAAAATGTCAGGGCAGTATGTCCATAAGGGACGAGGGAACAG
AAACAAGGGTGGAACGGGCGGGTAACCGATGGCTCGTTAGTACACCCGTCACTGAA
GCGTTAATGTCCTACGACCGTCATGACACAGCCACTAAGTTAACCTTACCTAACCAA
ACGGTATTTTTAACGGTTCCCCAAGGAGCTACCATTCATATCAATGACATTGTCCTCC
ATCATCTCAATCCCGACAGACACGATGCCGAGATTGAGATCATGGACGCATTTCGAG
GCCACGACCTCACTATTGAATACCCCCTCCAACAGCAACTTCTGGCTGAGGGGACTA

AACTGGTTCAGTTCAGCCTCAAACCTACTGGGCTTACTACCACGTTTTCCAACCATGT
AGGGAGTTCGGCATCTTACCAAGACCACCCTGTTAGCATGGTTGCCTTGGGCCTCCT
TCTTAGTGGCTGGATCATCACCGCGATAGTCGCACATTTCATGTACAGACACATACA
GAATTTGCAAACTAAACTTGACTCGATTCTGTTCGTATCACCACGTTTCACACAGCA
GCCCACATCGTCCCACTAG
> Loki-Ape consensus sequence Env protein
LFCCRMELLWVLGILLGLSTPLPTEVVESGPSTGIVLQGAPGLLITRGLFTQRVYVRLDP
WDVYRKHIVLPPQLTEGRLSGIQTHDTVERTKQTTAHILEQLQKFIVTEEDLSGQKRPKR
FLGGLLVAASAIGSLFSIGLSAANSVSLGALQRHMGELDQKMPEIQQRLLLQQEQLQNL
GKTFQGTIVTVHLHSALLNNTMHALENLSKVVQEDLTYVRIVRDLMQDLVREVSSSVN
SLSAGKIPAYLVSLGLVEQILKSTTSTVVHPSQVHLAYSLGSAIPISVDSKNLEIGFILNLPI
LERQNIYRLKSVLNVGFWKGNTHIHLKTPPTIAFHEDNPSLYLIPNLNMCTKTTKDIHWV
CPSNPFVRDVTNYLCGLRADTPEQKCQGSMSIRDEGTETRVERAGNRWLVSTPVTEAL
MSYDRHDTATKLTLPNQTVFLTVPQGATIHINDIVLHHLNPDRHDAEIEIMDAFRGHDLT
IEYPLQQQLLAEGTKLVQFSLKPTGLTTTFSNHVGSSASYQDHPVSMVALGLLLSGWIIT
AIVAHFMYRHIQNLQTKLDSILFVSPRFTQQPTSSH*
>Loki-Pfl consensus sequence
AATGGCTGGTGTTGGAACCTAACAGTTATGATCAAGACTGTTCTGTGAGGCCTGAAA
CTTCCCTGTGGCCAAAGAGCGGCCTGTAGGGAAGTCACCAGCCTGTGACCCCTCACT
CCTAAACAAGGAGTCTCCAAAATGGAGATTGTACCTCCAAGACATGAGGAAAATG
GCAGACTGGTGGAGGGTTTGAGAACTGCAACGACGGATTCTCACACCTTCGTGGAG
TTCGGGATTTCTGATTGGCCGCGGCTCCTTGAACGCACCACTCGCCGCCAGCAGGCG
GCGGAGGGAAGATAAAAATAGCCGGAGGTCACAGCCCGGGGTTCTTCGAGGTAATC
CATTTGCCGCAAGGAGACGCATCGTTCAGCGCTGTTCGAGAATCAACATTGTATCAT
TCGCTGGTCGAGTTAGACATTTGTATCGTTCGTGGTACAACAGGGTCCTGGTGGATA
CGAAGCCCGTATTACACCAAATCTGCAGAGGGGTGAATCAGCCCCGTCTCGAGGAA
AAGAGACAACGCGTTTTTCTTGCAGGTCGACTGGACCGTGTTTCCGCTCCACTGCCT
TGGACGGTACGGGTCGTTAATCTCTGGCGAGAGTTAACTCAGTTTTGTTCACCGGGA
AATCCGCCGGATAAGCAACGGACTGTACACCAAAAGTAAGAGGCTTTTATAGTTCT
GGGCAGATTTGAGAACTAGGGTGTTTATTAATGAGTAAAAGGTACTGCTAATTTTTG
TTTACCATTAATGTGTGATCGGACCCGCAAGGCGTCCTATGTGCTGTTGTGAAGAAT
ATCACTGATCCAGATATTGTTGAAAGTTGTGTTGAACTGTAACTGATACTCCCGCCG
AGGAGAAATTACTGTATACGCTGAGTCAGCTGCACGCTCAAAGAGTGTGCCGGGGA
ACAAACGCTCTCTACCGTTAAACAAACCTCAGATTCTGCTCTAACGCCACGTTGAGT
TAGAATCTCGAGTAAAACGAAGAGAGCGGTTAGTCCCGTTACCCCAGAGATCAAGG
TTTTAAGTGTTTGTTTCCATAAATTGTGTTTTCCCTCCTTTTGTTTCTTTCGTTTCTTTT
TACACACACACACACACACACACACACACACACACACACACGCACACATACCTCCA
AGCTAGTTAGCCACGTGGAGTAAGCTAGAGCTAACGCTTTAGCATAAACGTGTGCGT
TGCCTCGCAACCCCCACCTGAGCTTCTCCACCCCCCTTCGGGCGCACCCAGTACCAC
TACCGCCCTCTTGAGGCTAAATAGTTAGGTGTAGCTCACTGGAGTAGCCATTTTAGT
AGTTCATGGTAGACTACACTTCGACCCCCCCCCCCCTTTCACTCACTCACTCACACAC
ACACACCCACACACGCCGCACAGACACAGAAGTGTCTCTGCTGCGTTGGGAGTTTCT
TGCTGTGTTTTGATTGTGATATTGTAAAAGGGTATAGAAGTTTATTATTGAAGGATT
GGTGATTGGCTACTGATAATTGTGATGTTTTAATAAATCTGATTATACTTTAAATAGC
AGTTGCTTGTGATTATTTGTGTACTTACTGTAATGACTGCTGGCTAGACAGGTCTGTG
CTCGAAATCACCCCCTCTTCTTTAAATCCTTTCAAGGGATCTTAACAGAAATGAGAC

TGTTCCACAGTTGATCACTGGTCCCTGACTTACAGGGTGGTGCCCCGTTGAGTGTTTG
TTGATTTGATAAAGTTATTAATAACCATATTCATAACTTTATTAATATTGATAAAACA
TATTTTGATAATTTGCTAGTGATCAATCTGTACCCTACAAGAGAGAATCCAACATTT
GACGCCCCGTGTGAGCCTCTTAGAGCCAGCATTGTTGCAGTGATTATACATAATAAT
CCATAGTCTTTGAATATTAATATTAAAAGAAAAAAAAAGATTATTTATTATTTATTTT
TTAAATAAACAAATAAATTTTTCCGCAACCTGCCAAGTACATGTTCAGTACAGTCAA
TCCATAACCTTAAATTAAGTTTCTGTTTTGTAAAAGAAAAAGGGAATTAATGTTAAA
GGACAACAGTTTGGAAATTGATAAGTAAAAATTAGGTTCTTTTCGGTGAACTCGCCA
AGCATACGTTCAGCGGAAGTTGTGAATCTGTGTGTATTAGTGGCTAGGATTTTTCTTC
GTTGTTTAGACAACAACACTCTTTGCATCCTACTTAACCTTGAAATCTCTGCCTAACA
CCCTGACCCCTAAACAGTAAAACAGCTAGTTACTAGACCCAGACAAGATCTGATTAT
AGCAATTAGGAAAGCTATCAGTCAACACGTAAACCTAGGAGCCACAGGTTAAAGCC
TTCACCTGCGGGCTCGCGTCAACCTCTTCACAAGCATTCTACAGTTGCGACAAGCAA
TTGTAGCCACTGGTGATTGCCCTCTTGCTCAGCAGGTAAGAATGAGCGACCACAGTG
ACCATGAGCTAACAGAAGAAGCTCAGACCGAGTCAATAGACACAGTCATAACAGAC
ACAGTCTCTGCACTCATCCGTCTGCCCCCAGCGGACAGAGTTTTCCTGGAGGAATAC
GAGGTTGGGACCTGTGAAACCAAACTCAAGGAACTACTAGACTACGTGGATCAGCC
CACCGGGAAGCCGGAGCTCACCAAACCAGACGTGCTTGGTCAAATAGTCCTTCTCC
AGCAACGGAGATTCCACCTGCAGCTCACTGAGCTTCAGAGACACAGTCAAACTCAA
CAAGAGGCTATCCGCCGCTTGCGCCAAGAGATCCGGGACGCACACGATGCGTACTC
TGATCTCCAGGCCGAGTGTGTTCTCCTGAAGCAACTGAAAATTGCTAAACAGCAAGG
AGATGCAGGGGACGGGGACACTCAGACCAACCGGGTAAGTACAGAAGCGGCGGAA
TTCACAACGAAGACTGGCGACTCGGACGACGAGCGGGCCGCAGGGATAATTTCCTA
TCGGTCTCAACCCGCCAGCGAAGGTGTCACCCCGGATTGGAGGATCATCGACCAAG
AGAAACCCTCCTCCCAGTGGAGGGAGGAAAATGGCTCCACCATCAGAAGGGTAAGT
TATCAGCCCCCAAGGCCGAAGAACATTCCGGAGTTTCGGGTCTCAAGGAAAGAGGA
TCACTCCCCTGTTCGGTTGAAAAGAGAAGAAGTTTCATGCAGCCGAAGAGATGATTA
CTCTCCTGCTTGGCCGAAAAAGGAGGGAGTCTCCCGCAGCAGAAGAGATGACGACT
CTCCTGTTCAGCCGAGAAAGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCT
CCTGTTCGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCC
TGTTCGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCCTG
CTCGGTCAAGAAAGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCCTGTT
CGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACTACTCTCCTGTTCG
GTCAAGAAGGGAGGAACTCTCCCGCAGCAGGAGAGATGACGACTCCCCTGTCCGGT
CGAGGAGGGAGGAGCTCTCCCGTAACAGGAGAGATGAGGAGGGATGATTACTCCCC
TGCTCGGCCAGGAAGGGACCGCCACTCGGTCACAACGGGGAGGAGCCGACAAGGA
GGTGACTACCGCACCTCAGACGACTCTGATGAAAGTCTGAGGGAGAGGTCCCGTTC
CTCTGCCTGGAAGGATCTGAGTCAGCCTCAACGGAAAACCCACCGTCAACACCGGA
GAGGGAGAGGCAGCGGTACTTCCAGTGACTCTGACGCCTCACGCCACTCCCCAGCG
CAGGGCCTGCGCATGCGCCAGATTGACTCCTTCGCAAGGGAGCTAGAGCGTTTCGA
CCCCAACAGTCGAGAAACTAATGTCGACGACTATCTCCGCGATGTAGAACGGGTCC
TGTCAGACCTCCCCAATGCAACGTCGCGCGAGAAGCGCAAACTACTCTGGAAGACC
ACGGCAAGGAGCGCCCATGTCTTTATGGACGAACTCCCGCCCTCGACCTTAAACAGT
TACACAAAGCTCTGCCAAGCTTTACGAGAAGAGTTCGCCCTGTACACCGACCCAGCT
GCGGCCATCCAGGGAGCATATGCCATTGTGCATAAGAGGAATGAGGCCCCGAGGGA
ATATTATAGGCGTCTAAGGACCGTTTACTTCCAGGGCCGCCGCTCCCCAGGTCTAGA

GGAAGACCCCGCTTTCCAATCCCTGTTCGTACACAATCTCCACGATGCTGTACGGTA
CGATGTGCAGATGCATTGTAGAACACAAGAGCTCACCATGAAAGAGATTAGGATGT
TCGCACAAAGGGTGTGGGAGGCTCGAACACGACCGATCAGGAAAGGTGAGGAGGA
GCCGAGGGTACTACACATCCAGACCTCAGAGAACGCAGATCTAGCCCTCGAAGGGA
ACGAAGTTCCCCAGGGAACAGGGGACTCCAGACCACCAAACCAGGAACAACGATCC
TACCACACAGGCAGGCAAAATCGACAAAACAACTACAGAGGGAAGAAGAAGCCCT
ATGGTGGAAGAAAACCCGGTAAGTATGAGGGAGGACCTCAAAGCCAGGAAAAGGA
TGACGGTCGTAATGCCTCCCCAGCACAGCAGCAGGCCCCTCCGAAGGGACCTGAGA
CCACCTCCCAACCCGCGCGACTAAGCCAAGCGCCCGTTTCACTCACCGCCAAGGTGT
TCTTCATAGGGTGGGGAAACACGCCAGACAGTCTCCCAGAGTCTGTCACTCCTCTTC
CTCCTGCGGCTTCCAGGGTTTTCTTGGGCGACCTGGTCCATCGGGGCAACGCGCGCC
GCATCCACACGTCAGTGATGCTTGGTAGCTGTGTCAGTATTCTTGCATTGTTGGACTC
TGGCTCTGAAATAAGCCTGATGAGTTCCGACACTTTCTCAAAGGTGGCTAGAGCAAT
GCACCTCATTGGCAAACCCCTTCAAATCGAGCCCTGCGACGTTAGCATCACCAGCTA
CACGCAGGACAGGTCACCCATCACGAAGCGTGCTTGGGTTGACATAACTTTCAAAG
AAATGTCTCTTGTACACCCCATTTACATCTGCCTTCTTGACACCGAACCTCTGCTCAT
TGGCCAAGATCTCCTAGATCGACTGGCACCGCTCATCGACTGCCATCAGGGTCATAT
CTGGGCTCAGGTGAAAATCTCCTTCCCCTCTGGGGAAGGAGAGTTCGAACCCTCTCC
GCGGTCATGGTCTTTCCTTTGTTCCTTAGAAAACACCAACCAGGAGCTGTATAGTCC
CAGCGTGGTCAGGAAAGTCCAGCTCAACGGTCTTCTCATATCGGACGCGCGTCTGGC
GCTCTGGTCTGAGAGAACCGCGGTCAGCCAGGACCTGTACCACGAACTTCAAAAGC
GAGACCCACCCCCAGTTTTCTCCCAACGTAGCCACCGTGTTCTGTCGGCAACACACC
CACAAACCCTCCTCAAGGCCGTCGGTGTATGTGCCCTCTTGGTACTGATTGGGAAAA
AAAGGGTGTTTCATTTTGCCAGTGTGGTACCCGACCTACATGCCCCGTTCTTCATGG
GGGCGGACCTTCTGGTCAGGCTCGGGGTCCAGTTAGACACGGTCAATCAGGTTTTGT
GGTCACGAGCCAATGCAGAAGCTCATTCCCTAACCACCGAACCACAGCAAATGACT
TCGGGGCAGACCATTCCCCAAGCCTGCCGGGTAATGAGTGAAACTGTGGTGACCAT
CCCAGCACGGACTGCCGGTGTTCCCATCCGTCTTGCAATCTGTCAAGGCCAGAAGAT
ACCTGGCCCTCGGGCTTTCTTCCAGCCGTTGCACGGTTTTTGGGAACTGAACTTGGCC
GTCTGTGGCACGCCCCTGCTTGAATTGGACAATCGTTCAGCCTACCTCCTTGTTCAGA
ATCTTACGCAAGCTCCGATCCAGGTTGCAGCACGGCAACTCCTGGGGATGCTGATCG
ACCACTCTTTTCATGATTTTGAACTGACCATTCCAGTCATAGGGGAGATGCCCCCAT
CTTTGGACACCAATCTTGACGGTGTTCTCTTCACTTACCCAACTAAAATGATCACCGT
TGCCAGTGACGAAACGCTACAGAACGAAACAATCTGCAGCACAAGTTTAGGCAAAA
GAGGGGACATGGTGGTGTATGCCCTGGTGGCACAACCAGGAGGGTCCCCAGATACT
GGTGACAAAAGACAGGAGTTACCAGAGGAACCGTACCCTGGGTTCGAGTCTGAGAT
TGAGCAGCAGCTGGGGAAAGCAGATGCCCTGACAACAGAACATCAGCGGAACGAG
CTCCGGAAACTATTTCATGAGTTCCAGCCAATCTTCTCCAAAGACTCCCAGGATTGC
GGAGTCACGGACATCCACACTGTTCGCATTCCCACCGATCCAGGCGCAGCACCCACC
TTTGTACGCCAGTACAAAATCCCGTTGGCCGCCTATGAATCGATCCAAGAGATCCTC
GACAAGCTGTTGGAAAAACAGATTGTCAGGGAATGCAACTCCACATACAATTCACC
CATTTGGCCGGTGTTAAAACCGACCGGGAAATGGCGCCTCACTATTGACTATCGGCC
ATTAAACAAACAGGTCCCACTCTCCCGTTGGCCCATGATCCACTTGGATCAGGAACT
GGCCAAGGTCAAAGAGGCCCGGTTCTTTTCAACGGTGGACGTGGCAAACGGGTTCT
GGACCATGAGGGTCGACCCAGTGGACCAGTACAAGCTGGCTTTCTCTTTCTGCAATC
GCCAGTACACCTGGAACCGCTGCCCGTTCGGATATTCAAACTCCCCAGCTGAGTTCA

ACATCTTCCTCCACAAGGCAATGAGCGATGCCTCAGCCCGTGGAAACCTCATCTACG
TGGATGATATCCTCATGAGGAGTCGGTCCTTCGATGAACACCTGGCCGAGATCCGTC
ACATTCTGGGCCAGCTTGCGGCCGCAGGAGCAAAACTGGCAATTACCAAGGGCCAG
TGGTGCCGCACCAAGGTAGAGTACGTGGGGCTGACCGTGGGTGCCAATGGTATTGA
GCCTAAAGCAGAAAGGATGCGAGCTATCCAAGACATCATGGCCCCTGTGAACGTTT
CAAAACTCCGGAGCTTTCTGGGAGTCTGTAACTATTCCCGACAGTTCATCGCAGACT
ACGCAGAAATAGCCAGACCCCTCACAGAACTCCTCGGCAAAGACAAACCGTTCCTG
TGGGGTGAACCCCAAGAGCAGGCCTTCCGCCAAATGAAGGCAGAACTCTGCTCCGC
TCCCTGTCTGGCCTACCCGGATAAGGACAAGGAGTTCAACCTTGAGGCGAGTTTCTC
CACCCAGTGCCTTAGTGCCGCCTTGGCCCAGAAATATGATGCAGACAAACGGGTCG
TGGCTTACGCCAGCCGACCCCTCAGCAGCGTGGAGATGAAGTTTTCCGACTGCGAG
AAAGCCCTTCTGGCCACAGTGTGGGCGGTGGAACACTTCCGCAGCTACGTAGGAGG
GCAAAAGGTGCTCATTGAGACCTGCCACCAACCGGTCACCTTCCTGAACAGCCAGC
GGCTTCGGGAAGGGCGGGTGTCCAACAGTCGCATCGCCTCCTGGATGATGGCGTTGC
AAGGCTATGACATCGAGGTAAAGTATGCTCAGAACCACAAGATGGCACTAGGTCAA
GGCCTGGCTGAGTGTCAACACTGCGACTGTGATGAACAGCCAAGCCTCCAGTCCTTA
CTGGTCACCACCCCATCACTCCCGTCCAACCATCATTATTATGATGAAAATGTTTGTC
AAGGCCTCCGTGAAGTCTATGTGGATGGCTCATTCCACCACGAAAGTCAGACTCAAG
CTGGTGTGGGCATTTACTGGGTCAAGTACCGCAGTGAGGAACCCAGTCACTACCAAC
TTGGTCCCAAGACCAGCCAGTATGCGGAAATCGCTGCAGGGTTCATTGCGCTTCAAC
AAGCGGCCAGACAAAAGATCAGTCAGCTGGTGATCTGCTCTGACTCCAACTATGCC
CGCTACAGTTTCATTGCACACTTTCCTGCGTGGAAAGAGAAAGGTATGCTGAATGCC
CGCAACAAAGAGATCAAGCACGCAAAACTTTTCCTCGCATGCGACCGACTGGTCTC
ACAGCAAGGAATGACCGTGTACTGGAGAAAGGTCAAGGGCCATTCCCAAATCTCTG
GTGGGGATAAGGAAGGAAACGATGAAGCTGACCGCCTGGCCAAACTGGGCGCAAG
TCAAGGTACGCCCTGGCAGTTTCAGGAGGAGTGGCTTCCAGATTCCCAAACTCACGC
GGTGAATGCAATCACTCGCCGTCAGGCCAGAGAGAGACGAGAAAATTCTCAAGCCG
GTGCGCAAACCCTAAATCTGTGTCGCCAGCCTGAGGACACAGACCTAATCACCATG
CAGGAGAGGGACCCAGCTATCTGGGCCGTTCGGCAGCTGCTAGCAAACCCCCCACC
GCAAAACACCACTCAACCCTCACCAGGTGAGCCAGAGGAACTATANGCCCTCCGCC
TCAACCCGCAGTCTTTCAGGATTGAGAAGGGTCTGTTGGTCTACTCTGATGGTGATC
AGGGACCGACAAGGTGGGTCGTCCCCACTCACCACAGGGGGTTGATGTTAGCACAC
GCTCACGACTCGCCAGTGGGGGGCCATCGAGGCCCCCAACCCNTGGAAACCCTCCG
GCAAGTTGCCTATTGGCCGTCAACGAACCGCGATGTCAAAGCTTACGTTAGAGGATG
TCTGGTCTGCTGCCAGTTCCAGCCGTCCCAGCCATTGGGTAGAGCCCCGCTCCAGAA
AAGGGGGGTCACGTTCCCCTGGTCCCACCTCCAAGTGGACTGGGTGGGCCCCGTACC
CAGATCGTCAAGAGGAAATAAATACCTCCTCACGGTCACATGCGCTTTTACAAAATG
GGTGGAATGCCTTCCAGCACCCAACGACACGGCAACGACCACAGCAGTCTTGCTAC
TAAACCATGTTTTCAGTCGTTGGGGACTACCCCTTTCCGTTGATTCGGATCGAGGCA
CCCATTTCACGTCCGGGGTCATGACTGCAATGTATGACATTCTTGGAGTAAAAACAA
GATTCCACATCCCATATCACCCCCAATCATCAGGCCAAGTGGAACGCGCTAACCGG
ACCATTGTCAACATGCTCAAGAAGTATGTCAATGGTAGCGCTAGGGATTGGGACGT
GAAACTTCCCCTGGTACTGATGGCCATCAGGTCCACTCCACACCGCACCACGGGGAT
CACTCCATTCGAGATGATGACCGGAAGAGAGATGACGTTGCCACTGCATCTCTTGTA
CCACCCTCAAGATGTCAGTGTCTCAACGGCATACACTGCCACTCAATACGTCTCGAA
ATTGCGCGAACATCTTCGGTCCACATTTGCCTGGGCCCAAGAAAACCTGGAGACAA

GGGCAAAGGGGGCCAAGGCGTACTACGACCGAAAGACAACTCCCCGCGAGTATGA
AATAGGTGACAAAGTCTTCTATTTCAGGTTCACCAAACCGGTGGGAATCTCTCGAAA
GTTCCTACCAAGCTGGTCAGGCCCAGTTGAGATTGTGGGAAAACTTTCCCCGGTGGC
GTACCGGATCAGATTGGCTAAGCCGAAACAGCCACCGACCTACAAATGGGTCCATT
CAAACCAAATCAAGCCATACGAGGCGCCACCCCTCCCCACAGAAGGGGAAAACTCC
CCATCAGCAGAAAGGGGGAACAGCCCAACACAGACATAGGAATAACTGCATGACC
AATCACTTCCGCGAGCATGCACTAACCCCTAAAAGAAATCCCCTGTCCACCTAGTAA
TCAAAACTCTGTTCTGTTGCAGGATGGAGCTACTGTGGGTCCTTGGCAATCTCCTCG
GGCTATCAACCCCTCTCCCAACTGAGGTGGTTGAATCAGGCCCGCCCACCGGGATCG
TCCTCCAGGGGACGCCCGGACTGCTAATCACCCGGTGTGGTCTCTTCACCCAGAGAG
TCTATGTACGACTAGACCCCTGGGATGTCTACCGCAAACACATCGTTCTACCACCCC
AATTGAGCGAGGGACGGTTAAGTGGGGTCCAGACCCACGACATAGTAGAACGAGCT
AAACAGACCACTGCCCACATTCTGGAACAATTGCAAAAGTTCATAGTCACAGAAGA
AGACCTCAGTGGCCAGAGACGACCTAAACGTTTCTTAGGCGGCCTGCTCGTCGCAGC
CTCAGCCATCGGGTCATTATTTTCTATTGGCCTGTCCGCGGCCAACTCAATTAGCCTC
GGCGCGCTCCAGAGACACATGGGGGAACTCGACCAAGAAATGCCAGAGATCCAAC
AGAGACTCCTCCTCCAACAAGAACAGTTACAAAACCTAGGAAAAACTTTCCAGGGG
ACCATAGTTACTGTAAATTTGCACTCTGCCCTCCTGAATAACACCATGCATGCCTTA
GAGAATTTGTCCAAAGTGGTCCAAGAAGACCTAACGTACGTCCGCATAGTCAGAGA
CCTAATGCAGGACCTTGTCAGGGAGGTTAGCTCCTCAGTCAACAGCCTGAGCGCGG
GAAAAATTCCAGCTTACCTAATCTCATTAGGACTCGTCGAACAGATCTTGAAATCCG
CCACTAGTACGGTAGTACACCCTTCACAGGTACACCTGGCGTATAGCCTTGGCAGTG
CAATCCCCATCTCAGTAGACTCTAAGAACTTAGAGATAGGCTTCATCCTCAATCTTC
CGATTCTAGAGAGACAAAACATATACCGACTAAAGTCTGTACTGAACGTGGGCTTCT
GGAAAGGTAACACACACGTACACCTTAAGACCCCACCCACAATTGCTTTCCATGAG
GACAACCCCTCACTCTACCTCATTCCTAACCTTAACATGTGCACTAAAACAAAAGAC
ATACACTGGGTCTGCCCGAGCAACCCATTCGTTAGAGATGTCACTAACTACCTCTGC
GGCCTGAGGGCCGACACACCTGAACAAAAATGTCAGGGCAGTATGTCCATAAGGGA
CGAGGGAACAGAAACGAGGGTGGAACGGGCGGGTAACCGATGGCTCGTCAGTACA
CCCGTCACGGAAGCGTTAATGTCCTACGACCGTCATGACACAGCCACTAAGTTAACT
TTACCTAACCAAACGGTATTTTTAACAGTTCCCCAAGGAGCTACCATTCATATCGAT
GACATTGTCCTCCATCATCTCGATCCCGACAGACACGATGCCGAGATTGAGATCACG
GACGCATTTCGAGGCCACGACCTCACTATTGAATACCCCCTCCAACAGCAACTTCTG
GCTGAGGGGACTAAACTGGTTCAGTTCAGCCTCAAACCTACGGGGCTTACCACCAC
GTTTTCCAACCATGTAGGGAGTTCGGCGTCTTACCAAGACCACCCTGTTAGCATGGT
TGCCTTGGGCCTCCTTCTTAGTGGCTGGATCATCACCGCGATAGTCGCACATTTCATG
TACAGACACATACAGAATTTGCAAACTAAACTTGACTCGATTCTGTTCGTATCACCA
CGTTTCACACAGCAGACCACATCATCCCACTAGTCAGCGCCAGTCCTTTTCAGGGAT
AGGTTAGTGACACCCAAATGATCTCTCACACCCTGTCCTTTCTTCTTTTTTTTTTTTTT
TTCTTTATCTCTCAACTTTCGTTTTGGCATTATACCATAAGGTGCATTATGAGGGGGC
ACGCAGTCTTACCTGCTTCTTGAAGTATTTTCCAAGTGTTATTAGTGTGATGTATTTT
AGTAGTAGTTATGATTTATACAAGGTTTAAATCGTAGTTAGAAAGTCTGCCCAGACG
CTGCCTCAGAGTCTGCCCAGACACTATAGCCTCAGTAACCAGATGAACACTATGAAC
TGTTTGCCCATTCCCAGGAATGAATGATGCTATGTTGCTTTCAGAATAACTCTTGTTT
GTTTACATAACCTGTTGGTCCATGGTTGCTCCAAAGAGCGGGGTTCAACCTGCTCTT
CAGACCAAAAGGGGGACTGTTGGAACCTAACAGTTATGATCAAGACTGTTCTGTGA

GGCCTGAAACTTCCCTGTGGCCAAAGAGCGGCCTGTAGGGAAGTCACCAGCCTGTG
ACCCCTCACTCCTAAACAAAGGAGTCTCCAAAATGGAGATTGTACCTCCAAGACAT
GAGGAAAATGGCAGACTGGTGGAGGGTTTGAGAACTGCAACGACGGATTCTCACAC
CTTTGTGGAGTTCGGGATTTCTGATTGGCCGCGGCTCCTTGAACGCACCACTTGCCG
CCAGCAGGCGGCGGAGGGAAGATAAAAATAGCCGGAGGTCACAGCCCGGGGTTCTT
CGAGGTAATCCATTTGCCGCAAGGAGACGCATCGTTCAGCGCTGTTCGAGAATCAAC
ATTGTATCATTCGCTGGTCGAGTGAGACATTTGTATCGTTCGTGGTACAACAGGGTC
CTGGTGGATACGAAGCCCGTATTACACCAAATCTGCAGAGGGGTGAATCAGCCCCG
TCTCGAGGAAAAGAGACAACGCGTTTTTCTTGCAAGTCGACTGGACCGTGTTTCCGC
TCCACTGCCTTGGACGGTACGGGTCGTTAATCTCTGGCGAGAGTTAACTCAGTTTTG
TTCACCGGAGAATCCGCCGGATAAGCAACGGACTGTACACCAAAAGTAAGAGGCTT
TTATAGTTCTGGGCAGATTTGAGAACTAGGGTGTTTATTAATGAGTAAAAGGTACTG
CTAATTTTTGTTTACCATTAATGTGTGATCGGACCGCTGCGTCCTATGTGCTGTTGTG
AAGAATATCACTGATCCAGATATTGTTGAAAGTTGTGTTGAACTGTAACTGATACTC
CCGCCGAGGAGAAATTACTGTATACGCTGAGTCAGCTGCACGCTCAAAGAGTGTGC
CGGGGAACAAACGCTCTCTACCGTTAAACAAACCTCAGATTCTGCTCTAACGCCACG
TTGAGTTAGAATCTCGAGTAAAACGAAGAGAGCGTTTAGTCCCGTTACCCCAGAGAT
CAAGGTTTTAAGTGTTTGTTTCCATAAATCGTGTTTTCCCTCCTTTTCTTTCTTTCTTTT
CTTTTTACACACACACACACACACACACACGCACACATACCTCCAAGCTAGTT
AGCCACGTGGAGTAAGCTAGAGCTAACGCTTTAGCATAAACGTGTGCGTTGCCTCGC
AACCCCCACCTGAGCTTCTCCACCCCCCTTCGGGCGCACCCAGTACCACTACCGCCC
TCTTGAGGCTAAATAGTTAGGTGTAGCTCACTGGAGTAGCCATTTTAGTAGTTCATG
GTAGACTACACTTCGACCCCGCCCCCCCCTTTCACTCACTCACTCTCACACACACAC
CCACACACACGCCGCACAGACACAGACGTGTCTCTGCTGCGTTGGGAGTTTCTTGCT
GTGTTTTGATTGTGATATTGTAAAAGGGTATAGAAGTTTATTATTGAAGGATTGGTG
ATTGGCTACTGATAATTGTGATGTTTTAATAAATCTGATTATACTTTAAATAGCAGTT
GCTTGTGATTATTTGTGTACTTACTGTAATGACTGCTGGCTAGACAGGTCTGTGCTCG
AAATCACCCCCTCTTCTTTAAATCCTTTCAAGGGATCTTAACAGAAATGAGACTGTT
CCACAGTTGATCACTGGTCCCTGACTTACAGGGTGGTGCCCCGTTGAGTGTTTGTCG
ATTTGATAAAGTTATTAATAACCATATTCATAACTTTATTAATATTGATAAAACATAT
TTTGATAATTTGCTAGTGATCAATCTGTACCCTACAAGAGAGAATCCAACACTGGT
> Loki-Pfl consensus sequence 5’LTR
TGTTGGAACCTAACAGTTATGATCAAGACTGTTCTGTGAGGCCTGAAACTTCCCTGT
GGCCAAAGAGCGGCCTGTAGGGAAGTCACCAGCCTGTGACCCCTCACTCCTAAACA
AAGGAGTCTCCAAAATGGAGATTGTACCTCCAAGACATGAGGAAAATGGCAGACTG
GTGGAGGGTTTGAGAACTGCAACGACGGATTCTCACACCTTCGTGGAGTTCGGGATT
TCTGATTGGCCGCGGCTCCTTGAACGCACCACTCGCCGCCAGCAGGCGGCGGAGGG
AAGATAAAAATAGCCGGAGGTCACAGCCCGGGGTTCTTCGAGGTAATCCATTTGCC
GCAAGGAGACGCATCGTTCAGCGCTGTTCGAGAATCAACATTGTATCATTCGCTGGT
CGAGTTAGACATTTGTATCGTTCGTGGTACAACAGGGTCCTGGTGGATACGAAGCCC
GTATTACACCAAATCTGCAGAGGGGTGAATCAGCCCCGTCTCGAGGAAAAGAGACA
ACGCGTTTTTCTTGCAGGTCGACTGGACCGTGTTTCCGCTCCACTGCCTTGGACGGTA
CGGGTCGTTAATCTCTGGCGAGAGTTAACTCAGTTTTGTTCACCGGGAAATCCGCCG
GATAAGCAACGGACTGTACACCAAAAGTAAGAGGCTTTTATAGTTCTGGGCAGATTT
GAGAACTAGGGTGTTTATTAATGAGTAAAAGGTACTGCTAATTTTTGTTTACCATTA
ATGTGTGATCGGACCCGCAAGGCGTCCTATGTGCTGTTGTGAAGAATATCACTGATC

CAGATATTGTTGAAAGTTGTGTTGAACTGTAACTGATACTCCCGCCGAGGAGAAATT
ACTGTATACGCTGAGTCAGCTGCACGCTCAAAGAGTGTGCCGGGGAACAAACGCTC
TCTACCGTTAAACAAACCTCAGATTCTGCTCTAACGCCACGTTGAGTTAGAATCTCG
AGTAAAACGAAGAGAGCGGTTAGTCCCGTTACCCCAGAGATCAAGGTTTTAAGTGT
TTGTTTCCATAAATTGTGTTTTCCCTCCTTTTGTTTCTTTCGTTTCTTTTTACACACACA
CACACACACACACACACACACACACACACACGCACACATACCTCCAAGCTAGTTAG
CCACGTGGAGTAAGCTAGAGCTAACGCTTTAGCATAAACGTGTGCGTTGCCTCGCAA
CCCCCACCTGAGCTTCTCCACCCCCCTTCGGGCGCACCCAGTACCACTACCGCCCTC
TTGAGGCTAAATAGTTAGGTGTAGCTCACTGGAGTAGCCATTTTAGTAGTTCATGGT
AGACTACACTTCGACCCCCCCCCCCCCTTTCACTCACTCACTCACACACACACACCCA
CACACGCCGCACAGACACAGAAGTGTCTCTGCTGCGTTGGGAGTTTCTTGCTGTGTT
TTGATTGTGATATTGTAAAAGGGTATAGAAGTTTATTATTGAAGGATTGGTGATTGG
CTACTGATAATTGTGATGTTTTAATAAATCTGATTATACTTTAAATAGCAGTTGCTTG
TGATTATTTGTGTACTTACTGTAATGACTGCTGGCTAGACAGGTCTGTGCTCGAAATC
ACCCCCTCTTCTTTAAATCCTTTCAAGGGATCTTAACAGAAATGAGACTGTTCCACA
GTTGATCACTGGTCCCTGACTTACAGGGTGGTGCCCCGTTGAGTGTTTGTTGATTTGA
TAAAGTTATTAATAACCATATTCATAACTTTATTAATATTGATAAAACATATTTTGAT
AATTTGCTAGTGATCAATCTGTACCCTACAAGAGAGAATCCAACA
> Loki-Pfl consensus sequence 3’LTR
TGTTGGAACCTAACAGTTATGATCAAGACTGTTCTGTGAGGCCTGAAACTTCCCTGT
GGCCAAAGAGCGGCCTGTAGGGAAGTCACCAGCCTGTGACCCCTCACTCCTAAACA
AAGGAGTCTCCAAAATGGAGATTGTACCTCCAAGACATGAGGAAAATGGCAGACTG
GTGGAGGGTTTGAGAACTGCAACGACGGATTCTCACACCTTTGTGGAGTTCGGGATT
TCTGATTGGCCGCGGCTCCTTGAACGCACCACTTGCCGCCAGCAGGCGGCGGAGGG
AAGATAAAAATAGCCGGAGGTCACAGCCCGGGGTTCTTCGAGGTAATCCATTTGCC
GCAAGGAGACGCATCGTTCAGCGCTGTTCGAGAATCAACATTGTATCATTCGCTGGT
CGAGTGAGACATTTGTATCGTTCGTGGTACAACAGGGTCCTGGTGGATACGAAGCCC
GTATTACACCAAATCTGCAGAGGGGTGAATCAGCCCCGTCTCGAGGAAAAGAGACA
ACGCGTTTTTCTTGCAAGTCGACTGGACCGTGTTTCCGCTCCACTGCCTTGGACGGTA
CGGGTCGTTAATCTCTGGCGAGAGTTAACTCAGTTTTGTTCACCGGAGAATCCGCCG
GATAAGCAACGGACTGTACACCAAAAGTAAGAGGCTTTTATAGTTCTGGGCAGATTT
GAGAACTAGGGTGTTTATTAATGAGTAAAAGGTACTGCTAATTTTTGTTTACCATTA
ATGTGTGATCGGACCGCTGCGTCCTATGTGCTGTTGTGAAGAATATCACTGATCCAG
ATATTGTTGAAAGTTGTGTTGAACTGTAACTGATACTCCCGCCGAGGAGAAATTACT
GTATACGCTGAGTCAGCTGCACGCTCAAAGAGTGTGCCGGGGAACAAACGCTCTCT
ACCGTTAAACAAACCTCAGATTCTGCTCTAACGCCACGTTGAGTTAGAATCTCGAGT
AAAACGAAGAGAGCGTTTAGTCCCGTTACCCCAGAGATCAAGGTTTTAAGTGTTTGT
TTCCATAAATCGTGTTTTCCCTCCTTTTCTTTCTTTCTTTTCTTTTTACACACACACAC
ACACACACACACACACGCACACATACCTCCAAGCTAGTTAGCCACGTGGAGTAAGC
TAGAGCTAACGCTTTAGCATAAACGTGTGCGTTGCCTCGCAACCCCCACCTGAGCTT
CTCCACCCCCCTTCGGGCGCACCCAGTACCACTACCGCCCTCTTGAGGCTAAATAGT
TAGGTGTAGCTCACTGGAGTAGCCATTTTAGTAGTTCATGGTAGACTACACTTCGAC
CCCGCCCCCCCCTTTCACTCACTCACTCTCACACACACACCCACACACACGCCGCAC
AGACACAGACGTGTCTCTGCTGCGTTGGGAGTTTCTTGCTGTGTTTTGATTGTGATAT
TGTAAAAGGGTATAGAAGTTTATTATTGAAGGATTGGTGATTGGCTACTGATAATTG
TGATGTTTTAATAAATCTGATTATACTTTAAATAGCAGTTGCTTGTGATTATTTGTGT

ACTTACTGTAATGACTGCTGGCTAGACAGGTCTGTGCTCGAAATCACCCCCTCTTCTT
TAAATCCTTTCAAGGGATCTTAACAGAAATGAGACTGTTCCACAGTTGATCACTGGT
CCCTGACTTACAGGGTGGTGCCCCGTTGAGTGTTTGTCGATTTGATAAAGTTATTAAT
AACCATATTCATAACTTTATTAATATTGATAAAACATATTTTGATAATTTGCTAGTGA
TCAATCTGTACCCTACAAGAGAGAATCCAACA
> Loki-Pfl consensus sequence *mscramm* gene
CTGCGGGCTCGCGTCAACCTCTTCACAAGCATTCTACAGTTGCGACAAGCAATTGTA
GCCACTGGTGATTGCCCTCTTGCTCAGCAGGTAAGAATGAGCGACCACAGTGACCAT
GAGCTAACAGAAGAAGCTCAGACCGAGTCAATAGACACAGTCATAACAGACACAGT
CTCTGCACTCATCCGTCTGCCCCCAGCGGACAGAGTTTTCCTGGAGGAATACGAGGT
TGGGACCTGTGAAACCAAACTCAAGGAACTACTAGACTACGTGGATCAGCCCACCG
GGAAGCCGGAGCTCACCAAACCAGACGTGCTTGGTCAAATAGTCCTTCTCCAGCAA
CGGAGATTCCACCTGCAGCTCACTGAGCTTCAGAGACACAGTCAAACTCAACAAGA
GGCTATCCGCCGCTTGCGCCAAGAGATCCGGGACGCACACGATGCGTACTCTGATCT
CCAGGCCGAGTGTGTTCTCCTGAAGCAACTGAAAATTGCTAAACAGCAAGGAGATG
CAGGGGACGGGGACACTCAGACCAACCGGGTAAGTACAGAAGCGGCGGAATTCAC
AACGAAGACTGGCGACTCGGACGACGAGCGGGCCGCAGGGATAATTTCCTATCGGT
CTCAACCCGCCAGCGAAGGTGTCACCCCGGATTGGAGGATCATCGACCAAGAGAAA
CCCTCCTCCCAGTGGAGGGAGGAAAATGGCTCCACCATCAGAAGGGTAAGTTATCA
GCCCCCAAGGCCGAAGAACATTCCGGAGTTTCGGGTCTCAAGGAAAGAGGATCACT
CCCCTGTTCGGTTGAAAAGAGAAGAAGTTTCATGCAGCCGAAGAGATGATTACTCTC
CTGCTTGGCCGAAAAAGGAGGGAGTCTCCCGCAGCAGAAGAGATGACGACTCTCCT
GTTCAGCCGAGAAAGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCCTGT
TCGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCCTGTTC
GGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCCTGCTCGG
TCAAGAAAGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTCTCCTGTTCGGTC
AAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACTACTCTCCTGTTCGGTCAA
GAAGGGAGGAACTCTCCCGCAGCAGGAGAGATGACGACTCCCCTGTCCGGTCGAGG
AGGGAGGAGCTCTCCCGTAACAGGAGAGATGAGGAGGGATGA
> Loki-Pfl consensus sequence Mscramm protein
MRARVNLFTSILQLRQAIVATGDCPLAQQVRMSDHSDHELTEEAQTESIDTVITDTVSAL
IRLPPADRVFLEEYEVGTCETKLKELLDYVDQPTGKPELTKPDVLGQIVLLQQRRFHLQL
TELQRHSQTQQEAIRRLRQEIRDAHDAYSDLQAECVLLKQLKIAKQQGDAGDGDTQTN
RVSTEAAEFTTKTGDSDDERAAGIISYRSQPASEGVTPDWRIIDQEKPSSQWREENGSTIR
RVSYQPPRPKNIPEFRVSRKEDHSPVRLKREEVSCSRRDDYSPAWPKKEGVSRSRRDDD
SPVQPRKEEVSRSRRDDDSPVRSRREEVSRSRRDDDSPVRSRREEVSRSRRDDDSPARSR
KEEVSRSRRDDDSPVRSRREEVSRSRRDDYSPVRSRREELSRSRRDDDSPVRSRREELSR
NRRDEEG*
>Loki-Pfl consensus sequence *gag-pol* gene
ATGATTACTCTCCTGCTTGGCCGAAAAAGGAGGGAGTCTCCCGCAGCAGAAGAGAT
GACGACTCTCCTGTTCAGCCGAGAAAGGAGGAAGTCTCCCGCAGCAGGAGAGATGA
CGACTCTCCTGTTCGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACG
ACTCTCCTGTTCGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGAC
TCTCCTGCTCGGTCAAGAAAGGAGGAAGTCTCCCGCAGCAGGAGAGATGACGACTC
TCCTGTTCGGTCAAGAAGGGAGGAAGTCTCCCGCAGCAGGAGAGATGACTACTCTC
CTGTTCGGTCAAGAAGGGAGGAACTCTCCCGCAGCAGGAGAGATGACGACTCCCCT

GTCCGGTCGAGGAGGGAGGAGCTCTCCCGTAACAGGAGAGATGAGGAGGGATGATT
ACTCCCCTGCTCGGCCAGGAAGGGACCGCCACTCGGTCACAACGGGGAGGAGCCGA
CAAGGAGGTGACTACCGCACCTCAGACGACTCTGATGAAAGTCTGAGGGAGAGGTC
CCGTTCCTCTGCCTGGAAGGATCTGAGTCAGCCTCAACGGAAAACCCACCGTCAACA
CCGGAGAGGGAGAGGCAGCGGTACTTCCAGTGACTCTGACGCCTCACGCCACTCCC
CAGCGCAGGGCCTGCGCATGCGCCAGATTGACTCCTTCGCAAGGGAGCTAGAGCGT
TTCGACCCCAACAGTCGAGAAACTAATGTCGACGACTATCTCCGCGATGTAGAACG
GGTCCTGTCAGACCTCCCCAATGCAACGTCGCGCGAGAAGCGCAAACTACTCTGGA
AGACCACGGCAAGGAGCGCCCATGTCTTTATGGACGAACTCCCGCCCTCGACCTTAA
ACAGTTACACAAAGCTCTGCCAAGCTTTACGAGAAGAGTTCGCCCTGTACACCGACC
CAGCTGCGGCCATCCAGGGAGCATATGCCATTGTGCATAAGAGGAATGAGGCCCCG
AGGGAATATTATAGGCGTCTAAGGACCGTTTACTTCCAGGGCCGCCGCTCCCCAGGT
CTAGAGGAAGACCCCGCTTTCCAATCCCTGTTCGTACACAATCTCCACGATGCTGTA
CGGTACGATGTGCAGATGCATTGTAGAACACAAGAGCTCACCATGAAAGAGATTAG
GATGTTCGCACAAAGGGTGTGGGAGGCTCGAACACGACCGATCAGGAAAGGTGAGG
AGGAGCCGAGGGTACTACACATCCAGACCTCAGAGAACGCAGATCTAGCCCTCGAA
GGGAACGAAGTTCCCCAGGGAACAGGGGACTCCAGACCACCAAACCAGGAACAAC
GATCCTACCACACAGGCAGGCAAAATCGACAAAACAACTACAGAGGGAAGAAGAA
GCCCTATGGTGGAAGAAAACCCGGTAAGTATGAGGGAGGACCTCAAAGCCAGGAA
AAGGATGACGGTCGTAATGCCTCCCCAGCACAGCAGCAGGCCCCTCCGAAGGGACC
TGAGACCACCTCCCAACCCGCGCGACTAAGCCAAGCGCCCGTTTCACTCACCGCCAA
GGTGTTCTTCATAGGGTGGGGAAACACGCCAGACAGTCTCCCAGAGTCTGTCACTCC
TCTTCCTCCTGCGGCTTCCAGGGTTTTCTTGGGCGACCTGGTCCATCGGGGCAACGC
GCGCCGCATCCACACGTCAGTGATGCTTGGTAGCTGTGTCAGTATTCTTGCATTGTTG
GACTCTGGCTCTGAAATAAGCCTGATGAGTTCCGACACTTTCTCAAAGGTGGCTAGA
GCAATGCACCTCATTGGCAAACCCCTTCAAATCGAGCCCTGCGACGTTAGCATCACC
AGCTACACGCAGGACAGGTCACCCATCACGAAGCGTGCTTGGGTTGACATAACTTTC
AAAGAAATGTCTCTTGTACACCCCATTTACATCTGCCTTCTTGACACCGAACCTCTGC
TCATTGGCCAAGATCTCCTAGATCGACTGGCACCGCTCATCGACTGCCATCAGGGTC
ATATCTGGGCTCAGGTGAAAATCTCCTTCCCCTCTGGGGAAGGAGAGTTCGAACCCT
CTCCGCGGTCATGGTCTTTCCTTTGTTCCTTAGAAAACACCAACCAGGAGCTGTATA
GTCCCAGCGTGGTCAGGAAAGTCCAGCTCAACGGTCTTCTCATATCGGACGCGCGTC
TGGCGCTCTGGTCTGAGAGAACCGCGGTCAGCCAGGACCTGTACCACGAACTTCAA
AAGCGAGACCCACCCCCAGTTTTCTCCCAACGTAGCCACCGTGTTCTGTCGGCAACA
CACCCACAAACCCTCCTCAAGGCCGTCGGTGTATGTGCCCTCTTGGTACTGATTGGG
AAAAAAAGGGTGTTTCATTTTGCCAGTGTGGTACCCGACCTACATGCCCCGTTCTTC
ATGGGGGCGGACCTTCTGGTCAGGCTCGGGGTCCAGTTAGACACGGTCAATCAGGTT
TTGTGGTCACGAGCCAATGCAGAAGCTCATTCCCTAACCACCGAACCACAGCAAAT
GACTTCGGGGCAGACCATTCCCCAAGCCTGCCGGGTAATGAGTGAAACTGTGGTGA
CCATCCCAGCACGGACTGCCGGTGTTCCCATCCGTCTTGCAATCTGTCAAGGCCAGA
AGATACCTGGCCCTCGGGCTTTCTTCCAGCCGTTGCACGGTTTTTGGGAACTGAACTT
GGCCGTCTGTGGCACGCCCCTGCTTGAATTGGACAATCGTTCAGCCTACCTCCTTGTT
CAGAATCTTACGCAAGCTCCGATCCAGGTTGCAGCACGGCAACTCCTGGGGATGCT
GATCGACCACTCTTTTCATGATTTTGAACTGACCATTCCAGTCATAGGGGAGATGCC
CCCATCTTTGGACACCAATCTTGACGGTGTTCTCTTCACTTACCCAACTAAAATGATC
ACCGTTGCCAGTGACGAAACGCTACAGAACGAAACAATCTGCAGCACAAGTTTAGG

CAAAAGAGGGGACATGGTGGTGTATGCCCTGGTGGCACAACCAGGAGGGTCCCCAG
ATACTGGTGACAAAAGACAGGAGTTACCAGAGGAACCGTACCCTGGGTTCGAGTCT
GAGATTGAGCAGCAGCTGGGGAAAGCAGATGCCCTGACAACAGAACATCAGCGGA
ACGAGCTCCGGAAACTATTTCATGAGTTCCAGCCAATCTTCTCCAAAGACTCCCAGG
ATTGCGGAGTCACGGACATCCACACTGTTCGCATTCCCACCGATCCAGGCGCAGCAC
CCACCTTTGTACGCCAGTACAAAATCCCGTTGGCCGCCTATGAATCGATCCAAGAGA
TCCTCGACAAGCTGTTGGAAAAACAGATTGTCAGGGAATGCAACTCCACATACAATT
CACCCATTTGGCCGGTGTTAAAACCGACCGGGAAATGGCGCCTCACTATTGACTATC
GGCCATTAAACAAACAGGTCCCACTCTCCCGTTGGCCCATGATCCACTTGGATCAGG
AACTGGCCAAGGTCAAAGAGGCCCGGTTCTTTTCAACGGTGGACGTGGCAAACGGG
TTCTGGACCATGAGGGTCGACCCAGTGGACCAGTACAAGCTGGCTTTCTCTTTCTGC
AATCGCCAGTACACCTGGAACCGCTGCCCGTTCGGATATTCAAACTCCCCAGCTGAG
TTCAACATCTTCCTCCACAAGGCAATGAGCGATGCCTCAGCCCGTGGAAACCTCATC
TACGTGGATGATATCCTCATGAGGAGTCGGTCCTTCGATGAACACCTGGCCGAGATC
CGTCACATTCTGGGCCAGCTTGCGGCCGCAGGAGCAAAACTGGCAATTACCAAGGG
CCAGTGGTGCCGCACCAAGGTAGAGTACGTGGGGCTGACCGTGGGTGCCAATGGTA
TTGAGCCTAAAGCAGAAAGGATGCGAGCTATCCAAGACATCATGGCCCCTGTGAAC
GTTTCAAAACTCCGGAGCTTTCTGGGAGTCTGTAACTATTCCCGACAGTTCATCGCA
GACTACGCAGAAATAGCCAGACCCCTCACAGAACTCCTCGGCAAAGACAAACCGTT
CCTGTGGGGTGAACCCCAAGAGCAGGCCTTCCGCCAAATGAAGGCAGAACTCTGCT
CCGCTCCCTGTCTGGCCTACCCGGATAAGGACAAGGAGTTCAACCTTGAGGCGAGTT
TCTCCACCCAGTGCCTTAGTGCCGCCTTGGCCCAGAAATATGATGCAGACAAACGGG
TCGTGGCTTACGCCAGCCGACCCCTCAGCAGCGTGGAGATGAAGTTTTCCGACTGCG
AGAAAGCCCTTCTGGCCACAGTGTGGGCGGTGGAACACTTCCGCAGCTACGTAGGA
GGGCAAAAGGTGCTCATTGAGACCTGCCACCAACCGGTCACCTTCCTGAACAGCCA
GCGGCTTCGGGAAGGGCGGGTGTCCAACAGTCGCATCGCCTCCTGGATGATGGCGTT
GCAAGGCTATGACATCGAGGTAAAGTATGCTCAGAACCACAAGATGGCACTAGGTC
AAGGCCTGGCTGAGTGTCAACACTGCGACTGTGATGAACAGCCAAGCCTCCAGTCC
TTACTGGTCACCACCCCATCACTCCCGTCCAACCATCATTATTATGATGAAAATGTTT
GTCAAGGCCTCCGTGAAGTCTATGTGGATGGCTCATTCCACCACGAAAGTCAGACTC
AAGCTGGTGTGGGCATTTACTGGGTCAAGTACCGCAGTGAGGAACCCAGTCACTAC
CAACTTGGTCCCAAGACCAGCCAGTATGCGGAAATCGCTGCAGGGTTCATTGCGCTT
CAACAAGCGGCCAGACAAAAGATCAGTCAGCTGGTGATCTGCTCTGACTCCAACTA
TGCCCGCTACAGTTTCATTGCACACTTTCCTGCGTGGAAAGAGAAAGGTATGCTGAA
TGCCCGCAACAAAGAGATCAAGCACGCAAAACTTTTCCTCGCATGCGACCGACTGG
TCTCACAGCAAGGAATGACCGTGTACTGGAGAAAGGTCAAGGGCCATTCCCAAATC
TCTGGTGGGGATAAGGAAGGAAACGATGAAGCTGACCGCCTGGCCAAACTGGGCGC
AAGTCAAGGTACGCCCTGGCAGTTTCAGGAGGAGTGGCTTCCAGATTCCCAAACTCA
CGCGGTGAATGCAATCACTCGCCGTCAGGCCAGAGAGAGACGAGAAAATTCTCAAG
CCGGTGCGCAAACCCTAAATCTGTGTCGCCAGCCTGAGGACACAGACCTAATCACC
ATGCAGGAGAGGGACCCAGCTATCTGGGCCGTTCGGCAGCTGCTAGCAAACCCCCC
ACCGCAAAACACCACTCAACCCTCACCAGGTGAGCCAGAGGAACTATANGCCCTCC
GCCTCAACCCGCAGTCTTTCAGGATTGAGAAGGGTCTGTTGGTCTACTCTGATGGTG
ATCAGGGACCGACAAGGTGGGTCGTCCCCACTCACCACAGGGGGTTGATGTTAGCA
CACGCTCACGACTCGCCAGTGGGGGGCCATCGAGGCCCCCAACCCNTGGAAACCCT
CCGGCAAGTTGCCTATTGGCCGTCAACGAACCGCGATGTCAAAGCTTACGTTAGAGG

ATGTCTGGTCTGCTGCCAGTTCCAGCCGTCCCAGCCATTGGGTAGAGCCCCGCTCCA
GAAAAGGGGGGTCACGTTCCCCTGGTCCCACCTCCAAGTGGACTGGGTGGGCCCCG
TACCCAGATCGTCAAGAGGAAATAAATACCTCCTCACGGTCACATGCGCTTTTACAA
AATGGGTGGAATGCCTTCCAGCACCCAACGACACGGCAACGACCACAGCAGTCTTG
CTACTAAACCATGTTTTCAGTCGTTGGGGACTACCCCTTTCCGTTGATTCGGATCGAG
GCACCCATTTCACGTCCGGGGTCATGACTGCAATGTATGACATTCTTGGAGTAAAAA
CAAGATTCCACATCCCATATCACCCCCAATCATCAGGCCAAGTGGAACGCGCTAACC
GGACCATTGTCAACATGCTCAAGAAGTATGTCAATGGTAGCGCTAGGGATTGGGAC
GTGAAACTTCCCCTGGTACTGATGGCCATCAGGTCCACTCCACACCGCACCACGGGG
ATCACTCCATTCGAGATGATGACCGGAAGAGAGATGACGTTGCCACTGCATCTCTTG
TACCACCCTCAAGATGTCAGTGTCTCAACGGCATACACTGCCACTCAATACGTCTCG
AAATTGCGCGAACATCTTCGGTCCACATTTGCCTGGGCCCAAGAAAACCTGGAGAC
AAGGGCAAAGGGGGCCAAGGCGTACTACGACCGAAAGACAACTCCCCGCGAGTAT
GAAATAGGTGACAAAGTCTTCTATTTCAGGTTCACCAAACCGGTGGGAATCTCTCGA
AAGTTCCTACCAAGCTGGTCAGGCCCAGTTGAGATTGTGGGAAAACTTTCCCCGGTG
GCGTACCGGATCAGATTGGCTAAGCCGAAACAGCCACCGACCTACAAATGGGTCCA
TTCAAACCAAATCAAGCCATACGAGGCGCCACCCCTCCCCACAGAAGGGGAAAACT
CCCCATCAGCAGAAAGGGGGAACAGCCCAACACAGACATAG
> Loki-Pfl consensus sequence Gag-Pol polyprotein
MITLLLGRKRRESPAAEEMTTLLFSRERRKSPAAGEMTTLLFGQEGRKSPAAGEMTTLL
FGQEGRKSPAAGEMTTLLLGQERRKSPAAGEMTTLLFGQEGRKSPAAGEMTTLLFGQE
GRNSPAAGEMTTPLSGRGGRSSPVTGEMRRDDYSPARPGRDRHSVTTGRSRQGGDYRT
SDDSDESLRERSRSSAWKDLSQPQRKTHRQHRRGRGSGTSSDSDASRHSPAQGLRMRQI
DSFARELERFDPNSRETNVDDYLRDVERVLSDLPNATSREKRKLLWKTTARSAHVFMD
ELPPSTLNSYTKLCQALREEFALYTDPAAAIQGAYAIVHKRNEAPREYYRRLRTVYFQG
RRSPGLEEDPAFQSLFVHNLHDAVRYDVQMHCRTQELTMKEIRMFAQRVWEARTRPIR
KGEEEPRVLHIQTSENADLALEGNEVPQGTGDSRPPNQEQRSYHTGRQNRQNNYRGKK
KPYGGRKPGKYEGGPQSQEKDDGRNASPAQQQAPPKGPETTSQPARLSQAPVSLTAKV
FFIGWGNTPDSLPESVTPLPPAASRVFLGDLVHRGNARRIHTSVMLGSCVSILALLDSGSE
ISLMSSDTFSKVARAMHLIGKPLQIEPCDVSITSYTQDRSPITKRAWVDITFKEMSLVHPI
YICLLDTEPLLIGQDLLDRLAPLIDCHQGHIWAQVKISFPSGEGEFEPSPRSWSFLCSLENT
NQELYSPSVVRKVQLNGLLISDARLALWSERTAVSQDLYHELQKRDPPPVFSQRSHRVL
SATHPQTLLKAVGVCALLVLIGKKRVFHFASVVPDLHAPFFMGADLLVRLGVQLDTVN
QVLWSRANAEAHSLTTEPQQMTSGQTIPQACRVMSETVVTIPARTAGVPIRLAICQGQKI
PGPRAFFQPLHGFWELNLAVCGTPLLELDNRSAYLLVQNLTQAPIQVAARQLLGMLIDH
SFHDFELTIPVIGEMPPSLDTNLDGVLFTYPTKMITVASDETLQNETICSTSLGKRGDMVV
YALVAQPGGSPDTGDKRQELPEEPYPGFESEIEQQLGKADALTTEHQRNELRKLFHEFQP
IFSKDSQDCGVTDIHTVRIPTDPGAAPTFVRQYKIPLAAYESIQEILDKLLEKQIVRECNST
YNSPIWPVLKPTGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVKEARFFSTVDVAN
GFWTMRVDPVDQYKLAFSFCNRQYTWNRCPFGYSNSPAEFNIFLHKAMSDASARGNLI
YVDDILMRSRSFDEHLAEIRHILGQLAAAGAKLAITKGQWCRTKVEYVGLTVGANGIEP
KAERMRAIQDIMAPVNVSKLRSFLGVCNYSRQFIADYAEIARPLTELLGKDKPFLWGEP
QEQAFRQMKAELCSAPCLAYPDKDKEFNLEASFSTQCLSAALAQKYDADKRVVAYASR
PLSSVEMKFSDCEKALLATVWAVEHFRSYVGGQKVLIETCHQPVTFLNSQRLREGRVSN
SRIASWMMALQGYDIEVKYAQNHKMALGQGLAECQHCDCDEQPSLQSLLVTTPSLPSN
HHYYDENVCQGLREVYVDGSFHHESQTQAGVGIYWVKYRSEEPSHYQLGPKTSQYAEI

AAGFIALQQAARQKISQLVICSDSNYARYSFIAHFPAWKEKGMLNARNKEIKHAKLFLA
CDRLVSQQGMTVYWRKVKGHSQISGGDKEGNDEADRLAKLGASQGTPWQFQEEWLP
DSQTHAVNAITRRQARERRENSQAGAQTLNLCRQPEDTDLITMQERDPAIWAVRQLLA
NPPPQNTTQPSPGEPEELXALRLNPQSFRIEKGLLVYSDGDQGPTRWVVPTHHRGLMLA
HAHDSPVGGHRGPQPXETLRQVAYWPSTNRDVKAYVRGCLVCCQFQPSQPLGRAPLQ
KRGVTFPWSHLQVDWVGPVPRSSRGNKYLLTVTCAFTKWVECLPAPNDTATTTAVLLL
NHVFSRWGLPLSVDSDRGTHFTSGVMTAMYDILGVKTRFHIPYHPQSSGQVERANRTIV
NMLKKYVNGSARDWDVKLPLVLMAIRSTPHRTTGITPFEMMTGREMTLPLHLLYHPQD
VSVSTAYTATQYVSKLREHLRSTFAWAQENLETRAKGAKAYYDRKTTPREYEIGDKVF
YFRFTKPVGISRKFLPSWSGPVEIVGKLSPVAYRIRLAKPKQPPTYKWVHSNQIKPYEAPP
LPTEGENSPSAERGNSPTQT*
> Loki-Pfl consensus sequence *env* gene
CTGTTCTGTTGCAGGATGGAGCTACTGTGGGTCCTTGGCAATCTCCTCGGGCTATCA
ACCCCTCTCCCAACTGAGGTGGTTGAATCAGGCCCGCCCACCGGGATCGTCCTCCAG
GGGACGCCCGGACTGCTAATCACCCGGTGTGGTCTCTTCACCCAGAGAGTCTATGTA
CGACTAGACCCCTGGGATGTCTACCGCAAACACATCGTTCTACCACCCCAATTGAGC
GAGGGACGGTTAAGTGGGGTCCAGACCCACGACATAGTAGAACGAGCTAAACAGAC
CACTGCCCACATTCTGGAACAATTGCAAAAGTTCATAGTCACAGAAGAAGACCTCA
GTGGCCAGAGACGACCTAAACGTTTCTTAGGCGGCCTGCTCGTCGCAGCCTCAGCCA
TCGGGTCATTATTTTCTATTGGCCTGTCCGCGGCCAACTCAATTAGCCTCGGCGCGCT
CCAGAGACACATGGGGGAACTCGACCAAGAAATGCCAGAGATCCAACAGAGACTC
CTCCTCCAACAAGAACAGTTACAAAACCTAGGAAAAACTTTCCAGGGGACCATAGT
TACTGTAAATTTGCACTCTGCCCTCCTGAATAACACCATGCATGCCTTAGAGAATTT
GTCCAAAGTGGTCCAAGAAGACCTAACGTACGTCCGCATAGTCAGAGACCTAATGC
AGGACCTTGTCAGGGAGGTTAGCTCCTCAGTCAACAGCCTGAGCGCGGAAAAATT
CCAGCTTACCTAATCTCATTAGGACTCGTCGAACAGATCTTGAAATCCGCCACTAGT
ACGGTAGTACACCCTTCACAGGTACACCTGGCGTATAGCCTTGGCAGTGCAATCCCC
ATCTCAGTAGACTCTAAGAACTTAGAGATAGGCTTCATCCTCAATCTTCCGATTCTA
GAGAGACAAAACATATACCGACTAAAGTCTGTACTGAACGTGGGCTTCTGGAAAGG
TAACACACACGTACACCTTAAGACCCCACCCACAATTGCTTTCCATGAGGACAACCC
CTCACTCTACCTCATTCCTAACCTTAACATGTGCACTAAAACAAAAGACATACACTG
GGTCTGCCCGAGCAACCCATTCGTTAGAGATGTCACTAACTACCTCTGCGGCCTGAG
GGCCGACACACCTGAACAAAAATGTCAGGGCAGTATGTCCATAAGGGACGAGGGAA
CAGAAACGAGGGTGGAACGGGCGGGTAACCGATGGCTCGTCAGTACACCCGTCACG
GAAGCGTTAATGTCCTACGACCGTCATGACACAGCCACTAAGTTAACTTTACCTAAC
CAAACGGTATTTTTAACAGTTCCCCAAGGAGCTACCATTCATATCGATGACATTGTC
CTCCATCATCTCGATCCCGACAGACACGATGCCGAGATTGAGATCACGGACGCATTT
CGAGGCCACGACCTCACTATTGAATACCCCCTCCAACAGCAACTTCTGGCTGAGGGG
ACTAAACTGGTTCAGTTCAGCCTCAAACCTACGGGGCTTACCACCACGTTTTCCAAC
CATGTAGGGAGTTCGGCGTCTTACCAAGACCACCCTGTTAGCATGGTTGCCTTGGGC
CTCCTTCTTAGTGGCTGGATCATCACCGCGATAGTCGCACATTTCATGTACAGACAC
ATACAGAATTTGCAAACTAAACTTGACTCGATTCTGTTCGTATCACCACGTTTCACA
CAGCAGACCACATCATCCCACTAG
> Loki-Pfl consensus sequence Env protein
LFCCRMELLWVLGNLLGLSTPLPTEVVESGPPTGIVLQGTPGLLITRCGLFTQRVYVRLD
PWDVYRKHIVLPPQLSEGRLSGVQTHDIVERAKQTTAHILEQLQKFIVTEEDLSGQRRPK

RFLGGLLVAASAIGSLFSIGLSAANSISLGALQRHMGELDQEMPEIQQRLLLQQEQLQNL
GKTFQGTIVTVNLHSALLNNTMHALENLSKVVQEDLTYVRIVRDLMQDLVREVSSSVN
SLSAGKIPAYLISLGLVEQILKSATSTVVHPSQVHLAYSLGSAIPISVDSKNLEIGFILNLPIL
ERQNIYRLKSVLNVGFWKGNTHVHLKTPPTIAFHEDNPSLYLIPNLNMCTKTKDIHWVC
PSNPFVRDVTNYLCGLRADTPEQKCQGSMSIRDEGTETRVERAGNRWLVSTPVTEALM
SYDRHDTATKLTLPNQTVFLTVPQGATIHIDDIVLHHLDPDRHDAEIEITDAFRGHDLTIE
YPLQQQLLAEGTKLVQFSLKPTGLTTTFSNHVGSSASYQDHPVSMVALGLLLSGWIITAI
VAHFMYRHIQNLQTKLDSILFVSPRFTQQTTSSH*
>Loki-Pyo consensus sequence
GAAAATTTTAAATCGATTATAATTGATTGTTAAATAAAGAAATAATTTTTTTTCTTTG
CAAAATTTTACTTTTAAAATTCCCCTTTGCACCTTTGCCAAGCAGGTAATAGGTGCTC
TGTGTAGACTGTAGTTGGTTGCTGACTGCGTCTCAATCGTTAGTAGCCTGGATTGGA
GATCCACGACAAGTTATTTTGATTTTGAGGAGACTAATTGTTACCTACTAGGTCTAT
AGGATTGAGCCGCCGAGCTATCCTACCTAAGAGTACTCGATAAACACAGGGTGAAG
CTCGCCCGTGTTTCTTCTTGTGTAGTGTAACGTACGTGTTTTACCCGCTTGCATCCAA
GCGTGTTTGAGTAATCCCACAGGTGTCTTGTAGCACTCAGAAGATGAGTAGCGCGGA
CAACTTCCCTTCAGGGTCAGAAGACCTGCCCCACTTAGTCGCTGATCTGATAAATTT
ATCTATTGATGATGTGAGTGAGGTAAGTGCATATGATGTCAGTACCTGCGATAAGAA
GATTGAGGAACAGATTGGGTACATCAATAACCCTTCTGATCCCCAACGCTCAATATC
CCAAATTCTGGGTCACCTGACCCTCCTGTATGAACAGAGATCACGGCAGAAAGCCA
GAGAACACTCCGAAGCGATCAGTGCCGTGCGAACAGCATGTCAGACAGAGCTAGAT
GAAAAGTCCTATTTGCAATCGCAAAATTCCAGATTACAACAGGAAATGGATAAGAC
CCGTGAAGAATGGGACCAGACACGGCAGGAGTGTGAAGATCTGCGTGAACAGGTCA
GACATCAGAGTGAAAGAGAGGTGAGAGACACAGGAAACAGTCTGTGGACACCTGG
ACTGGAAACAGGGTCAAGAACTGCCGTCTCACCCGGTGAAGCCAGTGGAATAGAGC
TACTGGAAGAGCGACGGCGAGCACCGCCACGCGAACCTCCGATTGAGGAAGATTGG
CGCAGGCATTTTGATGACACTTATCAGAGGCCCCTTCCAGCCAATACAAGAGTCAGA
GAGGAGCCAGCTGAAACCCCAAGGAGTGCAGTGACTGTTGACTACCAGGCAACACC
AAACGCCTGTCCAGTGGACCCTGGTAATCCATTCAGACCGGAAAGAGCCCCTCTACC
AAGAAACCGACCTGCTCCATACCAGGATGATGGAAGAGCTCCGTCCTATACGCCTG
AAGGAGATTTCCTATCAGAGAGGAACCAGGCTAACATCGAACATCGGCCTACCACT
CAAGCAGATGCACGCTCTGGTCCCCGGTTGCGTCCTGAATGGCACTCCAGCTCCGGT
GCAGGGCGAGCAGATGGTTACCGCCACGATCAGTACTCAGACGACTCTGACAGCTC
AATAGCACCTGTGAAAGAAGGAATCCGCACCCGACAGCTCGAGTCCCTGTCCAAAG
ATATTGATCGCTACGATCCCGATGTTCGAGGGTCCAACGTGGACGATTACCTTCGGG
AGGTCGAACGGTGCCTCATGGATTTAGCCAACCCCTCACCCAGAGAAAAGTTGAAG
CTGATCTGGAAGACAACAGCCCGAAGTGTCCACGCGTTCATGGAAACGCTGTCCCCT
GACACTAGAGATCGGTATTCCACCCTCTGTCAGGCCCTGCGAGAAGAGTACTCACTC
TTCACCGACCAGGCCTCTGCAACTCTCGGTGCCTTTGCCGTCATGCAGAGAAAGAAT
GAATCTCCCAGAGAATATTACAGACGTCTGAGAACGGCGTACTTCCAAGGGCGTAA
TGCCCCTGGGTTGGAAGAAGAAACTGCTTTCAAATCCTTGTTCCTGCACAACCTCCA
TGATAGTGTGCGGTACGATGTCACCATGCACTGCAGAGCTGGTGATCTCACGATGCG
AGAGATGAGAAAGTACTCTCAACTAGCCTGGGAGACGCGAACACGCCCAGGTAAAG
CACCTGAACAGGACGCGAGGGTCCTGGGAATCCAAGCTCAACCCAATGTGGATCTG
GCGCTGGAGGGTAACGAGATACCCCTGGCCAAAACACATATCAGAAGGGAGAACC
GGGAACGCCGACCAGTCCAGCAGGATGCTAGGAGAAACCGAGAACGGGAGGGATC

AAACCAGCCCCGCAACCAGAGGTCTCAATATCCAACTAACTCTCAACAGAGAGGGA
GGGACTGGGCTGAGCGAGATCCACACCAACGCCACGAGAAGAAGTTCAGTGATAAA
TATCCCAAACAAGGGAGACCACAAGAACAAGAAAAGAGACAGGAACATTGGAAGC
AACAGGAACATGGTAGAAGACAGGAACCTTGGAATCAACGGGAACGTCGGAATCC
AGAGGAACATGGGAGAAGACAAGAACCATGGAATCAACAGGAGCCACAAATAAAT
CCACGAATGGAAGATTTTATACGCCGATGTGTGGCGGAGGTCGTAAGGGACCGAGA
TGTCCCTACCCGCACCCCTGCACCACCTGGCGATCAGAAAAATCCAGACAAGGGTC
CCCCATCAACGTGACTAGGACAGGTGTCAACAATCCCCACCTCCAGAGTGTACTTTG
TGGGTTGGGGAAACTCAAGCCTCATCCGAATAAACCCAGACGCTTTCACCACAGAC
ACTGCTCCCTTCCTAACTTTCTTGGGCGACCTCGTCCGGAAAGGCAACGCACAACGC
ATCTACACTTCAGTAGTAATAGGAGGCTGCTTCAGCTCCCTGGGCCTACTGGATTCA
GGCTCGGAAATCAGCCTGATGTGCGCCACCACTTTCTCCCAAGTTGAAACAACCATG
TTATCTCTCGGAAAACCCCTGCAGGTTGAAACCTGCAACATTAATATCACCAGCTAC
ACTCAGGACCAGTCATCCATCACAAGACGTGCATGGATTGAGATCACCTTCCAAGA
GATGAGTCTGGCACATCCCTTCTATATCTGTTCCATCGATACAGAACCACTACTCATC
GGTCAGGACCTGCTGGACCGACTCGCTCCACTCATCGACTGTCGCCAAGGACGCATT
TGGGCTCAAGTTGAAAACCCCAAACCTCTCGAGCCTAATGTCAATCTTCCTCCCCAA
GTTGATGCCATCCAGACGCCCAGTAGAATCTCACCGATGCATATGCCTCTTCCAGAG
ACCGACCCCAGTTCGGACCCCCTCTGCACCACGGCAAGTGGGCCTGTGCCACGACCT
GACTCAGAGACCACCTTCAAATCCCATGCTTCCTTCCTATGTTCACTGAACAATGAG
AACCCGTCATTGTACGTTCCCCGCATTACAGAAGATGTTTTCCTAAATGGCGTTCGA
ATCCCAAACACCCTAGTAGCCCTGTGGTCTGAAAAATCCGCAATCAGTACGGACAT
ATACACTGATCTATGCAAACTAGAACTTCCACCTATCTTTGTGCAACGCCGACACCG
TATACTCTCAGCAGAAATACCACAAACTCCTATCAAAGCGACTGGAGTGTGCGCTAT
CTCCGTGCGGATTGGAAAACGACAGATCTGGCACTTTGTGAGTGTGATTCCCCAACT
TCACTCATCATTCATGGTGGGGGCAGATCTCCTCATCAGACTCGGAACACAGGTGGA
TACCATAAACCAAGTGTTGTGGTCCCAGGTCAGGGCAGAAAGACACTCCCTGTCTGC
TGACCCAGACCAGATGCTCTCCGGACAAACGATTCCCCAGGCATGCCAGGTCGCCA
GTGAGGTTGACATGGTCATTCCTCCACGAACTGCTGGGGCCCCATCCGACTTGAAAT
CCTCAAAGGACAAAGGATCCCGGGCCCTCAGGCTTTCTTCCAGCCTTCGCCCTCCTT
CCGGGAACTCAACCTGGAAGTCTGTGGCACACCTCTGCTGGAGTTGCAGCAGATCC
ACTTATGTACTTGTGAAAAATCCCACACACTGCCCCGTTCCATTGAAGGCACACCAA
GCGATAGGAATGTTGATCGACAGCTCATTCCACGATTTCGAACTGACTATACCTGTC
ATTGGGGAGCTTCCCCCTTCTCTGATCAAGGATGGTGATGTGGATCAAGTCTCTCAT
ACATTTCCAACCAGCATGATCACTGTTATGCGTCATGAGTCGCTGCAAGACGAAGCC
ATTTGTGGAGCCTCTCTGACGACAGAAGGCGACATGATGGTGTATGCTCTGGCAGCT
CAACCAGGACAGGACGTCTCAGAGGATGCTCAACCAAACGAACTGTATCCAGGGTT
TGAGGCTGAGATCAAGCAACAACTGGAGAATGCAGATGCCTTGACATCAGACACAC
AGAGAGAAGTGCTTCGGGATCTGTTCCTTGATTTCCAGTCAATCTTCTCAAAAGATT
CTACTGACTGCGGGGTCACAAACCTCCACACTGTGCGCATCCCCACTGATCCAAATG
CCCTACCAACCTTCGTACGTCAGTACAAAATCCCATTGGCAGCCTACGAGTCCATCC
AAGACATCCTTGACACCCTGCTAGAGAAACGAATAATCAGGGAATGCAACTCAACA
TACAATTCCCCCATTTGGCCGGTGTTAAAACCAACTGGAAAATGGCGTCTCACAATT
GATTATCGTCCCCTGAACAAACAAGTTCCCCTGTCCCGCTGGCCGATGATTCACCTG
GATCAAGAACTGGCCAAGGTCAAGGGAGCCACGTTCTTCTCCACAGTAGACGTTGC
AAACGGATTTTGGACTATGAAGGTGGATCAAGCCGACCAGTACAAACTGGCTTTCTC

TTTCGGAGGCAGCCAATACACCTGGAACCGCTGTCCGTTTGGATATTCAAACTCCCC
ATCGGAGTTCAACATCTTTCTCCACAAAGCCATGAGTGACGCCGCAACCCGGGGTA
ACCTCATTTACGTGGATGACATCCTCATGAGGAGCCAAACCTTCGAGGACCACATAA
CAGAGATCCGACATGTACTCACTCAACTGTCCAACGCAGGAGCCAAGCTGGCTCTTA
CCAAAGGCCAGTGGTGTCGCACCAAGGTAAAATATGTGGGTCTGACAGTCGGTGCT
GATGGCATTGAACCCCAAGCAGGGCGGATTCGAGCCATTCAGGACATCAAAGCTCC
GACACGCCTAACAGAACTTAGGAGTTTCCTCGGGGTCTGTAACTACTCCCGCCAGTT
CATCGAAGACTATGCAGAGATAGCCAGACCTCTCAACGAACTCCTTCACAAAGACA
AGCCCTTCGAGTGGGAAGAACCCCAAGAACAATCCTTCCGTCAGATGAAAGAAAAA
CTCTGCTCGGCTTCCTGTCTTGCCTACCCAGACAAAGACAAAGAGTTCCACCTCGAA
GCAAGCTTCTCATCCCACTGTCTCAGTGCAGCCCTGGCACAGAAGTACGACACCGAC
AAACGGGTCGTTGCCTACGCCAGTCGGCCCCTTAGTAGTGTGGAAATCAAATTCTCC
GACTGTGAAAAAGCCCTCCTGTCCACGGTCTGGGCTGTAGAACACTTCCGCAGCTAC
ATCGGAGGACAGAAGGTTATTATTGAGACGTGTCACCAACCTGTTGCCTTCCTAAAC
AGCCAAAGGTTGAGAGAAGGACGAGTCACAAACAGCCGCATAGCCACTTGGATGAT
GGCACTACAAGGCTATGACATCGAGGTCAAGTACGCTCAAAATAACAAGATGGCCC
TGGGCAAAGGCCTGGCGGAGTGCCACCACTGTGATGAAGATGATCAGTCAGATCCT
ATATCTCTCCCAGCAACCACTCCATCGCTGCCTTCCAACCATCGCTATTATGAGGAG
AACGTCTGCCAAGATCTTCCCAAGGTCTATGTGGATGGATGCTCCTTCCACCACCAA
AGTCACCTCCGGGCTGGAGTTGGCATTGCCTGGGTCGACTACAACTGCAGCGAACC
AACTCACTACCAACTAGGACCCAAGACAAGTCAGTACGCAGAGATTGCCGCCATAC
TCATCGCTTTGCAACAAGCCGTTAAGTTGGCTATCACACAATTTGTAATCTGCTCTGA
TTCTAACTACGCCCGTCACAGTTTCATCTCCCATTTTCCTTTATGGAAGGAAAAAGG
GATGAAAAATGCCAGGAACAAAGAAGTCAAGCACTCAGAACTTTTCCTGACCTGTG
ATCAACTGGTAACAGAAAAAGGGTTGATAGTTTACTGGAAGAAGGTCAAAGGTCAC
TCACAAATTGCTGGGCCAGACAAAGACGGCAACGACGAAGCCGACCGACTAGCCAA
GCTAGGTGCAGAACAAGGAATGCACTGGGAGCTACAGGTCGAGGGATTTCCAAGTC
CCCAAACCTGCGTTGTGAACGCTATCACCCGACGACAGGCAAGGGAGAGACTAGAG
AACCCCCAAGCCTCTGATCAGATCATTCATCTGGGACGTAAACCTGGAGATGCAGAT
TTGGTTGCCATGCAGGAGCAGGATCCCGTCATCAGTGCCATACGCCAGATAGTCGCG
AACCCTCCCCCGCAGGACTCACAACCATTTACTGACAAACCCAAGGAGATACGGGC
ATTTCAACTGGCATTACCCCATCTGAAGATGGAAAAAGACTTGCTCGTCTACAACCG
TAATGGCCAAGGACCAAACCGGTGGGTCGTCCCAACCGACCATAGGGGGGTAATGT
TAACCCACGCTCACGATTCCCCTGTCGGTGGCCACAGAGGTACAAGGCTACTCTANA
GTACACTTCAGCAAGTCGGGTACTGGCCATCAATGTCCAACGACACTATGTCATACG
TCCGAGGTTGCCTCACCTGCTGCCAGTTCCAACCGTCACAACCATTGAATCGCGCCC
CACTTCAAGGAAGGGGGGTTTCATTCCCATGGTCCCATCTCCAGATCGACTGGATTG
GACCAGTCCCAAAATCCTCACGAGGTAACAAGTACCTACTCACGGTTACCTGTGCCT
TCACAAAATGGGTGGAATGCCTCCCAGCATCCAATGGCACAGCCGTCACCATAGCA
GTCCTACTGTTGAATCATGTGTTCAGTCGCTGGGGCCTATCAATCTCCATTGATTCAG
ACCGAGGAACACATTTCACGTCCCATGTCATGACATCGCTATGCAAAATCCTGGGAG
TGGAAGTCAGATTCCACATCTCTCATCACCCCCAATCGTCAGGCCAAGTTGAACGTG
CCAACCGAACTATCATTCAAATGCTGAAGAAATACGTGGGAAGTAACGGCAAGGAC
TGGGACATTAAACTTCCCCTGGTGCTAATGGCAATCAGGCTCACCCCGCACCGCTCC
ACAGGTGTCACCCCCTTTGAAATGATGACAGGTAGACAAATGACTCTCCCCCTGCAC
CTGTTATACTATGCTGAAGACCTCAGCGTCATGACCGCCTACACTGCACATCAATAT

GTGGCAGAGCTGCGTGACCACTTGAAGACTACCTTTGCATGGGCACAGGAAACCCT
GGAGGCCAGTGTGGAAGCTTCAAAGGCCTATTATGACAGGAAAACCTCCCACCATG
AATACCATGTGGGAGACAAGGTTTTCCACTTCAAGTTCATAAAACCAGTGGGGGTTT
CCAAGAAATTCCTCCCCAGCTGGATAGGTCCCTTTGAGGTTGTGGGAAAACTCTCCC
CGGTAGCATACCGAATCAGGATCCCGAAACAGAGAAGACCCCCCACATACAAATGG
GTCCACTCAAACCAAATCAAACCATACAAACCTTCCTCAATCTCAGGAGCAGAGGC
CACCCCTACAGCAGAAAGGGGGGAGCCATCCCAACTCCAAGAGGAAGAACACTGA
ACCACTCACAGCATCAATAGTTGCATGCGTCAATTCAAATCCGTCGGCATGCACTAG
TACCACCTCCATTCCACTCCATAATAACTAGGGTTTGTTTGTTTCAGGATGATGCTCC
CGTGGATCCTCGGCATACTCTTCACTCCACAAATGGTTCAACCGCTCCAGGTGGTTC
AGCAAGGCCCACCAACGGGCATTGTTTTTCAGAACACACCTGGACTCCTCATCACCC
GTTGCGGATTGTATACTCAAAGGGTCTACGTACGACTGGACCCCTGGGATGTGTATA
GAAAACACATCCGTAACCCCCCACAGTCGACTGAGGGAAGGTTAAGAGCAATCCAA
ACGCATGACACAGTGGGACATGCCAAGCAGACCACTAAATATATGCTCGAGCAAAT
GCAAAAGTTGTTGGTCACTGAAAAAGACCTCAGTAACCAGAAGAGACCAAAAAGGT
TCCTTGGGTCCCTCATCCTAGCAGCATCCGCTATCGGATCGCTTTTCTCTATTGGTGT
ATCAGTAGCCAATTCAATCAGCGTTGGTGCACTCCAGAGACAGATGAGTGACCTTGA
GGATGAAATGCCGGAAATCCTAGAGAGACTCGACCTCCAACAAGGTCAATTACAGG
ACCTGGGGAAAACACTCCGGGGTACTGTGTTGGCAGTAAACCTTCACTCCGCTCTCT
TAAATAATACCATGCATGCCTTAGACACTCTGTCAGTTATCGTGAGTGACGATATAA
CGTACGTCCAGGTAGTAAAAGACTTAATGCAGGACCTTTTAAGGGAAATTAGTTCCA
CAGTCAACAGCTTAAGCGCCGGTAGGATTCCCACCTATTTGGTGTCTTTAGACCTTG
TTGAAACAATCTTGCAGTCAGCTACCACTACGATTGTGCAGTCCTCTCAGGTACACC
TGGCATTTAGTCTCGGCAGTGCAATCCCAATCTCAGTAGACCCACAAAACCTAGAGA
TAGGTTTTATCATAAACCTACCTATCATCGAACAACAGAACGTGTATCACCTAAAGT
CCGTTCTAAACGTAGGATTTTGGCGGGGAAACACACATGTTAACATTAAAACACCAC
CATTACTCGCCTACCATGACGAAAATCCCTCGCTCTATCTCATCCCAAATTTAAGTCT
ATGTACCAAAACAAAGGACGTCCATTGGGTTTGTCCTAGCACCCCTTTTATCAGGGA
TGTTACATATTACCTCTGTGGCCTGCGAGCCGATGCCCCAGAACAAAAGTGCCAAGG
CAGCATGTCCACCAGGAATGAGGAGTCACCGACTAGAGTAGAGCGAGCTGGAGACA
AATGGTTAGTTAACACACCCGCCACTGAAGCCATAATGGTCTATGACCGTCATGATA
CCTCCACTAAACTAAGCCTGATAAACCAGACAATGTTCTTGGTGGTTCCTCAAGGAG
CCACTATACATATTGGTGACATTGTCCTCCATCACCTCAACTCTGACAGACACGATG
CGGAGATTGAGATCATGGACGCCTTTATGGGTCACAACCTCACGATCGATACAACCA
TCCAACAACAGCTCAGAGCTGAAGGAACAAAGGAGGTAAAGTTCAGTCTTAGATCA
TCTGGACTTATCACCAAATTTGGTAACCACTATCGTCAAACACCAAAGTTTGAAGAA
CGTACCATTAGCATGGTCGCGCTGGGACTGCTTTTTAGTGGATGGATCATCACGATA
GTTGTGGCTTATCTCTTGTACAAATACATTAAGAAACTCCAGGCAAAATTGGATTCA
CTTCTTTTCATCCCGAAAAAGAAGGCTGTGAAAGTTCCTCGTGCTTCCGAAGACATT
GCATTGACTTAAAAGAAAATCCCTTTCGGT
> Loki-Pyo consensus sequence *gag* gene
ATGAGTAGCGCGGACAACTTCCCTTCAGGGTCAGAAGACCTGCCCCACTTAGTCGCT
GATCTGATAAATTTATCTATTGATGATGTGAGTGAGGTAAGTGCATATGATGTCAGT
ACCTGCGATAAGAAGATTGAGGAACAGATTGGGTACATCAATAACCCTTCTGATCCC
CAACGCTCAATATCCCAAATTCTGGGTCACCTGACCCTCCTGTATGAACAGAGATCA
CGGCAGAAAGCCAGAGAACACTCCGAAGCGATCAGTGCCGTGCGAACAGCATGTCA

GACAGAGCTAGATGAAAAGTCCTATTTGCAATCGCAAAATTCCAGATTACAACAGG
AAATGGATAAGACCCGTGAAGAATGGGACCAGACACGGCAGGAGTGTGAAGATCT
GCGTGAACAGGTCAGACATCAGAGTGAAAGAGAGGTGAGAGACACAGGAAACAGT
CTGTGGACACCTGGACTGGAAACAGGGTCAAGAACTGCCGTCTCACCCGGTGAAGC
CAGTGGAATAGAGCTACTGGAAGAGCGACGGCGAGCACCGCCACGCGAACCTCCGA
TTGAGGAAGATTGGCGCAGGCATTTTGATGACACTTATCAGAGGCCCCTTCCAGCCA
ATACAAGAGTCAGAGAGGAGCCAGCTGAAACCCCAAGGAGTGCAGTGACTGTTGAC
TACCAGGCAACACCAAACGCCTGTCCAGTGGACCCTGGTAATCCATTCAGACCGGA
AAGAGCCCCTCTACCAAGAAACCGACCTGCTCCATACCAGGATGATGGAAGAGCTC
CGTCCTATACGCCTGAAGGAGATTTCCTATCAGAGAGGAACCAGGCTAACATCGAA
CATCGGCCTACCACTCAAGCAGATGCACGCTCTGGTCCCCGGTTGCGTCCTGAATGG
CACTCCAGCTCCGGTGCAGGGCGAGCAGATGGTTACCGCCACGATCAGTACTCAGA
CGACTCTGACAGCTCAATAGCACCTGTGAAAGAAGGAATCCGCACCCGACAGCTCG
AGTCCCTGTCCAAAGATATTGATCGCTACGATCCCGATGTTCGAGGGTCCAACGTGG
ACGATTACCTTCGGGAGGTCGAACGGTGCCTCATGGATTTAGCCAACCCCTCACCCA
GAGAAAAGTTGAAGCTGATCTGGAAGACAACAGCCCGAAGTGTCCACGCGTTCATG
GAAACGCTGTCCCCTGACACTAGAGATCGGTATTCCACCCTCTGTCAGGCCCTGCGA
GAAGAGTACTCACTCTTCACCGACCAGGCCTCTGCAACTCTCGGTGCCTTTGCCGTC
ATGCAGAGAAAGAATGAATCTCCCAGAGAATATTACAGACGTCTGAGAACGGCGTA
CTTCCAAGGGCGTAATGCCCCTGGGTTGGAAGAAGAAACTGCTTTCAAATCCTTGTT
CCTGCACAACCTCCATGATAGTGTGCGGTACGATGTCACCATGCACTGCAGAGCTGG
TGATCTCACGATGCGAGAGATGAGAAAGTACTCTCAACTAGCCTGGGAGACGCGAA
CACGCCCAGGTAAAGCACCTGAACAGGACGCGAGGGTCCTGGGAATCCAAGCTCAA
CCCAATGTGGATCTGGCGCTGGAGGGTAACGAGATACCCCTGGCCAAAACACATAT
CAGAAGGGAGAACCGGGAACGCCGACCAGTCCAGCAGGATGCTAGGAGAAACCGA
GAACGGGAGGGATCAAACCAGCCCCGCAACCAGAGGTCTCAATATCCAACTAACTC
TCAACAGAGAGGGAGGGACTGGGCTGAGCGAGATCCACACCAACGCCACGAGAAG
AAGTTCAGTGATAAATATCCCAAACAAGGGAGACCACAAGAACAAGAAAAGAGAC
AGGAACATTGGAAGCAACAGGAACATGGTAGAAGACAGGAACCTTGGAATCAACG
GGAACGTCGGAATCCAGAGGAACATGGGAGAAGACAAGAACCATGGAATCAACAG
GAGCCACAAATAAATCCACGAATGGAAGATTTTATACGCCGATGTGTGGCGGAGGT
CGTAAGGGACCGAGATGTCCCTACCCGCACCCCTGCACCACCTGGCGATCAGAAAA
ATCCAGACAAGGGTCCCCCATCAACGTGA
> Loki-Pyo consensus sequence Gag protein
MSSADNFPSGSEDLPHLVADLINLSIDDVSEVSAYDVSTCDKKIEEQIGYINNPSDPQRSIS
QILGHLTLLYEQRSRQKAREHSEAISAVRTACQTELDEKSYLQSQNSRLQQEMDKTREE
WDQTRQECEDLREQVRHQSEREVRDTGNSLWTPGLETGSRTAVSPGEASGIELLEERRR
APPREPPIEEDWRRHFDDTYQRPLPANTRVREEPAETPRSAVTVDYQATPNACPVDPGN
PFRPERAPLPRNRPAPYQDDGRAPSYTPEGDFLSERNQANIEHRPTTQADARSGPRLRPE
WHSSSGAGRADGYRHDQYSDDSDSSIAPVKEGIRTRQLESLSKDIDRYDPDVRGSNVDD
YLREVERCLMDLANPSPREKLKLIWKTTARSVHAFMETLSPDTRDRYSTLCQALREEYS
LFTDQASATLGAFAVMQRKNESPREYYRRLRTAYFQGRNAPGLEEETAFKSLFLHNLH
DSVRYDVTMHCRAGDLTMREMRKYSQLAWETRTRPGKAPEQDARVLGIQAQPNVDL
ALEGNEIPLAKTHIRRENRERRPVQQDARRNREREGSNQPRNQRSQYPTNSQQRGRDW
AERDPHQRHEKKFSDKYPKQGRPQEQEKRQEHWKQQEHGRRQEPWNQRERRNPEEHG
RRQEPWNQQEPQINPRMEDFIRRCVAEVVRDRDVPTRTPAPPGDQKNPDKGPPST*

> Loki-Pyo consensus sequence *pol* gene
TTGGGCGACCTCGTCCGGAAAGGCAACGCACAACGCATCTACACTTCAGTAGTAAT
AGGAGGCTGCTTCAGCTCCCTGGGCCTACTGGATTCAGGCTCGGAAATCAGCCTGAT
GTGCGCCACCACTTTCTCCCAAGTTGAAACAACCATGTTATCTCTCGGAAAACCCCT
GCAGGTTGAAACCTGCAACATTAATATCACCAGCTACACTCAGGACCAGTCATCCAT
CACAAGACGTGCATGGATTGAGATCACCTTCCAAGAGATGAGTCTGGCACATCCCTT
CTATATCTGTTCCATCGATACAGAACCACTACTCATCGGTCAGGACCTGCTGGACCG
ACTCGCTCCACTCATCGACTGTCGCCAAGGACGCATTTGGGCTCAAGTTGAAAACCC
CAAACCTCTCGAGCCTAATGTCAATCTTCCTCCCCAAGTTGATGCCATCCAGACGCC
CAGTAGAATCTCACCGATGCATATGCCTCTTCCAGAGACCGACCCCAGTTCGGACCC
CCTCTGCACCACGGCAAGTGGGCCTGTGCCACGACCTGACTCAGAGACCACCTTCAA
ATCCCATGCTTCCTTCCTATGTTCACTGAACAATGAGAACCCGTCATTGTACGTTCCC
CGCATTACAGAAGATGTTTTCCTAAATGGCGTTCGAATCCCAAACACCCTAGTAGCC
CTGTGGTCTGAAAAATCCGCAATCAGTACGGACATATACACTGATCTATGCAAACTA
GAACTTCCACCTATCTTTGTGCAACGCCGACACCGTATACTCTCAGCAGAAATACCA
CAAACTCCTATCAAAGCGACTGGAGTGTGCGCTATCTCCGTGCGGATTGGAAAACG
ACAGATCTGGCACTTTGTGAGTGTGATTCCCCAACTTCACTCATCATTCATGGTGGG
GGCAGATCTCCTCATCAGACTCGGAACACAGGTGGATACCATAAACCAAGTGTTGT
GGTCCCAGGTCAGGGCAGAAAGACACTCCCTGTCTGCTGACCCAGACCAGATGCTC
TCCGGACAAACGATTCCCCAGGCATGCCAGGTCGCCAGTGAGGTTGACATGGTCATT
CCTCCACGAACTGCTGGGGCCCCATCCGACTTGAAATCCTCAAAGGACAAAGGATC
CCGGGCCCTCAGGCTTTCTTCCAGCCTTCGCCCTCCTTCCGGGAACTCAACCTGGAA
GTCTGTGGCACACCTCTGCTGGAGTTGCAGCAGATCCACTTATGTACTTGTGAAAAA
TCCCACACACTGCCCCGTTCCATTGAAGGCACACCAAGCGATAGGAATGTTGATCGA
CAGCTCATTCCACGATTTCGAACTGACTATACCTGTCATTGGGGAGCTTCCCCCTTCT
CTGATCAAGGATGGTGATGTGGATCAAGTCTCTCATACATTTCCAACCAGCATGATC
ACTGTTATGCGTCATGAGTCGCTGCAAGACGAAGCCATTTGTGGAGCCTCTCTGACG
ACAGAAGGCGACATGATGGTGTATGCTCTGGCAGCTCAACCAGGACAGGACGTCTC
AGAGGATGCTCAACCAAACGAACTGTATCCAGGGTTTGAGGCTGAGATCAAGCAAC
AACTGGAGAATGCAGATGCCTTGACATCAGACACACAGAGAGAAGTGCTTCGGGAT
CTGTTCCTTGATTTCCAGTCAATCTTCTCAAAAGATTCTACTGACTGCGGGGTCACAA
ACCTCCACACTGTGCGCATCCCCACTGATCCAAATGCCCTACCAACCTTCGTACGTC
AGTACAAAATCCCATTGGCAGCCTACGAGTCCATCCAAGACATCCTTGACACCCTGC
TAGAGAAACGAATAATCAGGGAATGCAACTCAACATACAATTCCCCCATTTGGCCG
GTGTTAAAACCAACTGGAAAATGGCGTCTCACAATTGATTATCGTCCCCTGAACAAA
CAAGTTCCCCTGTCCCGCTGGCCGATGATTCACCTGGATCAAGAACTGGCCAAGGTC
AAGGGAGCCACGTTCTTCTCCACAGTAGACGTTGCAAACGGATTTTGGACTATGAAG
GTGGATCAAGCCGACCAGTACAAACTGGCTTTCTCTTTCGGAGGCAGCCAATACACC
TGGAACCGCTGTCCGTTTGGATATTCAAACTCCCCATCGGAGTTCAACATCTTTCTCC
ACAAAGCCATGAGTGACGCCGCAACCCGGGGTAACCTCATTTACGTGGATGACATC
CTCATGAGGAGCCAAACCTTCGAGGACCACATAACAGAGATCCGACATGTACTCAC
TCAACTGTCCAACGCAGGAGCCAAGCTGGCTCTTACCAAAGGCCAGTGGTGTCGCA
CCAAGGTAAAATATGTGGGTCTGACAGTCGGTGCTGATGGCATTGAACCCCAAGCA
GGGCGGATTCGAGCCATTCAGGACATCAAAGCTCCGACACGCCTAACAGAACTTAG
GAGTTTCCTCGGGGTCTGTAACTACTCCCGCCAGTTCATCGAAGACTATGCAGAGAT
AGCCAGACCTCTCAACGAACTCCTTCACAAAGACAAGCCCTTCGAGTGGGAAGAAC

CCCAAGAACAATCCTTCCGTCAGATGAAAGAAAAACTCTGCTCGGCTTCCTGTCTTG
CCTACCCAGACAAAGACAAAGAGTTCCACCTCGAAGCAAGCTTCTCATCCCACTGTC
TCAGTGCAGCCCTGGCACAGAAGTACGACACCGACAAACGGGTCGTTGCCTACGCC
AGTCGGCCCCTTAGTAGTGTGGAAATCAAATTCTCCGACTGTGAAAAAGCCCTCCTG
TCCACGGTCTGGGCTGTAGAACACTTCCGCAGCTACATCGGAGGACAGAAGGTTATT
ATTGAGACGTGTCACCAACCTGTTGCCTTCCTAAACAGCCAAAGGTTGAGAGAAGG
ACGAGTCACAAACAGCCGCATAGCCACTTGGATGATGGCACTACAAGGCTATGACA
TCGAGGTCAAGTACGCTCAAAATAACAAGATGGCCCTGGGCAAAGGCCTGGCGGAG
TGCCACCACTGTGATGAAGATGATCAGTCAGATCCTATATCTCTCCCAGCAACCACT
CCATCGCTGCCTTCCAACCATCGCTATTATGAGGAGAACGTCTGCCAAGATCTTCCC
AAGGTCTATGTGGATGGATGCTCCTTCCACCACCAAAGTCACCTCCGGGCTGGAGTT
GGCATTGCCTGGGTCGACTACAACTGCAGCGAACCAACTCACTACCAACTAGGACC
CAAGACAAGTCAGTACGCAGAGATTGCCGCCATACTCATCGCTTTGCAACAAGCCG
TTAAGTTGGCTATCACACAATTTGTAATCTGCTCTGATTCTAACTACGCCCGTCACAG
TTTCATCTCCCATTTTCCTTTATGGAAGGAAAAAGGGATGAAAAATGCCAGGAACAA
AGAAGTCAAGCACTCAGAACTTTTCCTGACCTGTGATCAACTGGTAACAGAAAAAG
GGTTGATAGTTTACTGGAAGAAGGTCAAAGGTCACTCACAAATTGCTGGGCCAGAC
AAAGACGGCAACGACGAAGCCGACCGACTAGCCAAGCTAGGTGCAGAACAAGGAA
TGCACTGGGAGCTACAGGTCGAGGGATTTCCAAGTCCCCAAACCTGCGTTGTGAACG
CTATCACCCGACGACAGGCAAGGGAGAGACTAGAGAACCCCCAAGCCTCTGATCAG
ATCATTCATCTGGGACGTAAACCTGGAGATGCAGATTTGGTTGCCATGCAGGAGCAG
GATCCCGTCATCAGTGCCATACGCCAGATAGTCGCGAACCCTCCCCCGCAGGACTCA
CAACCATTTACTGACAAACCCAAGGAGATACGGGCATTTCAACTGGCATTACCCCAT
CTGAAGATGGAAAAAGACTTGCTCGTCTACAACCGTAATGGCCAAGGACCAAACCG
GTGGGTCGTCCCAACCGACCATAGGGGGGTAATGTTAACCCACGCTCACGATTCCCC
TGTCGGTGGCCACAGAGGTACAAGGCTACTCTANAGTACACTTCAGCAAGTCGGGT
ACTGGCCATCAATGTCCAACGACACTATGTCATACGTCCGAGGTTGCCTCACCTGCT
GCCAGTTCCAACCGTCACAACCATTGAATCGCGCCCCACTTCAAGGAAGGGGGGTTT
CATTCCCATGGTCCCATCTCCAGATCGACTGGATTGGACCAGTCCCAAAATCCTCAC
GAGGTAACAAGTACCTACTCACGGTTACCTGTGCCTTCACAAAATGGGTGGAATGCC
TCCCAGCATCCAATGGCACAGCCGTCACCATAGCAGTCCTACTGTTGAATCATGTGT
TCAGTCGCTGGGGCCTATCAATCTCCATTGATTCAGACCGAGGAACACATTTCACGT
CCCATGTCATGACATCGCTATGCAAAATCCTGGGAGTGGAAGTCAGATTCCACATCT
CTCATCACCCCCAATCGTCAGGCCAAGTTGAACGTGCCAACCGAACTATCATTCAAA
TGCTGAAGAAATACGTGGGAAGTAACGGCAAGGACTGGGACATTAAACTTCCCCTG
GTGCTAATGGCAATCAGGCTCACCCCGCACCGCTCCACAGGTGTCACCCCCTTTGAA
ATGATGACAGGTAGACAAATGACTCTCCCCCTGCACCTGTTATACTATGCTGAAGAC
CTCAGCGTCATGACCGCCTACACTGCACATCAATATGTGGCAGAGCTGCGTGACCAC
TTGAAGACTACCTTTGCATGGGCACAGGAAACCCTGGAGGCCAGTGTGGAAGCTTC
AAAGGCCTATTATGACAGGAAAACCTCCCACCATGAATACCATGTGGGAGACAAGG
TTTTCCACTTCAAGTTCATAAAACCAGTGGGGGTTTCCAAGAAATTCCTCCCCAGCT
GGATAGGTCCCTTTGAGGTTGTGGGAAAACTCTCCCCGGTAGCATACCGAATCAGGA
TCCCGAAACAGAGAAGACCCCCCACATACAAATGGGTCCACTCAAACCAAATCAAA
CCATACAAACCTTCCTCAATCTCAGGAGCAGAGGCCACCCCTACAGCAGAAAGGGG
GGAGCCATCCCAACTCCAAGAGGAAGAACACTGA
> Loki-Pyo consensus sequence Pol protein

LGDLVRKGNAQRIYTSVVIGGCFSSLGLLDSGSEISLMCATTFSQVETTMLSLGKPLQVE
TCNINITSYTQDQSSITRRAWIEITFQEMSLAHPFYICSIDTEPLLIGQDLLDRLAPLIDCRQ
GRIWAQVENPKPLEPNVNLPPQVDAIQTPSRISPMHMPLPETDPSSDPLCTTASGPVPRPD
SETTFKSHASFLCSLNNENPSLYVPRITEDVFLNGVRIPNTLVALWSEKSAISTDIYTDLC
KLELPPIFVQRRHRILSAEIPQTPIKATGVCAISVRIGKRQIWHFVSVIPQLHSSFMVGADL
LIRLGTQVDTINQVLWSQVRAERHSLSADPDQMLSGQTIPQACQVASEVDMVIPPRTAG
APSDLKSSKDKGSRALRLSSSLRPPSGNSTWKSVAHLCWSCSRSTYVLVKNPTHCPVPL
KAHQAIGMLIDSSFHDFELTIPVIGELPPSLIKDGDVDQVSHTFPTSMITVMRHESLQDEAI
CGASLTTEGDMMVYALAAQPGQDVSEDAQPNELYPGFEAEIKQQLENADALTSDTQRE
VLRDLFLDFQSIFSKDSTDCGVTNLHTVRIPTDPNALPTFVRQYKIPLAAYESIQDILDTLL
EKRIIRECNSTYNSPIWPVLKPTGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVKGAT
FFSTVDVANGFWTMKVDQADQYKLAFSFGGSQYTWNRCPFGYSNSPSEFNIFLHKAMS
DAATRGNLIYVDDILMRSQTFEDHITEIRHVLTQLSNAGAKLALTKGQWCRTKVKYVGL
TVGADGIEPQAGRIRAIQDIKAPTRLTELRSFLGVCNYSRQFIEDYAEIARPLNELLHKDK
PFEWEEPQEQSFRQMKEKLCSASCLAYPDKDKEFHLEASFSSHCLSAALAQKYDTDKRV
VAYASRPLSSVEIKFSDCEKALLSTVWAVEHFRSYIGGQKVIIETCHQPVAFLNSQRLRE
GRVTNSRIATWMMALQGYDIEVKYAQNNKMALGKGLAECHHCDEDDQSDPISLPATTP
SLPSNHRYYEENVCQDLPKVYVDGCSFHHQSHLRAGVGIAWVDYNCSEPTHYQLGPKT
SQYAEIAAILIALQQAVKLAITQFVICSDSNYARHSFISHFPLWKEKGMKNARNKEVKHS
ELFLTCDQLVTEKGLIVYWKKVKGHSQIAGPDKDGNDEADRLAKLGAEQGMHWELQV
EGFPSPQTCVVNAITRRQARERLENPQASDQIIHLGRKPGDADLVAMQEQDPVISAIRQI
VANPPPQDSQPFTDKPKEIRAFQLALPHLKMEKDLLVYNRNGQGPNRWVVPTDHRGV
MLTHAHDSPVGGHRGTRLLXSTLQQVGYWPSMSNDTMSYVRGCLTCCQFQPSQPLNR
APLQGRGVSFPWSHLQIDWIGPVPKSSRGNKYLLTVTCAFTKWVECLPASNGTAVTIAV
LLLNHVFSRWGLSISIDSDRGTHFTSHVMTSLCKILGVEVRFHISHHPQSSGQVERANRTI
IQMLKKYVGSNGKDWDIKLPLVLMAIRLTPHRSTGVTPFEMMTGRQMTLPLHLLYYAE
DLSVMTAYTAHQYVAELRDHLKTTFAWAQETLEASVEASKAYYDRKTSHHEYHVGDK
VFHFKFIKPVGVSKKFLPSWIGPFEVVGKLSPVAYRIRIPKQRRPPTYKWVHSNQIKPYKP
SSISGAEATPTAERGEPSQLQEEEH*
> Loki-Pyo consensus sequence *env* gene
ATGATGCTCCCGTGGATCCTCGGCATACTCTTCACTCCACAAATGGTTCAACCGCTC
CAGGTGGTTCAGCAAGGCCCACCAACGGGCATTGTTTTCAGAACACACCTGGACTC
CTCATCACCCGTTGCGGATTGTATACTCAAAGGGTCTACGTACGACTGGACCCCTGG
GATGTGTATAGAAAACACATCCGTAACCCCCCACAGTCGACTGAGGGAAGGTTAAG
AGCAATCCAAACGCATGACACAGTGGGACATGCCAAGCAGACCACTAAATATATGC
TCGAGCAAATGCAAAAGTTGTTGGTCACTGAAAAAGACCTCAGTAACCAGAAGAGA
CCAAAAAGGTTCCTTGGGTCCCTCATCCTAGCAGCATCCGCTATCGGATCGCTTTTCT
CTATTGGTGTATCAGTAGCCAATTCAATCAGCGTTGGTGCACTCCAGAGACAGATGA
GTGACCTTGAGGATGAAATGCCGGAAATCCTAGAGAGACTCGACCTCCAACAAGGT
CAATTACAGGACCTGGGGAAAACACTCCGGGGTACTGTGTTGGCAGTAAACCTTCA
CTCCGCTCTCTTAAATAATACCATGCATGCCTTAGACACTCTGTCAGTTATCGTGAGT
GACGATATAACGTACGTCCAGGTAGTAAAAGACTTAATGCAGGACCTTTTAAGGGA
AATTAGTTCCACAGTCAACAGCTTAAGCGCCGGTAGGATTCCCACCTATTTGGTGTC
TTTAGACCTTGTTGAAACAATCTTGCAGTCAGCTACCACTACGATTGTGCAGTCCTCT
CAGGTACACCTGGCATTTAGTCTCGGCAGTGCAATCCCAATCTCAGTAGACCCACAA
AACCTAGAGATAGGTTTTATCATAAACCTACCTATCATCGAACAACAGAACGTGTAT

CACCTAAAGTCCGTTCTAAACGTAGGATTTTGGCGGGGAAACACACATGTTAACATT
AAAACACCACCATTACTCGCCTACCATGACGAAAATCCCTCGCTCTATCTCATCCCA
AATTTAAGTCTATGTACCAAAACAAAGGACGTCCATTGGGTTTGTCCTAGCACCCCT
TTTATCAGGGATGTTACATATTACCTCTGTGGCCTGCGAGCCGATGCCCCAGAACAA
AAGTGCCAAGGCAGCATGTCCACCAGGAATGAGGAGTCACCGACTAGAGTAGAGCG
AGCTGGAGACAAATGGTTAGTTAACACACCCGCCACTGAAGCCATAATGGTCTATG
ACCGTCATGATACCTCCACTAAACTAAGCCTGATAAACCAGACAATGTTCTTGGTGG
TTCCTCAAGGAGCCACTATACATATTGGTGACATTGTCCTCCATCACCTCAACTCTGA
CAGACACGATGCGGAGATTGAGATCATGGACGCCTTTATGGGTCACAACCTCACGA
TCGATACAACCATCCAACAACAGCTCAGAGCTGAAGGAACAAAGGAGGTAAAGTTC
AGTCTTAGATCATCTGGACTTATCACCAAATTTGGTAACCACTATCGTCAAACACCA
AAGTTTGAAGAACGTACCATTAGCATGGTCGCGCTGGGACTGCTTTTTAGTGGATGG
ATCATCACGATAGTTGTGGCTTATCTCTTGTACAAATACATTAAGAAACTCCAGGCA
AAATTGGATTCACTTCTTTTCATCCCGAAAAAGAAGGCTGTGAAAGTTCCTCGTGCT
TCCGAAGACATTGCATTGACTTAA
> Loki-Pyo consensus sequence Env protein
MMLPWILGILFTPQMVQPLQVVQQGPPTGIVFQNTPGLLITRCGLYTQRVYVRLDPWDV
YRKHIRNPPQSTEGRLRAIQTHDTVGHAKQTTKYMLEQMQKLLVTEKDLSNQKRPKRF
LGSLILAASAIGSLFSIGVSVANSISVGALQRQMSDLEDEMPEILERLDLQQGQLQDLGKT
LRGTVLAVNLHSALLNNTMHALDTLSVIVSDDITYVQVVKDLMQDLLREISSTVNSLSA
GRIPTYLVSLDLVETILQSATTTIVQSSQVHLAFSLGSAIPISVDPQNLEIGFIINLPIIEQQN
VYHLKSVLNVGFWRGNTHVNIKTPPLLAYHDENPSLYLIPNLSLCTKTKDVHWVCPSTP
FIRDVTYYLCGLRADAPEQKCQGSMSTRNEESPTRVERAGDKWLVNTPATEAIMVYDR
HDTSTKLSLINQTMFLVVPQGATIHIGDIVLHHLNSDRHDAEIEIMDAFMGHNLTIDTTIQ
QQLRAEGTKEVKFSLRSSGLITKFGNHYRQTPKFEERTISMVALGLLFSGWIITIVVAYLL
YKYIKKLQAKLDSLLFIPKKKAVKVPRASEDIALT*
>Loki-Str consensus sequence
ATGGTGCTANCTGTCATGACTGTCCTGATCAGGTCAGGGTACAGGAGACCACCACCC
TACAGATTATCTCCCAAACCCCCAACAGAGGAGGAGAGATCTAGGGGTCTGAAGAT
GTGGGGGTTTATGACACCTCATGCCCATAATACAGAGAAATTCCTTTGTCCTAACA
ATGGAGAACTGGCCTCAGAACCTTAAACATGCAATAAAGGAACTTTGGAACAATGG
TTTCCGTCAGCCACAATGGTGGTTATGACGAAAAGTGGAATATTAAAATGTATGTAA
CTTTTGTATTTGTTTTTAAAGGTTAAGAGATGACGTTATTATGAAAACATTGTACCTT
TAAGAGTTTTCCCAGTATATGCCTGATGTTTATACATTGTACGCTGTTTGGAAAATAT
CCAAATCAAACAGAATGTTTTGATAAAGATGAAATGTGAAGTTAGTGTCTAAAATC
GGATTTTTAGCCAAATCTAAGCCTTGCCCCCTGTACTTGGTCCGCCCAGAGAATCGC
CCTAAAGGCTGTTACACCCACTTCTGACCCGAGGGTATAAGACAGGAGAGTGAAGA
ATTAACATAGTAGACTATTGACCCCAAGCTGCAGCCAAGGTCTAACAAAGTCGACG
AACCCCAAAACGAAACACAAGGTTGAAGACAAAGAAATATTTTTCTACACGAGCTA
CGGACGAGTAGCTGTGTCTAAGCGGGTGAATTCAAGCCGAACCACCCAGCCTCCAC
TCTCCATTGAATCGTGGTATCGACACCATTCGAGCCGAAGCTGTGAGCTCTGAGCTA
CAGAGCTGTCTGTCCTCAGAAGACCCCTTTCCGATCAAGGGTGAGGATCAGACCACT
TAGCCAAGAAGGACACTGACATCGTGAGGACAACCAGAGAGTTGCGCCGGAGAACT
GCGTCATTTAGAAGCCTAAACGACCCACGCGGAGCTTCCCACCTGAGAACTACAAC
ACGTAATTACATCATTATATTCTGACCCATAAGAGCGGCAGTTCGGGGCAAGGCTAA
TTTTAAATAAGCATGGCTGACAAATGAACCCAAATGTATATTTCTCTCGTGTACTTCC

TTTCTTTCTCTCTCTTTAAAATCCCCATTTTGGGTAACAAGCGCCATAGTGTGTTGGC
CCGTTATACTAAGTCCTAATCAATAGCTAGACTGTGTTTTGTGTATGTGCATTTTTAT
CATCATTTTAGCTTGCTAGTAAATAAATAATCAACTAAGATTGGTGTGGTAAATTCA
GTGGTAAAGCCCGGGTCCGTGCAGATTCCCGGATTATACGACTTTCAGATTATGAGA
CTGTAGAGGAAATTGATTAATTTAGCGACTGTTGTAATCGATATTCTGATATCCTTTG
AGTTAATTTGGGAAATAGAAACTCAATCAAAACAATGTTCCCATGGTGCCCCAGGTT
AATGAGTTAATAATTGCTTGATTCATTGCTTAATTCATTTAATCACGTAATTATAAAC
CGTTAATCATTCGATGAGCAACAGTCGTCACATTAACTAATACAACGTCACGACATA
TGGCGCCCCCTGTGAGGAATCTAAGATAGGACTAGGCCGCACTGTTGTGTTATTCCA
TATCAATTGTGTAGCAATATATATACATTCCATTAATAGCTATAGGCAGGACTAGTG
CGGGAGAGAAGTGTGTGTGTGTCGTTCCATTTGACCCAGATACTGGTCGGGAGAGA
ACGACCCCTGTTGTATATCTGACCAAAGCCAGTGTGAAGGGGGTCTACTGAATGCTA
CGAGCTAGGGCATATGTACCAGCCGATGCAGGCATTCTGAGAAATAATACTAGTTG
GGCCTAACGTAGTGTTATTTCTGTCGATCGCTTTCAGGAGTGATCTGACTCGGTCATA
AGTAATTCCTAGAGCAGAGGACATATGAGCCAGCCATTACGGAAATTAGAGTAGAT
ATATTCGTTTGACCGAAGCGAAGTTATTATCTCTCTACCTCTCGCGTTGGGTAACGG
GGAGAGGTTAAGGGTTGAACGTGAGACGTCACCAAGTAAACCCGTTCCCTTGTTGG
AGGGAACATCCCTTGTGCGGCGAGGGGGAAACCCCGCGCAAGGAAAGCTAAGCATT
GCGCCAGATGCTAAGCTAACAAAGGGGAGCCGCCATTTTTGTTTTCCTCTTTGTTCAT
AGTGAATTCCCGCGTTAGACGGGAATCCTGTGTGATCTCAGCTTGGGCTATCCGTAA
CCTCTGGCGAAGAATCGCCAGTCATTTTTCGTGAAATAAGTCAGGTTACGCTGTAGG
ATATCAAACCAGTCTTAACTGATTAGATATCATTGTCTTGTTCAATTTTATACATACA
TTAAATTGATACACTACCTCTCCAGGTTGGTTATGGGAATAATACACACACGAATGT
GCCATAACCAATGAGACATGGTTAAACTGTTCCACGTTAACTTAGATATAGCAACTA
TTTCAGGTTAATTAAAGCTAATTCCTTTAGCCTATACGGGGTTGACTAATCATTCAAA
TTGATTAATAGATTAAAGCTGTTTTACCAGCTTAATGGTGTTTAGGTAGACTTTTTAA
CAGTTTTGTAAAATTCTATTTTCACTGGTACCCTGTCGATTTTAGCTTGTGTTAACAG
TGACCTCCGGTGGTGAAGAATCACCAGTAATTATTTTTAAATAATTCAGGTTATGCT
GTACGATATCTAACCAGTCTCAACTGATTGGATATCATTGTCTCGTTCAATTTAATAT
ACATTAAATTGCTACACTACCGTCCAGGTTGGTTATTGGAATAATACGCAAACTAAT
GTGTTCTAACCAATGAGACATAGTTAAACTTTTCAATGTTAACTTAGATATAGCAAT
TACTTCAGGTTAATTAAAGCTAATTCCTTTAGCCTATACGGGGTTGACTAATCATTCA
AATTGATTAATAGATTAAAGCTGTTTTACCAGCTTAATGGTGTTTAGGTAGACTTTTT
AACAGTATTGTAAAATTCTATTTTCACTGGTACCCTGTCGATTTTAGCTTGTGTTAAC
AGTGACCTCCGGTGGTGAAGAATCACCAGTAATTATTTTTAAATAATTCAGGTTATG
CTGTACGATATCTAACCAGTCTCAACTGATTGGATATCATTGTCTCGTTCAATTTAAT
ATACATTAAATTGCTACACTACCGTCCAGGTTGGTTATTGGAATAATACGCAAACTA
ATGTGTACTAACCAATGAGACATAGTTAAACTTTTCAACGTTAACTTAGAGATAGCA
ACTCTTTCAGGTGAATTAAAATTAAAATTCCCAAATAGCCGATACAGGTTAGATTTA
TCATTAAAATCGATTTAATCAATTAAACCTGTTTTGGCAAGTTCAGTAATAACCAAC
TCACATTACTGACCCAGTCTTGATGATTCATTGACGTTTACAAACTGAGTTAAATCGT
GTTTGCAGCCTCCAACTAAAAACCCCAAACATTACGTAAATTGAAATAACTGACAAT
CATAATAATGAGCAATCCATCAGATGGGTTTCCACCCATTTCCCCTCTAATCAGTGG
ACCTTCACCCACGGAAAGATTGATCGCGGACCTCAGCAACGATGCAAATCCTAACT
ATGCCGAGGGGCTGAAAATGTTAGATTTCAATGCCATAAGCGAGAAAAATGCAGAC
ATTGCGACAGACGCTCTTCAAAATAGACCATCAGGCGGAGGCCTTGTGAAGGTTCTC

TCCAGTTTAGCCCTCAACTATGCCCACCAGCAGAGATGGACAGTGCACCAGTACCTC
AGTGCCCAGCAGAGGCACGAGCAGGAAGCTGGGGAGTTCAAGGTACAGGTGGAGA
GAACCAGGGAGCTGGCATCGAAAGCCGAAAAAGATAAGCTACTGCTAGCTGAGGA
ACTGTGTGTGAGAACAGATAAATTGGAAGATCTGTCTAACACTTATAACAAAGAAC
GCGAACAAACTCTTGACCAAGTCCGTATTCTAAAAGAACAACTCGGTTTAGCCAAA
ACTAAATACGATGAAGTCCACGCGAAACTAGATGAATCGGACGAGCTAGCTAGAAA
ACGGGGCGAGCAAATTGATGCCCTACAAACGGTTGTGAATGAAACCACTCAAAGCA
ATAATGCTCTATTCCAAAAGCTGCAGACCAAAGACGATCAGTTAATGAACACAATG
ACCAAGTTGGAGGACAAAACAGCAGAACTCATTGAAGTCGCTGATTTAATGAAGGA
TGAGAAAGACCAGGTGAGAAAGTTGGAAAATGTGACCTCTAAACAAAAGGAGGAC
ATCGCATCACTGGATCTCTCACTCCACACTCGCGACCTCTCCCTGCAAACCATGACA
GATAAGTATGAGGCCGAGCGAAAAAAGGGCGACACCTACGTGTCGAAAATAAACA
CCCTCAACTCCCAGGTGGACTCTGCGATGCAGCATAACACCACCCTTAGGTATCATC
TGGAGGAACTCCAGAACAGTCACGCGCTATCGCGGGACAACCAAACCAAGCAACCT
AGCTCACATCCGGAGCTAAGAGAACGAGGGAATGTAGATGACAGGAGATTTCAGCC
TCTTTCTCTTGGCCATTCGGCCCACAGTGAATTGGGGCCTCCTCGCCAACACAACCA
CAGCTTGACTTTTCCTCTTGGCCTCTCGGCCCACAATTCTCCACGGGAGAGTGCAGA
TGCTCCCCGCGCACTTGGCGGGGATCGCCTTGACAAAATCGTCAAAAACTTCCGCCT
CTTCGACCCCGTTCCGGGAAAGCCAAACGACACTGAGACATTCTTAGCAGACATAG
AGGACGCCTTGGATGGCTACCCAAACGCTACGAATGCAGACAGGCTCTATCTTTTGA
AGCGAACGTCCAACAGACACGTGACAAGGTTCATCCGTCTACAAGAGCAACACGTG
CAAAACGACTATGCTAAACTTGCCACGGTTTTGAAAATAGAATTCAGTGGTTCTGCG
ACTCGCAAACACGATAGTTCGTTGGCTAACAATATCAGACAAGCTCGAAATGAACA
CCCACAAGCCTACTATCACCGGCTTCGTTCAGCTTACTTTGGCATGCTCACTGAAAC
AGGAATGGAAGACCTATTACCGTTCAAACAACTTTTTCTGTCAAACATGTCTCCCAA
CTTCATTAACTACTTGGGCCCTACTGCCCACGTCGGGTTGCCAATCTCGACGCTCCG
AGAGCTGGCAAGCACCGCTTTTGAAGCATCAAAAGCAAGCCGGGCTAAGAGCCCGG
ACACCTCGACTTTCGAGCTTAGGCGGGAACCATCACTCGAATTAGAAGGGGCTGCAT
CAGGGACATATGCTGTAAAAGAGGACGATCAAAGGGGGTGGCTACCAAGTAACCAC
CACCCCGACAACCACCGCCGTAACAATAGCTACGATGAACGTCCCCAATCTAACAG
GTCCCGTAGAGATCAACCTAGCCGACATGCCCCTCCGCCTAACTCTCACAAAGGGTC
AGGACACAAGGGTAACAAGCGTAACAAGGGCAACAAAGTGTTCAACAATGATTTAA
ACGCTAACCGAAATGATATAGAAGCTCTGATAAAAGATGTACTGCATCGTAAGAAA
TTTGAGAAAGAGGACAAGGATAAATCCTCAGGACTAGGCGTAAAATCGTTGGAAAC
CGAGTCCGCCGAAATTCATGCAATTCAAGAGGACGCAAACATTTCCATTACCCCCGC
CCCCACTCGAATCCCTGTCCAGGTACCGCCCGTTCTAGACACAAGCCCGCAGGTTTT
GTCTACAGACTTGGACACGGAGGTGGAGATGCACGAGATAAGCAGCTTTGCAACAG
ATGATTCCTCTACCATTCCGATAGAGGACCCACTGGGTCACGTAGGTAACTTATCTG
TCTTGGAAATCGACGCAGATTACAAACCGCTCCGTTCTCCCAACCCACTCCATTTTG
TTGGGGATATGACGAGAAAAACCAACTCGAACAGGCCATACCTTAGAACAGTCCTG
GAGGACTGTGTGACCTGTCACGCTCTGATAGATTCGGGTTCAACAATTTCCCTCATA
TCCGAAACCTTGCTGGATGAACTAAAAAGGGCAGTGGACCCAACAAAACGTTGGTT
GAAAACGGAACATTGCGATACGAATCTTCGAGGTTTCACTCAAGCTACCTCGCCCCT
AACCAAACGTCTCCTACTTAAACTCAACTTCGAAAGCGTGTCACTGATACACCCCGT
GTATGTGACTAGCATCGAAATCGAACGAATGCTGATTGGGAGTGATTTAATTGACCG
TTTGGTACCACTAATGGATTGGAAAACCAATCAAATCTGGTCACAAATACCTATGCC

AACTCCGTTGACTTCGCCGCATTCTTCCAAGGCTACCTGCCATACATTGTGCAACGA
CGTGGGTCCGACGTCAGCATCTGACGCAAACCCTGTGAACTTGGCCATGAGCCCGCC
ATTTGTGGCAAAGGACAATGTGAGTTCGACTCGGAACTTAACCGAACATGTGGTTTT
CTTGTGCGGCTTGGGTGATTCACCAGCCGACACTTACCGCCCTCCTCTGATAGGGGG
CGTGTGCCTAAACGGCACTATGGCTGAGGATACCAGGCTGGCCACCTGGTCAGAAA
AATCCGCAATCAGTTTGGAAATGTTCAACGAAATTTCGAAAACGGCTAACTGCCATC
CCCTTGTGCCGAAAACGTTTAGATTTCCACTGGGAACGACTCCCCAGACAACACTCA
CCGCAATAGGGGTGTGTGCGCTATCGATCCGCATTGGAACCAAGGAATTATCCCACT
ACCTACTGGTTGTCGCGGACCTCCCACATACAGTCTATGTGGGAGCGGACATTTTGG
TAAGATTGGGTGTTAAACTTGATACCATACACCAAGTTCTTTGGTCTTTGGCGCAGC
CAAACCAACATGCCTTGTCTTTCGACCCGGTGCGAATGGCTTCCGGGCAGACTATTC
CTGAAGCTTGCAAGACGATAACTGAGTCCGCCATGTTGATACCGGCAAGAACCACA
GAGGTTTCTGTAAGACTGAACCTGGCGCCCGGATACCGGATGGAAGGCACCACTGC
GTTCTTCCAACCCTCTCCAAAACTATTCGACTTGGGGCTAACTATCAATGGCAACCC
ACTGTTGGAACTAACCGCTAGATCCACTTACCTCCTGGTACAAAATCTGACTCAAGC
GGATATTTCCATTCCTCGGCACACTCAGCTAGGAACATTGATCGATTACGCCTTCCA
TGATTTTGAACTAGTTGTTCCTGTGATTGGGCCTCTTCCGTCCTCCCTAGACCTAGAT
GGCGAGGGAGGTACGCTGTTTACTTGTCAGTCAAAAGCAATAGCACTAACTCCCGTA
TTGCCACTGGACGACACATCGGCATTCCGGCTGGATGTCGATCCTGACAGCAACCTG
TTAATATACTCCATCGTCACGGAGGATGATATTGCGTCTCCTCCTGCATGCGAAGTA
CACTCTTGTGAGGAAACAACCGATGACTCTTCCAAAGGTGACATGCCGTCACCACCC
GATGAAGACCTGTACAATGTCGCCGAACCATATCCTGGTTTCCATGCACAGGTAATA
CAACTACTGTCGGAGGCTGATGCTCTTGTCAATGACACTGAACGCCATCAGTTGAGA
GAATTGTTTAACAAACACAGTGAGATCTGGTCGACAGACTCGCTGGATTGCGGGGTT
ACGGGTATCCATGTGGTGCGTATTCCTACGCCACCCGGAGCAACTCCTACGTTCGTT
AGACAATACAAAATCCCGCTCGCAGCATACGCAGCAGTACAAGAGATTATCGATTC
CTTATTAGCTAAACGGATAATCAGGGAATGTAACAGTACGTACAGCGCTCCCGTGTG
GCCCGTTCTGAAACCAACGGGTAAATGGCGTCTTACCATTGACTATAGACAACTCAA
CAAACTGGTTCCGTTGTCTCGTTGGCCAATGACACAGCTCGACCAAGAGTTGCCAAA
GGTGGCAAACGCCAAGTACTTCTCCACCGTCGACGTGGCAAATGGTTTCTGGACCAT
GACAGTCGACCCGCGTGACCAACACAAGTTGGCTTTCTCTTTCTCGAACAAGCTCTT
CACGTTCAATCGTTGCCCATTCGGGTATGCGAACTCCCCCTCAGAGTTCAACATCTT
CCTGCACAAGGCGATGCCAGACGCAGCCTCTAGGGGGACGATAATTTACGTGGACG
ACGTTCTCATGAGAAGTGAGACTTGGTCACATCACCTCAATGAAATGGACCACGTTC
TCACCCAACTCGGGACTGCAGGGGCTAAGTTGGCCATCATGAAAGGACAATGGTGC
AGGACCAAGGTAAACTACGTTGGGCTTCTGGTGGGTGCCGAGGGCATCCTGCCACA
GTCTAACCGAATCCAAGCAGTCCGCAACATCAAAACACCTACAAACCTTCACGAGG
TGCGCAGCTTCTTGGGAGTATGTAATTACTCCCGTCAATTCATCGAAAACTACGCCG
ACTTGTCAAAACCGTTGACTCACCTTCTTCAGAAAGACACCCCATTCGTTTGGGATG
TCACTCACAACGAAGCGGTGGCTTCCTTGAAAAACCTGCTTTGCGAGGCACCCTGCC
TGGTATACCCCAACAAAGACAAGACATTCTTTTTGGAAGTCGGTTTCTCAGAGCACT
GCCTTAGCGCTGGGTTGTACCAAAAATACGACCAAGACAAACGTGTGGTTGCTTACG
CGAGCAAAACACTCAACCCCGCCGAACACAAATACTCAGACTGCGAAAAAGCGCTG
CTGTCGACAGTCTGGGCGATCAAACACTTCACCAGTTATGTCGGCGGACAAAAGGT
GATCATAGAAACATGTCACCAGCCTGTGACTTTTCTGAAAAGCCAACGAATCCGTGA
GGGTGCTGTACACAACAGCAGGATAGCGGCTTGGCTCATGACGCTGCAGAGCCATG

ATGTAGTAATCAACTACGCAAAAGCAAAGAACCTGCCTTTGGGCAGTGCTTTGGCG
GTGTGTCAACACTGTGGTGACGATGAAACCGACTCCGGACCACCCCCGCGTGACAT
CGCTCCGCCATTGCCTTCGAACCATCACTATTTCGAAGAGAACGTGTGTCTCGACAT
GCCGATGGCATTTGTAGATGGCTGTTCTTACCGCCATCTAGACCACTTGCAAGCTGG
GGTGGGATTGGTATGGCACAACGACATTCCGTGCAAACCACTTCAATTTCAGCTTGG
CAACAAGACCAGCCAGTTTGCAGAAGTGGCTGGCGTGCTGATCACCCTGCAAACAG
CGGTGAAACACGAATTGGCAGAACTGGCGATCTGCACCGACTCTAACTACGCGAGA
CTCACCTTCTTATGCCACTTGCCGTTTTGGAAACAAAAGCACATGACTACTTCAAGT
GGTAAAGAGGTGAAAAACAAAGAGCTCATTCTAGCTTGTGATGATCTCATCACCAA
ACACGACATACAGGTGTATTGGAAGAAAGTCAAGGGACACTCCAAATCACCTGGTC
GTGACAAACTTGGCAACGACCATGCTGACAGCATGGCGAAATCTGGTGCAATTCAC
GGCACCCCTTGGGTGTTTGCTGCTCGAGACGAAAAACCCACTGAGGAAGCTATGGT
GTCTGCGGTAACGCGTAGCCATGTAGCACCACGGAAAACGTCGTCGCACAAACATA
ATGTGTTGCTAACGCATTCATTCATGAATGGCGACCTGGTAGCAATGCAACACCAAG
ATGAGTCCCTCGCCACTTTGATGGCATTCCTGTCGGATCCTGTCAACCACCCAGTGT
CCGAACTGGCTTTGGCAGGTTCACACGACTTGCGTTCGCTGTACGCAACTAAACAGC
ACCTCACACTTGTAGATGACCTATTGGTGTATGTTTCAGAAACCGACACCGCTCGAA
GGTGGGTGGTTCCTAAAACACAGAGGGGGATAATGATAGCTCATGCCCACGACGAG
CCATGTGGAGGCCATAGGGGGGTCAAGGCTACATGTGAAACATTGCGACAGGTGGC
CCACTGGCCTCACATGGAACAAGACGTGGCACGATACGTGAGGGGGTGTCTAGTTT
GCTGCCAGTTTCAACCCACCAAGCCACTTCACAGAGCACCCCTGCAACGCAAGGGT
GTGTCTTACCCCTGGAGCTCGATCCAGATCGACTGGGTCGGCCCAGTTGCCAGGTCT
GCCAGAGGCAACAAGTACCTCCTCACAGTGACTTGCGCATTCACAAAGTGGGTGGA
GTGCCTGCCGGCGACCAATGACACAGCGGAAACCACTGCCGTTCTTTTGCTAAACCA
CGTTTTCAGTCGTTTCGGCCTGCCAGGAGAAACCGTGGACAGCGACAGAGGAACCC
ACTTTTCGGCAGCTGTAATGACCGAACTGTGGAAGCTCCTGGGAGTCAAAGCTAAAC
TCCACATCGCGTACCACCCTAGGAGCTCGGGGGGGGTAGAAAGGAGCAACCAATCA
ATTGTCAGAATCTTGAGGAAATATGTGGCAGCCAACCACAAAGATTGGGACCTCAA
ACTCCCATTGGTACTGATGGCAATCAGAGCCACACGCAACAGGTCTACGGGAATGA
CACCCTTTGAGATGATGACGGGCCGGCAAATGACCCTTCCACTGCATCTCCTGTACC
AACCGGGAGATGTCGCCGCAGCCACCGCCTACACGGCTCATCAATACGTGACCGAC
TTGCGGAACCATCTCCAGACTACCTTTGCGTACGCTCAAGGACAATTGGAAAGAAGT
GCAGAAGGTGACAAGACCTATTACGACAAAAAAGCCTCACACCAGGTGTTCGAGGT
CGGAGATAAGGTGTGGTACTACATATACACCAAACCCGCGGGCGTCGCAACAAAGT
TCCTGCCCCACTGGACGGGGCCACACGAGATTGTGGTGAAGCTATCACCTGTAGCTT
ACCAGATCAAAATCAGCAAAGGTCGACAAAACGCCACATTAAAGTGGGTCCACCGG
AACCAAATCAAGTTGTACACTCCCCCCATGGGAATAGAAGGGGTGCTAGCCAGCAC
CGAATAGATAAAAACCCTAGCAAAAAACCAGAGGTACTAAGAAAATTCTTAAGTAC
CCTCAGCTGATCCCTTCCAGGAGACCAGTACGTTGCACCTAATATCTTTTATTCTCTT
CCCAGCTACCAGATATACATCTGTTGTCACTAGTGATATCGTCCTGCGCTGTACTGAT
TAAAAATAAGAAAAACAGAAAAATAGTTGTCAAAACAATTTTTGTTTACGAATACA
AATAACTGAAATAATCCAACAATCTCTGATTATGCGTTTCTGTAGGATGGAGTTCCT
TTGGATGATCCTGACACTCTGCGCCCTGGTCAAAACCAACCCAGCAGACGTAATCAC
GAACGGACCCCCTACGGGAATCGTTCTCCGCGACGATCCAGGACTCCTCATAACTAA
CTGCAGAGTACACACGCAGAGGGTATATGTCCGCCTAGATGCAGAGAATGTGTACC
GCCAACACATTCCACCTGCGGCGCATTTGAGTTGGGCCGGGGCTGACTGGACCAGC

CAGGCAATTAAGCACGCCCAGCTAGACACTGCCCATATGTTGGCCCAACTCCAAAA
GTTCACAGTAACCCAGTCAGAACTAGCTGAGCCCAGGCGAGAAAAACGATTCGTCG
GGGCGCTACTATCGATAGGGGCAGCCGTAGGGTCACTATTTGCCCTAGGTACCACTG
CCGTCAACGCAGTCAGTCTAGCCACAGTCAAGAGGAATGTTAGGGAGATTACTGAA
GAGATGCCACTTATCCAGCAGCAGATGAAGGCACAGGCCCTTAGGTTGCAACATGT
GGGAAAGTCTCTCCAGGACACTATTCTGGTGGTCAACACACACGCTACTCTCCTGAA
CAAAACTATCCTGTCTGTAGGAAAGCTAGCAGAGGTCATGAACCATGACTATGCCC
ATGTCCAATTGGTTCGATTGTTACTGGAGGATTTTCTCCGTGAAGTTAGCTCTTCTAT
GGACCACTTGGCCATGAATAGAATCCCCTCATATCTAGTGCCCCTCTCAATGGTGCA
CGACATATTGACCTCAGCTACTTCAACCACACTAAGGCCGTTACAATCACATTTGGC
CTATAGTCTGGGGTCGGCCATCCCAATACACGTTGATGTTGATCGTAATGAAGTGGG
TTTTCTACTGACGCTGCCGGTTGTTGAACTGGAGAATATCTATAGATTGAAAACGGT
TCTCAATGTAGGATTTTGGCGCGACAGTACCCACGTAAAGATTGCCACGCCCCCAGT
TGTAGCTTACCAGGAAAACAACCCAGATTTCTATCTCACCCCTAACCTGCTAATGTG
TACTCTAACCAAAGATGTCCACTACCTTTGTCCTAGCAAACCATTTGTCAGGGATAA
CACTGAGCGTCTTTGTGGTATTAAAGCTATCACCAGCGAAGAACGCTGTAGAGCCAA
ACTAACAGCCAGGGATGGGGCCACCACCACACAAGTGGAGGTGGTAGGGAACCGA
TGGTTGGTCAACACACCGTTCGCGTCCGCCACTATGACTTACGAACGCCATGACTCC
ATCACCCAATTGAACCTCCCCAACCAGACTGTTTTTGTTACGGTACCAGAGGGGGCG
ATTATTCACGTAGACGACCTCGCTCTATACCACCTTCCCGACGACCGGATAGACACA
GAGATTGAAATGCCGGACGCTTTTGCTAAACGTTCCCTTGAGTTACCACAAGCCATT
CAATACCAAATCCAGTACGAGGGACCCATGACTGTCGATCTTAGCGTACTGGATGA
GGCACTGTTAGTCGACCCGACTGACTACAGACCAAAAGATTGGTCTGTCGCCCGCTC
TTGGACGGTGCCGGACAGTATCCTGACTGTTACCATGATCCTTGGTCTCATTTCCCTT
GGCGTGTGCACCTACTACGTACACAGGCGCACACGTGCAATGGAAGACCTAATGAG
GTCTTATACTAGGATCACCCGTGGGACGGAAGCGATCCCGATTTTTGGAAAGGTGGC
AATGGATCCGGAAACCCTCTTACAGGGCCCAGTCCCAGCCTCCTCTCCCGAACTCGC
TAGGTCAGCCCAGTAATGTCCCTAATGTAAGCTTTGGCCTTCAGGGGCCAAGTTTTT
CCAAGTGCCTTAACTACCCACCTGCAAGTAGGATCAACCTAGACATGACATTAGAG
ACAATGCCATGATAGAGCTATTTAGCCAAGGCCTTGGACATATATAGAACATAGTCC
ATATTAGATGATTAAGGTGTGGTAGACATATTTTGTATACTTTTATGTTTTTATTCCA
ATTTTTCGTTTCATTACCTTTGACACTTAGGAAGCCTTAGAGCCAAGACACCGCTCCT
TTGGGGAGGAAATGTAATGATGTATTGCCTGAGTGTTGTGCGTTATTAACGTCTGCC
AAGTTACTGCCTAGGTCCACTAACCGGGATAACCGGACGACGGGCCGGAACACAGA
GCCACTGCCAGACCTCCTTGATCCATTCTGGTGGTGATCCACAGCTTGCGGAAAGCC
TACCAGGGGGAACCACTAAAGGGGGGGGGGCTGCTCTGATGATCCACAGACCAGGA
CATCCGATGGTCATTCTTATGGTGGGTTGCAACATGCAGAATCATTCCAAGTTGAAC
CTACCAGAGGGGGACTGTCATGACTGTCCTGATCAGGTCAGGGTACAGGAGACCAC
CACCCTACAGATTATCTCCCAAACCCCCAACAGAGGAGGAGAGATCTAGGGGTCTG
AAGATGTGGGGGTTTTATGACACCTCATGCCCATAATACAGAGAAATTCCTTTGTCC
TAACAATGGAGAACTGGCCTCAGAACATTAAACATGCAATAAAGGGACTTTGGAAC
AATGGTTTCCGTCAGCCACAATGGTGGTTATGACGAAAAGTGGAATATTAAAATGTA
TGTAACTTTTGTATTTGTTTTTAAAGGTTAAGAGATGACGTTATTATGAAAACATTGT
ACCTTTAAGAGTTTTCCCAGTATATGCCTGATGTTTATACATTGTACGCTGTTTGGAA
AATATCCAAATCAAACAGAATGTTTTGATAAAGATGAAATGTGAAGTTAGTGTCTAA
AATCGGATTTTTAGCCAAATCTAAGCCTTGCCCCCTTTACTTGGTCCGCCCAGAGAA

TCGCCCTAAAGGCTGTTACACCCACTTCTGACCCGAGGGTATAAGACAGGAGAGTG
AAGAATTAACATAGGAGACTATTGACCCCAAGCTGCAGCCAAGGTCTAACAAAGTC
GACGAACCCCAAAACGAAACACAAGGTTGAAGACAAAGAAATATTTTTCTACACGA
GCTACGGACGAGTAGCTGTGTCTAAGCGGGTGAATTCAAGCCGAACCACCCAGCCT
CCACTCTCCATTGAATCGTGGTATCGACACCATTCGAGCCGAAGCTGTGAGCTCTGA
GCTACAGAGCTGTCTGTCCTCAGAAGACCCCTTTCCGATCAAGGGTGAGGATCAGAC
CACTTAGCCAAGAAGGACACTGACATCGTGAGGACAACCAGAGAGTTGCGCCGGAG
AACTGCGTCATTTAGAAGCCTAAACGACCCACGCGGAGCTTCCCACCTGAGAACTCC
AACACGTAATTACATCATTATATTCTGACCCATAAGAGCGGCAGTTCGGGGCAAGGC
TAATTTTAAATAAGCATGGCTGACAAATGAACCCAAATGTATATTTCTCTCGTGTAC
TTCCTTTCTTTCTCTCTCTTTAAAATCCCCATTTTGGGTAACAAGCGCCATAGTGTGTT
GGCCCGTTATACTAAGTCCTAATCAATAGCTAGACTGTGTTTTGTGTATGTGCATTTT
TATCATCATTTTAGCTTGCTAGTAAATAAATAATCAACTAAGATTGGTGTGGTAAAT
TCAGTGGTAAAGCCCGGGTCCGTGCAGATTCCCGGATTATACGACTTTCAGATTATG
AGACTGTAGAGGAAATTGATTAATTTAGCGACTGTTGTAATCGATATTCTGATATCC
TTTGAGTTAATTTGGGAAATAGAAACTCAATCAAAACAATGTTCCCATGGTGCCCCA
GGTTAATGAGTTAATAATTGCTTGATTCATTGCTTAATTCATTTAATCACGTAATTAT
AAACCGTTAATCATTCGATGAGCAACAGTCGTCACATTAACTAATACAACGTCACGA
CATACCCTACA
> Loki-Str consensus sequence 5’LTR
TGTCATGACTGTCCTGATCAGGTCAGGGTACAGGAGACCACCACCCTACAGATTATC
TCCCAAACCCCCAACAGAGGAGGAGAGATCTAGGGGTCTGAAGATGTGGGGGTTTT
ATGACACCTCATGCCCATAATACAGAGAAATTCCTTTGTCCTAACAATGGAGAACTG
GCCTCAGAACCTTAAACATGCAATAAAGGAACTTTGGAACAATGGTTTCCGTCAGCC
ACAATGGTGGTTATGACGAAAAGTGGAATATTAAAATGTATGTAACTTTTGTATTTG
TTTTTAAAGGTTAAGAGATGACGTTATTATGAAAACATTGTACCTTTAAGAGTTTTCC
CAGTATATGCCTGATGTTTATACATTGTACGCTGTTTGGAAAATATCCAAATCAAAC
AGAATGTTTTGATAAAGATGAAATGTGAAGTTAGTGTCTAAAATCGGATTTTTAGCC
AAATCTAAGCCTTGCCCCCTGTACTTGGTCCGCCCAGAGAATCGCCCTAAAGGCTGT
TACACCCACTTCTGACCCGAGGGTATAAGACAGGAGAGTGAAGAATTAACATAGTA
GACTATTGACCCCAAGCTGCAGCCAAGGTCTAACAAAGTCGACGAACCCCAAAACG
AAACACAAGGTTGAAGACAAAGAAATATTTTTCTACACGAGCTACGGACGAGTAGC
TGTGTCTAAGCGGGTGAATTCAAGCCGAACCACCCAGCCTCCACTCTCCATTGAATC
GTGGTATCGACACCATTCGAGCCGAAGCTGTGAGCTCTGAGCTACAGAGCTGTCTGT
CCTCAGAAGACCCCTTTCCGATCAAGGGTGAGGATCAGACCACTTAGCCAAGAAGG
ACACTGACATCGTGAGGACAACCAGAGAGTTGCGCCGGAGAACTGCGTCATTTAGA
AGCCTAAACGACCCACGCGGAGCTTCCCACCTGAGAACTACAACACGTAATTACAT
CATTATATTCTGACCCATAAGAGCGGCAGTTCGGGGCAAGGCTAATTTTAAATAAGC
ATGGCTGACAAATGAACCCAAATGTATATTTCTCTCGTGTACTTCCTTTCTTTCTCTC
TCTTTAAAATCCCCATTTTGGGTAACAAGCGCCATAGTGTGTTGGCCCGTTATACTA
AGTCCTAATCAATAGCTAGACTGTGTTTTGTGTATGTGCATTTTTATCATCATTTTAG
CTTGCTAGTAAATAAATAATCAACTAAGATTGGTGTGGTAAATTCAGTGGTAAAGCC
CGGGTCCGTGCAGATTCCCGGATTATACGACTTTCAGATTATGAGACTGTAGAGGAA
ATTGATTAATTTAGCGACTGTTGTAATCGATATTCTGATATCCTTTGAGTTAATTTGG
GAAATAGAAACTCAATCAAAACAATGTTCCCATGGTGCCCCAGGTTAATGAGTTAAT

AATTGCTTGATTCATTGCTTAATTCATTTAATCACGTAATTATAAACCGTTAATCATT
CGATGAGCAACAGTCGTCACATTAACTAATACAACGTCACGACA
> Loki-Str consensus sequence 3’LTR
TGTCATGACTGTCCTGATCAGGTCAGGGTACAGGAGACCACCACCCTACAGATTATC
TCCCAAACCCCCAACAGAGGAGGAGAGATCTAGGGGTCTGAAGATGTGGGGGTTTT
ATGACACCTCATGCCCATAATACAGAGAAATTCCTTTGTCCTAACAATGGAGAACTG
GCCTCAGAACATTAAACATGCAATAAAGGGACTTTGGAACAATGGTTTCCGTCAGCC
ACAATGGTGGTTATGACGAAAAGTGGAATATTAAAATGTATGTAACTTTTGTATTTG
TTTTTAAAGGTTAAGAGATGACGTTATTATGAAAACATTGTACCTTTAAGAGTTTTCC
CAGTATATGCCTGATGTTTATACATTGTACGCTGTTTGGAAAATATCCAAATCAAAC
AGAATGTTTTGATAAAGATGAAATGTGAAGTTAGTGTCTAAAATCGGATTTTTAGCC
AAATCTAAGCCTTGCCCCCTTTACTTGGTCCGCCCAGAGAATCGCCCTAAAGGCTGT
TACACCCACTTCTGACCCGAGGGTATAAGACAGGAGAGTGAAGAATTAACATAGGA
GACTATTGACCCCAAGCTGCAGCCAAGGTCTAACAAAGTCGACGAACCCCAAAACG
AAACACAAGGTTGAAGACAAAGAAATATTTTTCTACACGAGCTACGGACGAGTAGC
TGTGTCTAAGCGGGTGAATTCAAGCCGAACCACCCAGCCTCCACTCTCCATTGAATC
GTGGTATCGACACCATTCGAGCCGAAGCTGTGAGCTCTGAGCTACAGAGCTGTCTGT
CCTCAGAAGACCCCTTTCCGATCAAGGGTGAGGATCAGACCACTTAGCCAAGAAGG
ACACTGACATCGTGAGGACAACCAGAGAGTTGCGCCGGAGAACTGCGTCATTTAGA
AGCCTAAACGACCCACGCGGAGCTTCCCACCTGAGAACTCCAACACGTAATTACAT
CATTATATTCTGACCCATAAGAGCGGCAGTTCGGGGCAAGGCTAATTTTAAATAAGC
ATGGCTGACAAATGAACCCAAATGTATATTTCTCTCGTGTACTTCCTTTCTTTCTCTC
TCTTTAAAATCCCCATTTTGGGTAACAAGCGCCATAGTGTGTTGGCCCGTTATACTA
AGTCCTAATCAATAGCTAGACTGTGTTTTGTGTATGTGCATTTTTATCATCATTTTAG
CTTGCTAGTAAATAAATAATCAACTAAGATTGGTGTGGTAAATTCAGTGGTAAAGCC
CGGGTCCGTGCAGATTCCCGGATTATACGACTTTCAGATTATGAGACTGTAGAGGAA
ATTGATTAATTTAGCGACTGTTGTAATCGATATTCTGATATCCTTTGAGTTAATTTGG
GAAATAGAAACTCAATCAAAACAATGTTCCCATGGTGCCCCAGGTTAATGAGTTAAT
AATTGCTTGATTCATTGCTTAATTCATTTAATCACGTAATTATAAACCGTTAATCATT
CGATGAGCAACAGTCGTCACATTAACTAATACAACGTCACGACA
> Loki-Str consensus sequence *smc-gag-pol* gene
CTGACAATCATAATAATGAGCAATCCATCAGATGGGTTTCCACCCATTTCCCCTCTA
ATCAGTGGACCTTCACCCACGGAAAGATTGATCGCGGACCTCAGCAACGATGCAAA
TCCTAACTATGCCGAGGGGCTGAAAATGTTAGATTTCAATGCCATAAGCGAGAAAA
ATGCAGACATTGCGACAGACGCTCTTCAAAATAGACCATCAGGCGGAGGCCTTGTG
AAGGTTCTCTCCAGTTTAGCCCTCAACTATGCCCACCAGCAGAGATGGACAGTGCAC
CAGTACCTCAGTGCCCAGCAGAGGCACGAGCAGGAAGCTGGGGAGTTCAAGGTACA
GGTGGAGAGAACCAGGGAGCTGGCATCGAAAGCCGAAAAAGATAAGCTACTGCTA
GCTGAGGAACTGTGTGTGAGAACAGATAAATTGGAAGATCTGTCTAACACTTATAA
CAAAGAACGCGAACAAACTCTTGACCAAGTCCGTATTCTAAAAGAACAACTCGGTT
TAGCCAAAACTAAATACGATGAAGTCCACGCGAAACTAGATGAATCGGACGAGCTA
GCTAGAAAACGGGGCGAGCAAATTGATGCCCTACAAACGGTTGTGAATGAAACCAC
TCAAAGCAATAATGCTCTATTCCAAAAGCTGCAGACCAAAGACGATCAGTTAATGA
ACACAATGACCAAGTTGGAGGACAAAACAGCAGAACTCATTGAAGTCGCTGATTTA
ATGAAGGATGAGAAAGACCAGGTGAGAAAGTTGGAAAATGTGACCTCTAAACAAA
AGGAGGACATCGCATCACTGGATCTCTCACTCCACACTCGCGACCTCTCCCTGCAAA

CCATGACAGATAAGTATGAGGCCGAGCGAAAAAAGGGCGACACCTACGTGTCGAAA
ATAAACACCCTCAACTCCCAGGTGGACTCTGCGATGCAGCATAACACCACCCTTAGG
TATCATCTGGAGGAACTCCAGAACAGTCACGCGCTATCGCGGACAACCAAACCAA
GCAACCTAGCTCACATCCGGAGCTAAGAGAACGAGGGAATGTAGATGACAGGAGAT
TTCAGCCTCTTTCTCTTGGCCATTCGGCCCACAGTGAATTGGGGCCTCCTCGCCAACA
CAACCACAGCTTGACTTTTCCTCTTGGCCTCTCGGCCCACAATTCTCCACGGGAGAG
TGCAGATGCTCCCCGCGCACTTGGCGGGGATCGCCTTGACAAAATCGTCAAAAACTT
CCGCCTCTTCGACCCCGTTCCGGGAAAGCCAAACGACACTGAGACATTCTTAGCAGA
CATAGAGGACGCCTTGGATGGCTACCCAAACGCTACGAATGCAGACAGGCTCTATC
TTTTGAAGCGAACGTCCAACAGACACGTGACAAGGTTCATCCGTCTACAAGAGCAA
CACGTGCAAAACGACTATGCTAAACTTGCCACGGTTTTGAAAATAGAATTCAGTGGT
TCTGCGACTCGCAAACACGATAGTTCGTTGGCTAACAATATCAGACAAGCTCGAAAT
GAACACCCACAAGCCTACTATCACCGGCTTCGTTCAGCTTACTTTGGCATGCTCACT
GAAACAGGAATGGAAGACCTATTACCGTTCAAACAACTTTTTCTGTCAAACATGTCT
CCCAACTTCATTAACTACTTGGGCCCTACTGCCCACGTCGGGTTGCCAATCTCGACG
CTCCGAGAGCTGGCAAGCACCGCTTTTGAAGCATCAAAAGCAAGCCGGGCTAAGAG
CCCGGACACCTCGACTTTCGAGCTTAGGCGGGAACCATCACTCGAATTAGAAGGGG
CTGCATCAGGGACATATGCTGTAAAAGAGGACGATCAAAGGGGGTGGCTACCAAGT
AACCACCACCCCGACAACCACCGCCGTAACAATAGCTACGATGAACGTCCCCAATC
TAACAGGTCCCGTAGAGATCAACCTAGCCGACATGCCCCTCCGCCTAACTCTCACAA
AGGGTCAGGACACAAGGGTAACAAGCGTAACAAGGGCAACAAAGTGTTCAACAAT
GATTTAAACGCTAACCGAAATGATATAGAAGCTCTGATAAAAGATGTACTGCATCGT
AAGAAATTTGAGAAAGAGGACAAGGATAAATCCTCAGGACTAGGCGTAAAATCGTT
GGAAACCGAGTCCGCCGAAATTCATGCAATTCAAGAGGACGCAAACATTTCCATTA
CCCCCGCCCCCACTCGAATCCCTGTCCAGGTACCGCCCGTTCTAGACACAAGCCCGC
AGGTTTTGTCTACAGACTTGGACACGGAGGTGGAGATGCACGAGATAAGCAGCTTT
GCAACAGATGATTCCTCTACCATTCCGATAGAGGACCCACTGGGTCACGTAGGTAAC
TTATCTGTCTTGGAAATCGACGCAGATTACAAACCGCTCCGTTCTCCCAACCCACTC
CATTTTGTTGGGGATATGACGAGAAAAACCAACTCGAACAGGCCATACCTTAGAAC
AGTCCTGGAGGACTGTGTGACCTGTCACGCTCTGATAGATTCGGGTTCAACAATTTC
CCTCATATCCGAAACCTTGCTGGATGAACTAAAAAGGGCAGTGGACCCAACAAAAC
GTTGGTTGAAAACGGAACATTGCGATACGAATCTTCGAGGTTTCACTCAAGCTACCT
CGCCCCTAACCAAACGTCTCCTACTTAAACTCAACTTCGAAAGCGTGTCACTGATAC
ACCCCGTGTATGTGACTAGCATCGAAATCGAACGAATGCTGATTGGGAGTGATTTAA
TTGACCGTTTGGTACCACTAATGGATTGGAAAACCAATCAAATCTGGTCACAAATAC
CTATGCCAACTCCGTTGACTTCGCCGCATTCTTCCAAGGCTACCTGCCATACATTGTG
CAACGACGTGGGTCCGACGTCAGCATCTGACGCAAACCCTGTGAACTTGGCCATGA
GCCCGCCATTTGTGGCAAAGGACAATGTGAGTTCGACTCGGAACTTAACCGAACAT
GTGGTTTTCTTGTGCGGCTTGGGTGATTCACCAGCCGACACTTACCGCCCTCCTCTGA
TAGGGGGCGTGTGCCTAAACGGCACTATGGCTGAGGATACCAGGCTGGCCACCTGG
TCAGAAAAATCCGCAATCAGTTTGGAAATGTTCAACGAAATTTCGAAAACGGCTAA
CTGCCATCCCCTTGTGCCGAAAACGTTTAGATTTCCACTGGGAACGACTCCCCAGAC
AACACTCACCGCAATAGGGGTGTGTGCGCTATCGATCCGCATTGGAACCAAGGAAT
TATCCCACTACCTACTGGTTGTCGCGGACCTCCCACATACAGTCTATGTGGGAGCGG
ACATTTTGGTAAGATTGGGTGTTAAACTTGATACCATACACCAAGTTCTTTGGTCTTT
GGCGCAGCCAAACCAACATGCCTTGTCTTTCGACCCGGTGCGAATGGCTTCCGGGCA

GACTATTCCTGAAGCTTGCAAGACGATAACTGAGTCCGCCATGTTGATACCGGCAAG
AACCACAGAGGTTTCTGTAAGACTGAACCTGGCGCCCGGATACCGGATGGAAGGCA
CCACTGCGTTCTTCCAACCCTCTCCAAAACTATTCGACTTGGGGCTAACTATCAATG
GCAACCCACTGTTGGAACTAACCGCTAGATCCACTTACCTCCTGGTACAAAATCTGA
CTCAAGCGGATATTTCCATTCCTCGGCACACTCAGCTAGGAACATTGATCGATTACG
CCTTCCATGATTTTGAACTAGTTGTTCCTGTGATTGGGCCTCTTCCGTCCTCCCTAGA
CCTAGATGGCGAGGGAGGTACGCTGTTTACTTGTCAGTCAAAAGCAATAGCACTAA
CTCCCGTATTGCCACTGGACGACACATCGGCATTCCGGCTGGATGTCGATCCTGACA
GCAACCTGTTAATATACTCCATCGTCACGGAGGATGATATTGCGTCTCCTCCTGCAT
GCGAAGTACACTCTTGTGAGGAAACAACCGATGACTCTTCCAAAGGTGACATGCCG
TCACCACCCGATGAAGACCTGTACAATGTCGCCGAACCATATCCTGGTTTCCATGCA
CAGGTAATACAACTACTGTCGGAGGCTGATGCTCTTGTCAATGACACTGAACGCCAT
CAGTTGAGAGAATTGTTTAACAAACACAGTGAGATCTGGTCGACAGACTCGCTGGA
TTGCGGGGTTACGGGTATCCATGTGGTGCGTATTCCTACGCCACCCGGAGCAACTCC
TACGTTCGTTAGACAATACAAAATCCCGCTCGCAGCATACGCAGCAGTACAAGAGA
TTATCGATTCCTTATTAGCTAAACGGATAATCAGGGAATGTAACAGTACGTACAGCG
CTCCCGTGTGGCCCGTTCTGAAACCAACGGGTAAATGGCGTCTTACCATTGACTATA
GACAACTCAACAAACTGGTTCCGTTGTCTCGTTGGCCAATGACACAGCTCGACCAAG
AGTTGCCAAAGGTGGCAAACGCCAAGTACTTCTCCACCGTCGACGTGGCAAATGGTT
TCTGGACCATGACAGTCGACCCGCGTGACCAACACAAGTTGGCTTTCTCTTTCTCGA
ACAAGCTCTTCACGTTCAATCGTTGCCCATTCGGGTATGCGAACTCCCCCTCAGAGT
TCAACATCTTCCTGCACAAGGCGATGCCAGACGCAGCCTCTAGGGGGACGATAATTT
ACGTGGACGACGTTCTCATGAGAAGTGAGACTTGGTCACATCACCTCAATGAAATG
GACCACGTTCTCACCCAACTCGGGACTGCAGGGGCTAAGTTGGCCATCATGAAAGG
ACAATGGTGCAGGACCAAGGTAAACTACGTTGGGCTTCTGGTGGGTGCCGAGGGCA
TCCTGCCACAGTCTAACCGAATCCAAGCAGTCCGCAACATCAAAACACCTACAAAC
CTTCACGAGGTGCGCAGCTTCTTGGGAGTATGTAATTACTCCCGTCAATTCATCGAA
AACTACGCCGACTTGTCAAAACCGTTGACTCACCTTCTTCAGAAAGACACCCCATTC
GTTTGGGATGTCACTCACAACGAAGCGGTGGCTTCCTTGAAAAACCTGCTTTGCGAG
GCACCCTGCCTGGTATACCCCAACAAAGACAAGACATTCTTTTTGGAAGTCGGTTTC
TCAGAGCACTGCCTTAGCGCTGGGTTGTACCAAAAATACGACCAAGACAAACGTGT
GGTTGCTTACGCGAGCAAAACACTCAACCCCGCCGAACACAAATACTCAGACTGCG
AAAAAGCGCTGCTGTCGACAGTCTGGGCGATCAAACACTTCACCAGTTATGTCGGCG
GACAAAAGGTGATCATAGAAACATGTCACCAGCCTGTGACTTTTCTGAAAAGCCAA
CGAATCCGTGAGGGTGCTGTACACAACAGCAGGATAGCGGCTTGGCTCATGACGCT
GCAGAGCCATGATGTAGTAATCAACTACGCAAAAGCAAAGAACCTGCCTTTGGGCA
GTGCTTTGGCGGTGTGTCAACACTGTGGTGACGATGAAACCGACTCCGGACCACCCC
CGCGTGACATCGCTCCGCCATTGCCTTCGAACCATCACTATTTCGAAGAGAACGTGT
GTCTCGACATGCCGATGGCATTTGTAGATGGCTGTTCTTACCGCCATCTAGACCACTT
GCAAGCTGGGGTGGGATTGGTATGGCACAACGACATTCCGTGCAAACCACTTCAATT
TCAGCTTGGCAACAAGACCAGCCAGTTTGCAGAAGTGGCTGGCGTGCTGATCACCCT
GCAAACAGCGGTGAAACACGAATTGGCAGAACTGGCGATCTGCACCGACTCTAACT
ACGCGAGACTCACCTTCTTATGCCACTTGCCGTTTTGGAAACAAAAGCACATGACTA
CTTCAAGTGGTAAAGAGGTGAAAAACAAAGAGCTCATTCTAGCTTGTGATGATCTCA
TCACCAAACACGACATACAGGTGTATTGGAAGAAAGTCAAGGGACACTCCAAATCA
CCTGGTCGTGACAAACTTGGCAACGACCATGCTGACAGCATGGCGAAATCTGGTGC

AATTCACGGCACCCCTTGGGTGTTTGCTGCTCGAGACGAAAAACCCACTGAGGAAG
CTATGGTGTCTGCGGTAACGCGTAGCCATGTAGCACCACGGAAAACGTCGTCGCAC
AAACATAATGTGTTGCTAACGCATTCATTCATGAATGGCGACCTGGTAGCAATGCAA
CACCAAGATGAGTCCCTCGCCACTTTGATGGCATTCCTGTCGGATCCTGTCAACCAC
CCAGTGTCCGAACTGGCTTTGGCAGGTTCACACGACTTGCGTTCGCTGTACGCAACT
AAACAGCACCTCACACTTGTAGATGACCTATTGGTGTATGTTTCAGAAACCGACACC
GCTCGAAGGTGGGTGGTTCCTAAAACACAGAGGGGGATAATGATAGCTCATGCCCA
CGACGAGCCATGTGGAGGCCATAGGGGGGTCAAGGCTACATGTGAAACATTGCGAC
AGGTGGCCCACTGGCCTCACATGGAACAAGACGTGGCACGATACGTGAGGGGGTGT
CTAGTTTGCTGCCAGTTTCAACCCACCAAGCCACTTCACAGAGCACCCCTGCAACGC
AAGGGTGTGTCTTACCCCTGGAGCTCGATCCAGATCGACTGGGTCGGCCCAGTTGCC
AGGTCTGCCAGAGGCAACAAGTACCTCCTCACAGTGACTTGCGCATTCACAAAGTG
GGTGGAGTGCCTGCCGGCGACCAATGACACAGCGGAAACCACTGCCGTTCTTTTGCT
AAACCACGTTTTCAGTCGTTTCGGCCTGCCAGGAGAAACCGTGGACAGCGACAGAG
GAACCCACTTTTCGGCAGCTGTAATGACCGAACTGTGGAAGCTCCTGGGAGTCAAA
GCTAAACTCCACATCGCGTACCACCCTAGGAGCTCGGGGGGGGTAGAAAGGAGCAA
CCAATCAATTGTCAGAATCTTGAGGAAATATGTGGCAGCCAACCACAAAGATTGGG
ACCTCAAACTCCCATTGGTACTGATGGCAATCAGAGCCACACGCAACAGGTCTACG
GGAATGACACCCTTTGAGATGATGACGGGCCGGCAAATGACCCTTCCACTGCATCTC
CTGTACCAACCGGGAGATGTCGCCGCAGCCACCGCCTACACGGCTCATCAATACGT
GACCGACTTGCGGAACCATCTCCAGACTACCTTTGCGTACGCTCAAGGACAATTGGA
AAGAAGTGCAGAAGGTGACAAGACCTATTACGACAAAAAAGCCTCACACCAGGTGT
TCGAGGTCGGAGATAAGGTGTGGTACTACATATACACCAAACCCGCGGGCGTCGCA
ACAAAGTTCCTGCCCCACTGGACGGGGCCACACGAGATTGTGGTGAAGCTATCACC
TGTAGCTTACCAGATCAAAATCAGCAAAGGTCGACAAAACGCCACATTAAAGTGGG
TCCACCGGAACCAAATCAAGTTGTACACTCCCCCCATGGGAATAGAAGGGGTGCTA
GCCAGCACCGAATAG
> Loki-Str consensus sequence Smc-Gag-Pol polyprotein
LTIIIMSNPSDGFPPISPLISGPSPTERLIADLSNDANPNYAEGLKMLDFNAISEKNADIATD
ALQNRPSGGGLVKVLSSLALNYAHQQRWTVHQYLSAQQRHEQEAGEFKVQVERTREL
ASKAEKDKLLLAEELCVRTDKLEDLSNTYNKEREQTLDQVRILKEQLGLAKTKYDEVH
AKLDESDELARKRGEQIDALQTVVNETTQSNNALFQKLQTKDDQLMNTMTKLEDKTAE
LIEVADLMKDEKDQVRKLENVTSKQKEDIASLDLSLHTRDLSLQTMTDKYEAERKKGD
TYVSKINTLNSQVDSAMQHNTTLRYHLEELQNSHALSRDNQTKQPSSHPELRERGNVD
DRRFQPLSLGHSAHSELGPPRQHNHSLTFPLGLSAHNSPRESADAPRALGGDRLDKIVKN
FRLFDPVPGKPNDTETFLADIEDALDGYPNATNADRLYLLKRTSNRHVTRFIRLQEQHV
QNDYAKLATVLKIEFSGSATRKHDSSLANNIRQARNEHPQAYYHRLRSAYFGMLTETG
MEDLLPFKQLFLSNMSPNFINYLGPTAHVGLPISTLRELASTAFEASKASRAKSPDTSTFE
LRREPSLELEGAASGTYAVKEDDQRGWLPSNHHPDNHRRNNSYDERPQSNRSRRDQPS
RHAPPPNSHKGSGHKGNKRNKGNKVFNNDLNANRNDIEALIKDVLHRKKFEKEDKDKS
SGLGVKSLETESAEIHAIQEDANISITPAPTRIPVQVPPVLDTSPQVLSTDLDTEVEMHEISS
FATDDSSTIPIEDPLGHVGNLSVLEIDADYKPLRSPNPLHFVGDMTRKTNSNRPYLRTVL
EDCVTCHALIDSGSTISLISETLLDELKRAVDPTKRWLKTEHCDTNLRGFTQATSPLTKRL
LLKLNFESVSLIHPVYVTSIEIERMLIGSDLIDRLVPLMDWKTNQIWSQIPMPTPLTSPHSS
KATCHTLCNDVGPTSASDANPVNLAMSPPFVAKDNVSSTRNLTEHVVFLCGLGDSPAD
TYRPPLIGGVCLNGTMAEDTRLATWSEKSAISLEMFNEISKTANCHPLVPKTFRFPLGTT

PQTTLTAIGVCALSIRIGTKELSHYLLVVADLPHTVYVGADILVRLGVKLDTIHQVLWSL
AQPNQHALSFDPVRMASGQTIPEACKTITESAMLIPARTTEVSVRLNLAPGYRMEGTTAF
FQPSPKLFDLGLTINGNPLLELTARSTYLLVQNLTQADISIPRHTQLGTLIDYAFHDFELV
VPVIGPLPSSLDLDGEGGTLFTCQSKAIALTPVLPLDDTSAFRLDVDPDSNLLIYSIVTEDD
IASPPACEVHSCEETTDDSSKGDMPSPPDEDLYNVAEPYPGFHAQVIQLLSEADALVNDT
ERHQLRELFNKHSEIWSTDSLDCGVTGIHVVRIPTPPGATPTFVRQYKIPLAAYAAVQEII
DSLLAKRIIRECNSTYSAPVWPVLKPTGKWRLTIDYRQLNKLVPLSRWPMTQLDQELPK
VANAKYFSTVDVANGFWTMTVDPRDQHKLAFSFSNKLFTFNRCPFGYANSPSEFNIFLH
KAMPDAASRGTIIYVDDVLMRSETWSHHLNEMDHVLTQLGTAGAKLAIMKGQWCRTK
VNYVGLLVGAEGILPQSNRIQAVRNIKTPTNLHEVRSFLGVCNYSRQFIENYADLSKPLT
HLLQKDTPFVWDVTHNEAVASLKNLLCEAPCLVYPNKDKTFFLEVGFSEHCLSAGLYQ
KYDQDKRVVAYASKTLNPAEHKYSDCEKALLSTVWAIKHFTSYVGGQKVIIETCHQPV
TFLKSQRIREGAVHNSRIAAWLMTLQSHDVVINYAKAKNLPLGSALAVCQHCGDDETD
SGPPPRDIAPPLPSNHHYFEENVCLDMPMAFVDGCSYRHLDHLQAGVGLVWHNDIPCK
PLQFQLGNKTSQFAEVAGVLITLQTAVKHELAELAICTDSNYARLTFLCHLPFWKQKHM
TTSSGKEVKNKELILACDDLITKHDIQVYWKKVKGHSKSPGRDKLGNDHADSMAKSGA
IHGTPWVFAARDEKPTEEAMVSAVTRSHVAPRKTSSHKHNVLLTHSFMNGDLVAMQH
QDESLATLMAFLSDPVNHPVSELALAGSHDLRSLYATKQHLTLVDDLLVYVSETDTARR
WVVPKTQRGIMIAHAHDEPCGGHRGVKATCETLRQVAHWPHMEQDVARYVRGCLVC
CQFQPTKPLHRAPLQRKGVSYPWSSIQIDWVGPVARSARGNKYLLTVTCAFTKWVECL
PATNDTAETTAVLLLNHVFSRFGLPGETVDSDRGTHFSAAVMTELWKLLGVKAKLHIA
YHPRSSGGVERSNQSIVRILRKYVAANHKDWDLKLPLVLMAIRATRNRSTGMTPFEMM
TGRQMTLPLHLLYQPGDVAAATAYTAHQYVTDLRNHLQTTFAYAQGQLERSAEGDKT
YYDKKASHQVFEVGDKVWYYIYTKPAGVATKFLPHWTGPHEIVVKLSPVAYQIKISKG
RQNATLKWVHRNQIKLYTPPMGIEGVLASTE*
> Loki-Str consensus sequence *env* gene
CTGATTATGCGTTTCTGTAGGATGGAGTTCCTTTGGATGATCCTGACACTCTGCGCCC
TGGTCAAAACCAACCCAGCAGACGTAATCACGAACGGACCCCCTACGGGAATCGTT
CTCCGCGACGATCCAGGACTCCTCATAACTAACTGCAGAGTACACACGCAGAGGGT
ATATGTCCGCCTAGATGCAGAGAATGTGTACCGCCAACACATTCCACCTGCGGCGCA
TTTGAGTTGGGCCGGGGCTGACTGGACCAGCCAGGCAATTAAGCACGCCCAGCTAG
ACACTGCCCATATGTTGGCCCAACTCCAAAAGTTCACAGTAACCCAGTCAGAACTAG
CTGAGCCCAGGCGAGAAAAACGATTCGTCGGGGCGCTACTATCGATAGGGGCAGCC
GTAGGGTCACTATTTGCCCTAGGTACCACTGCCGTCAACGCAGTCAGTCTAGCCACA
GTCAAGAGGAATGTTAGGGAGATTACTGAAGAGATGCCACTTATCCAGCAGCAGAT
GAAGGCACAGGCCCTTAGGTTGCAACATGTGGGAAAGTCTCTCCAGGACACTATTCT
GGTGGTCAACACACACGCTACTCTCCTGAACAAAACTATCCTGTCTGTAGGAAAGCT
AGCAGAGGTCATGAACCATGACTATGCCCATGTCCAATTGGTTCGATTGTTACTGGA
GGATTTTCTCCGTGAAGTTAGCTCTTCTATGGACCACTTGGCCATGAATAGAATCCC
CTCATATCTAGTGCCCCTCTCAATGGTGCACGACATATTGACCTCAGCTACTTCAAC
CACACTAAGGCCGTTACAATCACATTTGGCCTATAGTCTGGGGTCGGCCATCCCAAT
ACACGTTGATGTTGATCGTAATGAAGTGGGTTTTCTACTGACGCTGCCGGTTGTTGA
ACTGGAGAATATCTATAGATTGAAAACGGTTCTCAATGTAGGATTTTGGCGCGACAG
TACCCACGTAAAGATTGCCACGCCCCCAGTTGTAGCTTACCAGGAAAACAACCCAG
ATTTCTATCTCACCCCTAACCTGCTAATGTGTACTCTAACCAAAGATGTCCACTACCT
TTGTCCTAGCAAACCATTTGTCAGGGATAACACTGAGCGTCTTTGTGGTATTAAAGC

TATCACCAGCGAAGAACGCTGTAGAGCCAAACTAACAGCCAGGGATGGGGCCACCA
CCACACAAGTGGAGGTGGTAGGGAACCGATGGTTGGTCAACACACCGTTCGCGTCC
GCCACTATGACTTACGAACGCCATGACTCCATCACCCAATTGAACCTCCCCAACCAG
ACTGTTTTTGTTACGGTACCAGAGGGGGCGATTATTCACGTAGACGACCTCGCTCTA
TACCACCTTCCCGACGACCGGATAGACACAGAGATTGAAATGCCGGACGCTTTTGCT
AAACGTTCCCTTGAGTTACCACAAGCCATTCAATACCAAATCCAGTACGAGGGACCC
ATGACTGTCGATCTTAGCGTACTGGATGAGGCACTGTTAGTCGACCCGACTGACTAC
AGACCAAAAGATTGGTCTGTCGCCCGCTCTTGGACGGTGCCGGACAGTATCCTGACT
GTTACCATGATCCTTGGTCTCATTTCCCTTGGCGTGTGCACCTACTACGTACACAGGC
GCACACGTGCAATGGAAGACCTAATGAGGTCTTATACTAGGATCACCCGTGGGACG
GAAGCGATCCCGATTTTTGGAAAGGTGGCAATGGATCCGGAAACCCTCTTACAGGG
CCCAGTCCCAGCCTCCTCTCCCGAACTCGCTAGGTCAGCCCAGTAA
> Loki-Str consensus sequence Env protein
LIMRFCRMEFLWMILTLCALVKTNPADVITNGPPTGIVLRDDPGLLITNCRVHTQRVYVR
LDAENVYRQHIPPAAHLSWAGADWTSQAIKHAQLDTAHMLAQLQKFTVTQSELAEPRR
EKRFVGALLSIGAAVGSLFALGTTAVNAVSLATVKRNVREITEEMPLIQQQMKAQALRL
QHVGKSLQDTILVVNTHATLLNKTILSVGKLAEVMNHDYAHVQLVRLLLEDFLREVSSS
MDHLAMNRIPSYLVPLSMVHDILTSATSTTLRPLQSHLAYSLGSAIPIHVDVDRNEVGFL
LTLPVVELENIYRLKTVLNVGFWRDSTHVKIATPPVVAYQENNPDFYLTPNLLMCTLTK
DVHYLCPSKPFVRDNTERLCGIKAITSEERCRAKLTARDGATTTQVEVVGNRWLVNTPF
ASATMTYERHDSITQLNLPNQTVFVTVPEGAIIHVDDLALYHLPDDRIDTEIEMPDAFAK
RSLELPQAIQYQIQYEGPMTVDLSVLDEALLVDPTDYRPKDWSVARSWTVPDSILTVTM
ILGLISLGVCTYYVHRRTRAMEDLMRSYTRITRGTEAIPIFGKVAMDPETLLQGPVPASSP
ELARSAQ*
>Loki-Tfu consensus sequence
GGCAGAAAGCCCCAGACAACATAGCTCCTAGGATCATTCAGGCATGCAAGCCCCTC
CACCACAATAAGGTGGCGGTTCAAGGAGGGGCTATTAGCGTTGGCCACCAATATTG
TAAGAACTTGCATTGGCCCTTTGTCTGTGTCTCTTTTAGCCAACCGTGCCAGCATGTC
CCGACCATCCAGCCCCGAGGCACACTGGGATCAGCTAGAGACCTGAATGAACGGCA
TGGCGAGCACCTTCCTACCGAAACTACTGCTGATCTGCAACACTTGAGCCGGGAAA
AGCTTGATGACAACCTGAGTGTGGTGATGTCTCATAACGCAGCACAGGACTATAGCC
ACAAGGAGCTGGCCAAGATCTTCGGAGCACTGACCTACAACCTCATCGCCCAGACA
AAAATGGAGGTAGAAATAAACTCTTGCCTACAACAGGATTCGTTGTCATTACAGATC
CAACTTGGGGAGGCTCAGAGGAACCTAGCACAGGCTCACGACCAGCTAGACCAAGT
CCAGACAGGGACTCAGAGCCAACGAGAAACGGTAGATGAAGAGGACACAGAACTG
CAGGAGAAAGTAAAAGGACTACAACAGGCCATGACCGACCTCCGTGCAGACACAG
CACGCAGAGAACAACAGGAAAAGGACACCCAGGAGGAAAGACGAACTTGCACAAC
TCACAGAGAACTAAAACACGCATATAGATTACAAAGGGACCCTGAAAAGGAAAGG
CCCAACGTGGAATTTCCCTTAACCAGTGGACTCATACCCCCAAAACAAGAATCAAC
ATTTCAAAAGGGGGGTGAGAAATTCCGACTCGAGACCTCCCCAGGGTGGGAGACTT
TTCAAGAAACCCCGCGTTCCAGTCACAGGGTCGCCCCTAAGGAATTGGACAAACTG
GCTAGAAACATCCCTACTTTCACACCAGACCCTGCAGGAGGCTATGACATACATGCC
TACCTGCAAGACATTGACTTTCATTTACAGACTGTAGCTAACATTACTACACGAGAC
CAACTCTACCTCCTCAGGATCACCTCCAGCCGCGAAGTACGGAGCTTCCTGGATCGG
CAACCAGACAAGATCAAAGAGGATTACCAGCAGCTTCAAGAAGCCCTGATTAAGGA
ATTCTCTGACCCAGAGTCAGAGCAAGGACTGTTCACTGCCCTGGACCTAAAACAGA

GCAGACATGAAACCACACAAGCTTATTACAACAGGCTCCGACAGGCCTACTTCGGA
GCCCGAAATGAACCAGGAATGGAGGAGGACTTCAACTTCAAAACTTTGTTCCTCAG
AAACCTTCACCCTGCTGTAAGTCACCACCTGGGGGTCCTGGCCTGCCCACGGGAAAT
GTCCATCCATCAGCTTCGAGACTTAGCCCGAAAAGCCTATGCCAAGCAAAGGACTG
TGCCCGAAAAACCTGTAAAAGGGGCAGCAATCTATTCTGTCTCCGAGCAATGTTCAG
AACTAGCACTAGAGGGCACCCAGCAACGCTACCATGACAAACCTGTCCAGCAGGTG
CCACGGCGCTTTCCAATCAACAGAGAACCACAGGGTCACGGCGGCACCCGTCCCAA
ATACCAAGCTGGATGCCCTGGGAGATCCTGNGACAGGCCACGACCACCACGCAACC
CAAAAGACGGAGCCACCTGGGAAGGATCTAACTGGCCCAAAGAGAACTGGTACTCG
CCAACACAAACATCAGGTTTCGACTGGAAGCAACAAAAACCCTCGCAACATGCACC
AGATAAACCAAAACACAGACCACCAATAACACAGGACCCAGAAAACACAGAGATC
TTGAGAATTCTAAAGGAAATCATCCAAGACAAACGCCATAAAAAGGACAAGGAGG
ACGAATCAGATAAATTATGNCTAAGCCAGGAACCTGACCGAACACCCACAATATGT
CCTCACAAACCAAACCATCAGAGTCCGCCGACCTCCAATACCAAATCTCCAGCCAG
AACCCCAATATTCCAGAGAGTGCTGTTTTAATGGTCTGCCTCCCCAATGAGCCACAG
GAATCCAGTCCCGAACTGCATGCCAGTCACTACACCTGCCCCAGGCCCAAGACTCCT
AGGGAATCTGTTTGAAAGAGGAATGGCCAAAAAACTGTACCTAGCCATCACATTAG
AGAATGAGCTGAGAGTAGAAGCTCTTGTGGACACTGGTGCAGATCTGACATTAATG
TCGACCCAACTCTTTTACAAACTACAGGTGGAAGCTAAGCGACGGAATCGAACCCTT
AAAACGCAAAAGTGTACATTGAATGTACAATCTTACAGCCAAAACGAAGTCAAGCT
TGAACAGATAGCTCCCATTCACCTGACCATCAGGCCCATGAGTGTAGTGCACCCCGT
GTACATCTCCCCAATGGACACCTACCCCCTGCTCATCGGAAAAGACCTGCTAGACTG
GTTTGAACCCCTGCTGGATTTTAAACGGCTGAAACTCTGGGGCTCAAGTGCGAAAAA
CTCTACCACCTCGGCCACCGAGAGCATCCAAGACAGACTGCCAGGTCACAGGGGTC
ATGGACAATCACACCACCAGCCGGGAAGGCAATTTGACCCGAGAACAAAGTTCAAG
GTCGCTGCTCTGTGCGTTCCAGCTAACAAGGACTCTGACTCATACTGTCCGCAAGCC
ATGGGAGGCCTACAGCTTGATGATATTACAACGACGGACATTATACTAGCACTATGG
GCAGACAACTCAGCAATCAGTCTCACACTGTTCAACACCCTCAAAATAGACACAGC
AAGCATCCCTTGGACCAAGAAGCGCACGCGATTCCCGCTGGTTCCCTGCCCGACCAC
CATGGTGAGAGCTAATGGCATTTGTGCCCTGAATGTAAAGTGGAATGGCAGGTGCA
TGACCCATTACTTCCTGATCCTGCCAGACCCACCACATGACGTCTACATTGGAGCTG
ATATTTTAATCCGTCTCAGGGCCCATGTGGATATGATCACTGAAGTAGTGTGGGCCC
CAGTTACACAACAGGCCCCCACAGCCGAACAAGAGAACAAGAGACTGTANATCAGC
GAACAAGAGACTGTAATACCTGCTTACACCAAGGGCGTTGCGGTTCGACTCAACCTG
AAGTGCGGTCAAGCCTTAAACCAGACGCTAGGCTTCTACCAACCGTCACCTGAGAG
TAAGGAACTGGGACTCACCTTGGAAGCTACACCCCTAGTTGAAATAACGTCCCGAG
CAGTTCACATCTTATTCAATAACTGCACAGCAAAGGACATCCGGGTACCTAAAGCAT
ACCGCCTGGGCTGGTTAGTGAGCTATGGCTTCCACGACTTCGAACTCACCTTGCCAA
TAATCGGCCCCATCCCACATGAAATGATCCCAGAAGAAAGTACAGATCAAGCCGTG
GAGACAGCACCTTTCAAGATGATCACCATTATCTCAGTCATGCCGGCCGTCCAAGAT
CAAGTGTGCAGAACAGAGCTAACCGATGATCACTACCTCTCTGTGTACACGGTCTCC
ACGAGCCCCACCTGCGAGACGGACAACACCACTGCACAGCTAGTAGCACAGGAAAC
ACACGGATTAACAACAGAAGAACCGTACCCGGGATTTGAGTCACAGGTGCAACAAA
TCCTGAAGGATGCCGACGCCATTCAAGATGAGGGAGACCGGCACAAACTACGTCAA
GTCCTGTATAAATATAAAACATCATTTGCAAAGGACTCTTTAGACTGTGGCCTAACC
AACCTGCATGTGATGCGTATCCCCACGCACCCACAGGCTCCGCCCACCTTTGTCAAA

CAATACAAGATCCCCCTTGAATCATACGAACCTGTGCAGGAGATCATCAACTCCATG
CTTGAAAAGGGTGTCATCCGCTTATGCAACAGCACCTGCTCAGCCCCCATATGGCCT
GTCCTAACACCAAACGGCAAGTGGCGACCGACCATCGACTATAGGAAACTAAACCA
ACAGGTATCATTGTCACGTTGGCCCATGACCCAGCTGGAACAAGAAATACCAAGAG
TTAAGGGAGCCACAATTTTCTCCACACTTGACGTGGCCTCAGGATTTTGGACTATAC
CAGTGCACCCCGATGATCAACACAAGCTGGCCTTCATGTTCTGCAACCGGTANTATA
CCTTCACAAGGTGTCCATTCGGCTATGCTAACTCGCCAGCTGAATTTAACATCTTCCT
CAACAAGGCATGCCCAGATGCTAGAGTAAGGGGCAACCTCATCTACGTGGACGATG
TCCTCATGAAAAGCACAACCATGACTGATCATCTAAAAGAAATCGACCACGTCCTA
CAACAACTAACCAAGGCTGGGGCCAAAATCGCACTCCACAAGGGACAATGGTGCAA
AACCAAAGTCAATTATGTGGGTCCGCTCATTGCACCACAGGGCATCGAACGACAAA
CCAACCGCATCCAGGCGATTCAAAACATCAAGCCACCTGCCAACATCTCAGAACTC
CGAAGTTTCCTAGGTGTATGCAGCTATTCTCGACAGTTCATTAAGGACTACTCAGAT
ATTGCACGCCCTCTCATGGCACTTCTAAAGAAAGATTCCCCCTTCACCTGGACAGAC
GCCCAGCAAAATGCTATGAACAAACTGAAGAAACACCTGTGCACCGCTCCATGCCT
GGCANAACCCCACCCACAGAGGGAGTTCTACCTGGAGGCTGGATTCTCTAGTCACT
GCCTGAGTGCTGGACTTTATCAGCGACATGACCAGGACAAAAGGGTGGTCGCATAC
GCCAGCAAGACACTCTTGCCACCAGAGTGCAAATACTCAGACTGTGAGAAAGCCCT
GCTCTGCACTGTATGGGCCATTCAGAGATTCTCTAACTATATCGGTGCCCAGAAAGT
GATCATAGAGACCTGCCACCAGCCTACCACATTTCTTAACAGCCAATGCATCCGTGA
TGGTGTGGTCACCAATGCCCGCCTCGCCACATGGCTAATGGCCCTCCAAGGGCGCAA
CGTGGAAACACGGTATGCCCAAAACCACAAGTCAGCACTGGGCAATGGCCTGGCAG
CATGCCAAAGCTGCTCAGACGATACACCGACTGTGACTACAGATATGGAAAGGCCC
CAACTGCAGCCACCCACCTGTCACAAGTACTTTGAAGTCAACGTGNTATGCCAAGGC
ATGCCCACAGCCTACGTTGATGGATGCTCCTACAACCATGAGGGTAACCTTAAAGCG
GGTGCAGGTGTCCTCTGGATGNAACCATGCCAACCACAACACTTCAGGCTGGGGCC
CCAGTCCTCACAATATGCGGAGATAGCAGCTGTACTCATAGCCCTACAGATAGCATC
AGCCCACAACATCAGAGACCTGCTCATTTGCACAGACTCCAACTATGCCAGACTTAG
TTTCACATGCCACCTCCCAAGCTGGAAGAAAAACGGCTTCAGGACTGCAAACAACA
AACAAGTTAAACATAAACACCTGTTCCTGGCATGTGAGGACATCACAAACGAACAC
GATATGACCATCTACTGTAAGAAGGTAAAAGGCCACTCGAAGCAACCAGGCACAAA
TAAAGATCTAAATGACCAGACAGACGCCCTCGCCAAAGCAGGCGCTTTACACGGCA
ATTTGTGGTCACCACCAACCGGACCACACAAACCCACCATCACAGCCATAACCCGC
AGCTACCGAACTAAATCAACCACACTTCCCACCTCACAGCCACTTACGCTGACTCCA
CAAATCTCCAAAAATGACATCGCTGACATGCAGGCATCTGATTCCACTATACAGACC
ATCCTTCGCCACCTCTCTAGCCCCTCCGACCCCCCGGTCACAGCTGTTAATCTAGCCA
CCTCCCCATCCCTAAAACGGCTACATGATGTCAAGCCCATGCTGCGTGTGATAGACG
GCATTTTATCGTATGCCCCTGATGACCGCACCGTGCCAAAACTTGTGGTGCCACAGT
GCCAAAGGGGGGTCATGTTAATGTATGCCCACGACGCACCATGTGCCGGACACCAT
GGCACAAAAGCCACGTATGAGACATTGAAACAAGTGGTATACTGGCCTGGCATGCA
ACAAGATGTGACAGAGTATGTCAAAGGATGCCTTAGGTGCTGTCAATTCCAACCAG
CAAACCCAAACTACCGAGCCCCACTGCAACGCAGAGGAGTCACATTTCCGTGGTCA
GACCTCCAAATTGACTGTGTGGGACCCCTTCCCAGATCAACAAGAGGCAACAAATA
CTTCATCACAGTCGTGTGTCAGTTCACCAAGTGGGTGGAGTGCCTCCCAGCACCCAA
TGACACTGCTCACACCACCGCCTATCTTCTGATGAATCATATTTTCAACCGGTTTGGG
CTACCACTGAGAGTCAATTCCGACAGAGGCACCCATTTCACGGCAGAGGTCATGCA

GCAGATTTGGAAGCTACTGGGGGTGAAAGCAAAACTTCACATCAGTCACCACCCTA
TCTCCTCAGGTCAAGTAGAACGAACCAATCGGACGGTAGTCAGCATGTTAAGGAAA
TACGTGGCTGCCAATCAANAAGACTGGGATGTGAAACTTCCTCTGGTGCTCATGGCC
ATCAAGGCAACACCACAGGAGTCAACCGGCATCTCCCCTTTCGAACTAATGACTGG
ACGGCAGATGACATTACCACTACACCTGCTATACCAACCAGGAGACTCTAACCTTGT
CACAGCTTATACCACCTGCCAATACCTCGAGGAACTGCACAAGCACCTGAGATCCA
CCTTTTCATTTGCTCAACAACACCTCCAAAAGAGTGCCGAGGGTCAGAAAGCATATT
ATGACCGTAAAGCATCGCATGAGGAACTGAACGTGGGGGATAAGGTCTGGTACTAC
TGCTTCAACCAGCCACCTCAAACGGCCTCTAATCGGCGATCAAAGAAATTCCACATT
GGGCAGGACCGCACGAGATTGTAGACAAACTCTCCTCGGTTGCCTACCGAATCCGA
ATAAGCCAGGGACAGAAAGAACCGGTCCTCAAATGGGTCCATCGCAACCAGATCAA
GAGACACAAGGCCATAAACAAGGAAAAACTGTCTGGGGCTATTGCAAACTCGGACT
AACTACTAATTCCACAATGGAAAACATGTCTTGACCAATCAAGAATTTCCCCCACTG
ACCTACTGTTTTTCTCCCCCTCTGTATCAGGATGCTGCTGTGGCTCTCCCTTCTGTGCC
TCCAAATAACCAGAGGACAGCCGCAACCTGACCTCACCTCTCCGGGCCCGACATCA
GGCATCGTATTGAAGGATCAACCGGGACTTATTACCAACTGCAGAACTCATGCCCAA
AAGGTATATGTGCGACTCGACCCCCGAAATGTCTACCAAACGCAGTACCCACAGAC
GGCCGCACACTACAGTTGGGCAAGGGCCCGATGGACCGAGAATGCACTTGCACATG
CCGAGGCTGATATCAAACACATGCTCACTCAGTTATANAAGATGACTGTCACTCAGG
CAGAGCTAAATGGGCACAACCGAAGGAACAAACGGTTCCTGGGGGCCTTGCTGGGA
ACCGCGGCTGCGGTGGGCACTCTTTTCAACCTAGGTCTGTCCAGTGTAAACTCTGTA
AGCCTGTCGACTCTGAGACAGCATATGAGAGAAATCCAGACTGAGATACCCCAGTT
GAGAGAACAACTCACACAACAAGGGCAAACCCTCCAGACCATAGGCAAGTCATTAA
AAGGCACCTTAGTGGTCCTCAACACCCACAGCATCCTGATGAACCAAACCGTGAAC
TCCATAAAACAGCTGTACTCGATCCTACAAAATGATGTGGCCCAAATCCAGCTGGTC
ACAGCACTGATGGCCGACATGCTGCGAGAAGTGAGTTCTTCCATCGACAGCCTAGC
CATGGGCAGGATCCCTCCATACCTGGTGCCCTTAAGCCTTGCACAAACTGCCCTGTC
CTCCGCTACCACTCATTCAACTGATTCCCTTCAAGCCCATCTTGCATACGCACTGGG
AGGATCCGTCCTACTTCATCGACCCAGAACAGGGTGGATGGCCTTCCTCCTAAACCT
ACCAATCATTGAATCTGACAACATCTATCGGCTCAAAGATGTGGTCAATGTCGGTTT
TTGGAAAGGAAACACCCACATCAAGGTGCACACCCCTGAGGTGGTAGCCTACCACG
ACAGCAATCCACAGTTATATCTGGCCCCAAACCTGCGCATGTGTACCCGCACCAAGG
ACATTCACTATCTCTGCCCAAGCAAGCCGTTCCGCCGAGACAACACCAAAGGAATCT
GTGGGTTGCAACCCATGCAACAGGACACCCGCTGCCCTGCTGAAGCCAAGCCATGC
AGCCAGGACACGCAAGCAGAAATCGTTGGAGATCGGTGGCTGGTCAACACTCCCAT
CCGTACCGCTACATTAACCTACGACAAACATGATACTGCCACTCGCGTGAACTTACC
CAGCCAAAATATGTGGATTCAAGTCCCCAAGGGTGCCATCCTCCACCTCGGGGACCT
CGCTCTATTCCACCTCCCCAGTGAAGAGTATCAATCAGAGGTGGAAATCACGGACTT
CTTCAAAATTCACAATCTCACCCTGGACCCTGAATTGGAACTGAGAATCGAGGAGG
GGGGGAGGCAAATGATCGACATCACCCCAGTCGACACGGCCCTCCAAGCATTGTCC
CGACTACCAATCCCAGCTAACTTCCCCGTTGCCCGAGCCTGGACTGCCGCTGACACA
GCACTCTGCATCTCCACAGTCATCGGGTATGCACTGAGCCTCGGCCTCGCCCTACTT
CTGTTCAAACGAATCAATGGAATACAAAGATCCATGACCAGGTGTACAGCTGCTCTT
CCCCGTGCTTTCAACCGTAACCACCGGAGAAGCGAGACCCAACCAGAATCAACGAT
AAGTCTCACCAAACTCAACCCCATGACTGAGGAACCCAGGGCTGAGGGAAGCTAA
> Loki-Tfu consensus sequence *smc-gag* gene

TTGCATTGGCCCTTTGTCTGTGTCTCTTTTAGCCAACCGTGCCAGCATGTCCCGACCA
TCCAGCCCCGAGGCACACTGGGATCAGCTAGAGACCTGAATGAACGGCATGGCGAG
CACCTTCCTACCGAAACTACTGCTGATCTGCAACACTTGAGCCGGGAAAAGCTTGAT
GACAACCTGAGTGTGGTGATGTCTCATAACGCAGCACAGGACTATAGCCACAAGGA
GCTGGCCAAGATCTTCGGAGCACTGACCTACAACCTCATCGCCCAGACAAAAATGG
AGGTAGAAATAAACTCTTGCCTACAACAGGATTCGTTGTCATTACAGATCCAACTTG
GGGAGGCTCAGAGGAACCTAGCACAGGCTCACGACCAGCTAGACCAAGTCCAGACA
GGGACTCAGAGCCAACGAGAAACGGTAGATGAAGAGGACACAGAACTGCAGGAGA
AAGTAAAAGGACTACAACAGGCCATGACCGACCTCCGTGCAGACACAGCACGCAGA
GAACAACAGGAAAAGGACACCCAGGAGGAAAGACGAACTTGCACAACTCACAGAG
AACTAAAACACGCATATAGATTACAAAGGGACCCTGAAAAGGAAAGGCCCAACGTG
GAATTTCCCTTAACCAGTGGACTCATACCCCCAAAACAAGAATCAACATTTCAAAAG
GGGGTGAGAAATTCCGACTCGAGACCTCCCCAGGGTGGGAGACTTTTCAAGAAAC
CCCGCGTTCCAGTCACAGGGTCGCCCCTAAGGAATTGGACAAACTGGCTAGAAACA
TCCCTACTTTCACACCAGACCCTGCAGGAGGCTATGACATACATGCCTACCTGCAAG
ACATTGACTTTCATTTACAGACTGTAGCTAACATTACTACACGAGACCAACTCTACC
TCCTCAGGATCACCTCCAGCCGCGAAGTACGGAGCTTCCTGGATCGGCAACCAGAC
AAGATCAAAGAGGATTACCAGCAGCTTCAAGAAGCCCTGATTAAGGAATTCTCTGA
CCCAGAGTCAGAGCAAGGACTGTTCACTGCCCTGGACCTAAAACAGAGCAGACATG
AAACCACACAAGCTTATTACAACAGGCTCCGACAGGCCTACTTCGGAGCCCGAAAT
GAACCAGGAATGGAGGAGGACTTCAACTTCAAAACTTTGTTCCTCAGAAACCTTCAC
CCTGCTGTAAGTCACCACCTGGGGGTCCTGGCCTGCCCACGGGAAATGTCCATCCAT
CAGCTTCGAGACTTAGCCCGAAAAGCCTATGCCAAGCAAAGGACTGTGCCCGAAAA
ACCTGTAAAAGGGGCAGCAATCTATTCTGTCTCCGAGCAATGTTCAGAACTAGCACT
AGAGGGCACCCAGCAACGCTACCATGACAAACCTGTCCAGCAGGTGCCACGGCGCT
TTCCAATCAACAGAGAACCACAGGGTCACGGCGGCACCCGTCCCAAATACCAAGCT
GGATGCCCTGGGAGATCCTGNGACAGGCCACGACCACCACGCAACCCAAAAGACGG
AGCCACCTGGGAAGGATCTAACTGGCCCAAAGAGAACTGGTACTCGCCAACACAAA
CATCAGGTTTCGACTGGAAGCAACAAAAACCCTCGCAACATGCACCAGATAAACCA
AAACACAGACCACCAATAACACAGGACCCAGAAAACACAGAGATCTTGAGAATTCT
AAAGGAAATCATCCAAGACAAACGCCATAAAAAGGACAAGGAGGACGAATCAGAT
AAATTATGNCTAAGCCAGGAACCTGACCGAACACCCACAATATGTCCTCACAAACC
AAACCATCAGAGTCCGCCGACCTCCAATACCAAATCTCCAGCCAGAACCCCAATATT
CCAGAGAGTGCTGTTTTAA
> Loki-Tfu consensus sequence Smc-Gag polyprotein
LHWPFVCVSFSQPCQHVPTIQPRGTLGSARDLNERHGEHLPTETTADLQHLSREKLDDN
LSVVMSHNAAQDYSHKELAKIFGALTYNLIAQTKMEVEINSCLQQDSLSLQIQLGEAQR
NLAQAHDQLDQVQTGTQSQRETVDEEDTELQEKVKGLQQAMTDLRADTARREQQEKD
TQEERRTCTTHRELKHAYRLQRDPEKERPNVEFPLTSGLIPPKQESTFQKGGEKFRLETSP
GWETFQETPRSSHRVAPKELDKLARNIPTFTPDPAGGYDIHAYLQDIDFHLQTVANITTR
DQLYLLRITSSREVRSFLDRQPDKIKEDYQQLQEALIKEFSDPESEQGLFTALDLKQSRHE
TTQAYYNRLRQAYFGARNEPGMEEDFNFKTLFLRNLHPAVSHHLGVLACPREMSIHQL
RDLARKAYAKQRTVPEKPVKGAAIYSVSEQCSELALEGTQQRYHDKPVQQVPRRFPINR
EPQGHGGTRPKYQAGCPGRSXDRPRPPRNPKDGATWEGSNWPKENWYSPTQTSGFDW
KQQKPSQHAPDKPKHRPPITQDPENTEILRILKEIIQDKRHKKDKEDESDKLXLSQEPDRT
PTICPHKPNHQSPPTSNTKSPARTPIFQRVLF*

> Loki-Tfu consensus sequence *pol* gene
ATGAGCCACAGGAATCCAGTCCCGAACTGCATGCCAGTCACTACACCTGCCCCAGG
CCCAAGACTCCTAGGGAATCTGTTTGAAAGAGGAATGGCCAAAAAACTGTACCTAG
CCATCACATTAGAGAATGAGCTGAGAGTAGAAGCTCTTGTGGACACTGGTGCAGAT
CTGACATTAATGTCGACCCAACTCTTTTACAAACTACAGGTGGAAGCTAAGCGACGG
AATCGAACCCTTAAAACGCAAAAGTGTACATTGAATGTACAATCTTACAGCCAAAA
CGAAGTCAAGCTTGAACAGATAGCTCCCATTCACCTGACCATCAGGCCCATGAGTGT
AGTGCACCCCGTGTACATCTCCCCAATGGACACCTACCCCCTGCTCATCGGAAAAGA
CCTGCTAGACTGGTTTGAACCCCTGCTGGATTTTAAACGGCTGAAACTCTGGGGCTC
AAGTGCGAAAAACTCTACCACCTCGGCCACCGAGAGCATCCAAGACAGACTGCCAG
GTCACAGGGGTCATGGACAATCACACCACCAGCCGGGAAGGCAATTTGACCCGAGA
ACAAAGTTCAAGGTCGCTGCTCTGTGCGTTCCAGCTAACAAGGACTCTGACTCATAC
TGTCCGCAAGCCATGGGAGGCCTACAGCTTGATGATATTACAACGACGGACATTATA
CTAGCACTATGGGCAGACAACTCAGCAATCAGTCTCACACTGTTCAACACCCTCAAA
ATAGACACAGCAAGCATCCCTTGGACCAAGAAGCGCACGCGATTCCCGCTGGTTCC
CTGCCCGACCACCATGGTGAGAGCTAATGGCATTTGTGCCCTGAATGTAAAGTGGAA
TGGCAGGTGCATGACCCATTACTTCCTGATCCTGCCAGACCCACCACATGACGTCTA
CATTGGAGCTGATATTTTAATCCGTCTCAGGGCCCATGTGGATATGATCACTGAAGT
AGTGTGGGCCCCAGTTACACAACAGGCCCCCACAGCCGAACAAGAGAACAAGAGA
CTGTANATCAGCGAACAAGAGACTGTAATACCTGCTTACACCAAGGGCGTTGCGGTT
CGACTCAACCTGAAGTGCGGTCAAGCCTTAAACCAGACGCTAGGCTTCTACCAACC
GTCACCTGAGAGTAAGGAACTGGGACTCACCTTGGAAGCTACACCCCTAGTTGAAA
TAACGTCCCGAGCAGTTCACATCTTATTCAATAACTGCACAGCAAAGGACATCCGGG
TACCTAAAGCATACCGCCTGGGCTGGTTAGTGAGCTATGGCTTCCACGACTTCGAAC
TCACCTTGCCAATAATCGGCCCCATCCCACATGAAATGATCCCAGAAGAAAGTACA
GATCAAGCCGTGGAGACAGCACCTTTCAAGATGATCACCATTATCTCAGTCATGCCG
GCCGTCCAAGATCAAGTGTGCAGAACAGAGCTAACCGATGATCACTACCTCTCTGTG
TACACGGTCTCCACGAGCCCCACCTGCGAGACGGACAACACCACTGCACAGCTAGT
AGCACAGGAAACACACGGATTAACAACAGAAGAACCGTACCCGGGATTTGAGTCAC
AGGTGCAACAAATCCTGAAGGATGCCGACGCCATTCAAGATGAGGGAGACCGGCAC
AAACTACGTCAAGTCCTGTATAAATATAAAACATCATTTGCAAAGGACTCTTTAGAC
TGTGGCCTAACCAACCTGCATGTGATGCGTATCCCCACGCACCCACAGGCTCCGCCC
ACCTTTGTCAAACAATACAAGATCCCCCTTGAATCATACGAACCTGTGCAGGAGATC
ATCAACTCCATGCTTGAAAAGGGTGTCATCCGCTTATGCAACAGCACCTGCTCAGCC
CCCATATGGCCTGTCCTAACACCAAACGGCAAGTGGCGACCGACCATCGACTATAG
GAAACTAAACCAACAGGTATCATTGTCACGTTGGCCCATGACCCAGCTGGAACAAG
AAATACCAAGAGTTAAGGGAGCCACAATTTTCTCCACACTTGACGTGGCCTCAGGAT
TTTGGACTATACCAGTGCACCCCGATGATCAACACAAGCTGGCCTTCATGTTCTGCA
ACCGGTANTATACCTTCACAAGGTGTCCATTCGGCTATGCTAACTCGCCAGCTGAAT
TTAACATCTTCCTCAACAAGGCATGCCCAGATGCTAGAGTAAGGGGCAACCTCATCT
ACGTGGACGATGTCCTCATGAAAAGCACAACCATGACTGATCATCTAAAAGAAATC
GACCACGTCCTACAACAACTAACCAAGGCTGGGGCCAAAATCGCACTCCACAAGGG
ACAATGGTGCAAAACCAAAGTCAATTATGTGGGTCCGCTCATTGCACCACAGGGCA
TCGAACGACAAACCAACCGCATCCAGGCGATTCAAAACATCAAGCCACCTGCCAAC
ATCTCAGAACTCCGAAGTTTCCTAGGTGTATGCAGCTATTCTCGACAGTTCATTAAG
GACTACTCAGATATTGCACGCCCTCTCATGGCACTTCTAAAGAAAGATTCCCCCTTC

ACCTGGACAGACGCCCAGCAAAATGCTATGAACAAACTGAAGAAACACCTGTGCAC
CGCTCCATGCCTGGCANAACCCCACCCACAGAGGGAGTTCTACCTGGAGGCTGGAT
TCTCTAGTCACTGCCTGAGTGCTGGACTTTATCAGCGACATGACCAGGACAAAAGGG
TGGTCGCATACGCCAGCAAGACACTCTTGCCACCAGAGTGCAAATACTCAGACTGT
GAGAAAGCCCTGCTCTGCACTGTATGGGCCATTCAGAGATTCTCTAACTATATCGGT
GCCCAGAAAGTGATCATAGAGACCTGCCACCAGCCTACCACATTTCTTAACAGCCA
ATGCATCCGTGATGGTGTGGTCACCAATGCCCGCCTCGCCACATGGCTAATGGCCCT
CCAAGGGCGCAACGTGGAAACACGGTATGCCCAAAACCACAAGTCAGCACTGGGCA
ATGGCCTGGCAGCATGCCAAAGCTGCTCAGACGATACACCGACTGTGACTACAGAT
ATGGAAAGGCCCCAACTGCAGCCACCCACCTGTCACAAGTACTTTGAAGTCAACGT
GNTATGCCAAGGCATGCCCACAGCCTACGTTGATGGATGCTCCTACAACCATGAGG
GTAACCTTAAAGCGGGTGCAGGTGTCCTCTGGATGNAACCATGCCAACCACAACAC
TTCAGGCTGGGGCCCCAGTCCTCACAATATGCGGAGATAGCAGCTGTACTCATAGCC
CTACAGATAGCATCAGCCCACAACATCAGAGACCTGCTCATTTGCACAGACTCCAAC
TATGCCAGACTTAGTTTCACATGCCACCTCCCAAGCTGGAAGAAAAACGGCTTCAGG
ACTGCAAACAACAAACAAGTTAAACATAAACACCTGTTCCTGGCATGTGAGGACAT
CACAAACGAACACGATATGACCATCTACTGTAAGAAGGTAAAAGGCCACTCGAAGC
AACCAGGCACAAATAAAGATCTAAATGACCAGACAGACGCCCTCGCCAAAGCAGGC
GCTTTACACGGCAATTTGTGGTCACCACCAACCGGACCACACAAACCCACCATCACA
GCCATAACCCGCAGCTACCGAACTAAATCAACCACACTTCCCACCTCACAGCCACTT
ACGCTGACTCCACAAATCTCCAAAAATGACATCGCTGACATGCAGGCATCTGATTCC
ACTATACAGACCATCCTTCGCCACCTCTCTAGCCCCTCCGACCCCCCGGTCACAGCT
GTTAATCTAGCCACCTCCCCATCCCTAAAACGGCTACATGATGTCAAGCCCATGCTG
CGTGTGATAGACGGCATTTTATCGTATGCCCCTGATGACCGCACCGTGCCAAAACTT
GTGGTGCCACAGTGCCAAAGGGGGGTCATGTTAATGTATGCCCACGACGCACCATG
TGCCGGACACCATGGCACAAAAGCCACGTATGAGACATTGAAACAAGTGGTATACT
GGCCTGGCATGCAACAAGATGTGACAGAGTATGTCAAAGGATGCCTTAGGTGCTGT
CAATTCCAACCAGCAAACCCAAACTACCGAGCCCCACTGCAACGCAGAGGAGTCAC
ATTTCCGTGGTCAGACCTCCAAATTGACTGTGTGGGACCCCTTCCCAGATCAACAAG
AGGCAACAAATACTTCATCACAGTCGTGTGTCAGTTCACCAAGTGGGTGGAGTGCCT
CCCAGCACCCAATGACACTGCTCACACCACCGCCTATCTTCTGATGAATCATATTTT
CAACCGGTTTGGGCTACCACTGAGAGTCAATTCCGACAGAGGCACCCATTTCACGGC
AGAGGTCATGCAGCAGATTTGGAAGCTACTGGGGGTGAAAGCAAAACTTCACATCA
GTCACCACCCTATCTCCTCAGGTCAAGTAGAACGAACCAATCGGACGGTAGTCAGC
ATGTTAAGGAAATACGTGGCTGCCAATCAANAAGACTGGGATGTGAAACTTCCTCT
GGTGCTCATGGCCATCAAGGCAACACCACAGGAGTCAACCGGCATCTCCCCTTTCGA
ACTAATGACTGGACGGCAGATGACATTACCACTACACCTGCTATACCAACCAGGAG
ACTCTAACCTTGTCACAGCTTATACCACCTGCCAATACCTCGAGGAACTGCACAAGC
ACCTGAGATCCACCTTTTCATTTGCTCAACAACACCTCCAAAAGAGTGCCGAGGGTC
AGAAAGCATATTATGACCGTAAAGCATCGCATGAGGAACTGAACGTGGGGGATAAG
GTCTGGTACTACTGCTTCAACCAGCCACCTCAAACGGCCTCTAATCGGCGATCAAAG
AAATTCCACATTGGGCAGGACCGCACGAGATTGTAG
> Loki-Tfu consensus sequence Pol protein
MSHRNPVPNCMPVTTPAPGPRLLGNLFERGMAKKLYLAITLENELRVEALVDTGADLT
LMSTQLFYKLQVEAKRRNRTLKTQKCTLNVQSYSQNEVKLEQIAPIHLTIRPMSVVHPV
YISPMDTYPLLIGKDLLDWFEPLLDFKRLKLWGSSAKNSTTSATESIQDRLPGHRGHGQS

HHQPGRQFDPRTKFKVAALCVPANKDSDSYCPQAMGGLQLDDITTTDIILALWADNSAI
SLTLFNTLKIDTASIPWTKKRTRFPLVPCPTTMVRANGICALNVKWNGRCMTHYFLILPD
PPHDVYIGADILIRLRAHVDMITEVVWAPVTQQAPTAEQENKRLXISEQETVIPAYTKGV
AVRLNLKCGQALNQTLGFYQPSPESKELGLTLEATPLVEITSRAVHILFNNCTAKDIRVP
KAYRLGWLVSYGFHDFELTLPIIGPIPHEMIPEESTDQAVETAPFKMITIISVMPAVQDQV
CRTELTDDHYLSVYTVSTSPTCETDNTTAQLVAQETHGLTTEEPYPGFESQVQQILKDAD
AIQDEGDRHKLRQVLYKYKTSFAKDSLDCGLTNLHVMRIPTHPQAPPTFVKQYKIPLES
YEPVQEIINSMLEKGVIRLCNSTCSAPIWPVLTPNGKWRPTIDYRKLNQQVSLSRWPMTQ
LEQEIPRVKGATIFSTLDVASGFWTIPVHPDDQHKLAFMFCNRXYTFTRCPFGYANSPAE
FNIFLNKACPDARVRGNLIYVDDVLMKSTTMTDHLKEIDHVLQQLTKAGAKIALHKGQ
WCKTKVNYVGPLIAPQGIERQTNRIQAIQNIKPPANISELRSFLGVCSYSRQFIKDYSDIAR
PLMALLKKDSPFTWTDAQQNAMNKLKKHLCTAPCLAXPHPQREFYLEAGFSSHCLSAG
LYQRHDQDKRVVAYASKTLLPPECKYSDCEKALLCTVWAIQRFSNYIGAQKVIIETCHQ
PTTFLNSQCIRDGVVTNARLATWLMALQGRNVETRYAQNHKSALGNGLAACQSCSDD
TPTVTTDMERPQLQPPTCHKYFEVNVXCQGMPTAYVDGCSYNHEGNLKAGAGVLWM
XPCQPQHFRLGPQSSQYAEIAAVLIALQIASAHNIRDLLICTDSNYARLSFTCHLPSWKKN
GFRTANNKQVKHKHLFLACEDITNEHDMTIYCKKVKGHSKQPGTNKDLNDQTDALAK
AGALHGNLWSPPTGPHKPTITAITRSYRTKSTTLPTSQPLTLTPQISKNDIADMQASDSTI
QTILRHLSSPSDPPVTAVNLATSPSLKRLHDVKPMLRVIDGILSYAPDDRTVPKLVVPQC
QRGVMLMYAHDAPCAGHHGTKATYETLKQVVYWPGMQQDVTEYVKGCLRCCQFQP
ANPNYRAPLQRRGVTFPWSDLQIDCVGPLPRSTRGNKYFITVVCQFTKWVECLPAPNDT
AHTTAYLLMNHIFNRFGLPLRVNSDRGTHFTAEVMQQIWKLLGVKAKLHISHHPISSGQ
VERTNRTVVSMLRKYVAANQXDWDVKLPLVLMAIKATPQESTGISPFELMTGRQMTLP
LHLLYQPGDSNLVTAYTTCQYLEELHKHLRSTFSFAQQHLQKSAEGQKAYYDRKASHE
ELNVGDKVWYYCFNQPPQTASNRRSKKFHIGQDRTRL*
> Loki-Tfu consensus sequence *env* gene
ATGCTGCTGTGGCTCTCCCTTCTGTGCCTCCAAATAACCAGAGGACAGCCGCAACCT
GACCTCACCTCTCCGGGCCCGACATCAGGCATCGTATTGAAGGATCAACCGGGACTT
ATTACCAACTGCAGAACTCATGCCCAAAAGGTATATGTGCGACTCGACCCCCGAAA
TGTCTACCAAACGCAGTACCCACAGACGGCCGCACACTACAGTTGGGCAAGGGCCC
GATGGACCGAGAATGCACTTGCACATGCCGAGGCTGATATCAAACACATGCTCACT
CAGTTATANAAGATGACTGTCACTCAGGCAGAGCTAAATGGGCACAACCGAAGGAA
CAAACGGTTCCTGGGGGCCTTGCTGGGAACCGCGGCTGCGGTGGGCACTCTTTTCAA
CCTAGGTCTGTCCAGTGTAAACTCTGTAAGCCTGTCGACTCTGAGACAGCATATGAG
AGAAATCCAGACTGAGATACCCCAGTTGAGAGAACAACTCACACAACAAGGGCAA
ACCCTCCAGACCATAGGCAAGTCATTAAAAGGCACCTTAGTGGTCCTCAACACCCAC
AGCATCCTGATGAACCAAACCGTGAACTCCATAAAACAGCTGTACTCGATCCTACA
AAATGATGTGGCCCAAATCCAGCTGGTCACAGCACTGATGGCCGACATGCTGCGAG
AAGTGAGTTCTTCCATCGACAGCCTAGCCATGGGCAGGATCCCTCCATACCTGGTGC
CCTTAAGCCTTGCACAAACTGCCCTGTCCTCCGCTACCACTCATTCAACTGATTCCCT
TCAAGCCCATCTTGCATACGCACTGGGAGGATCCGTCCTACTTCATCGACCCAGAAC
AGGGTGGATGGCCTTCCTCCTAAACCTACCAATCATTGAATCTGACAACATCTATCG
GCTCAAAGATGTGGTCAATGTCGGTTTTTGGAAAGGAAACACCCACATCAAGGTGC
ACACCCCTGAGGTGGTAGCCTACCACGACAGCAATCCACAGTTATATCTGGCCCCAA
ACCTGCGCATGTGTACCCGCACCAAGGACATTCACTATCTCTGCCCAAGCAAGCCGT
TCCGCCGAGACAACACCAAAGGAATCTGTGGGTTGCAACCCATGCAACAGGACACC

CGCTGCCCTGCTGAAGCCAAGCCATGCAGCCAGGACACGCAAGCAGAAATCGTTGG
AGATCGGTGGCTGGTCAACACTCCCATCCGTACCGCTACATTAACCTACGACAAACA
TGATACTGCCACTCGCGTGAACTTACCCAGCCAAAATATGTGGATTCAAGTCCCCAA
GGGTGCCATCCTCCACCTCGGGGACCTCGCTCTATTCCACCTCCCCAGTGAAGAGTA
TCAATCAGAGGTGGAAATCACGGACTTCTTCAAAATTCACAATCTCACCCTGGACCC
TGAATTGGAACTGAGAATCGAGGAGGGGGGAGGCAAATGATCGACATCACCCCAG
TCGACACGGCCCTCCAAGCATTGTCCCGACTACCAATCCCAGCTAACTTCCCCGTTG
CCCGAGCCTGGACTGCCGCTGACACAGCACTCTGCATCTCCACAGTCATCGGGTATG
CACTGAGCCTCGGCCTCGCCCTACTTCTGTTCAAACGAATCAATGGAATACAAAGAT
CCATGACCAGGTGTACAGCTGCTCTTCCCCGTGCTTTCAACCGTAACCACCGGAGAA
GCGAGACCCAACCAGAATCAACGATAAGTCTCACCAAACTCAACCCCATGACTGAG
GAACCCAGGGCTGAGGGAAGCTAA
> Loki-Tfu consensus sequence Env protein
MLLWLSLLCLQITRGQPQPDLTSPGPTSGIVLKDQPGLITNCRTHAQKVYVRLDPRNVY
QTQYPQTAAHYSWARARWTENALAHAEADIKHMLTQLXKMTVTQAELNGHNRRNKR
FLGALLGTAAAVGTLFNLGLSSVNSVSLSTLRQHMREIQTEIPQLREQLTQQGQTLQTIG
KSLKGTLVVLNTHSILMNQTVNSIKQLYSILQNDVAQIQLVTALMADMLREVSSSIDSLA
MGRIPPYLVPLSLAQTALSSATTHSTDSLQAHLAYALGGSVLLHRPRTGWMAFLLNLPII
ESDNIYRLKDVVNVGFWKGNTHIKVHTPEVVAYHDSNPQLYLAPNLRMCTRTKDIHYL
CPSKPFRRDNTKGICGLQPMQQDTRCPAEAKPCSQDTQAEIVGDRWLVNTPIRTATLTY
DKHDTATRVNLPSQNMWIQVPKGAILHLGDLALFHLPSEEYQSEVEITDFFKIHNLTLDP
ELELRIEEGGRQMIDITPVDTALQALSRLPIPANFPVARAWTAADTALCISTVIGYALSLG
LALLLFKRINGIQRSMTRCTAALPRAFNRNHRRSETQPESTISLTKLNPMTEEPRAEGS*
>Loki-Sau consensus sequence
ATAATCAATAACGATTCCCAACAAATGGTGCCCCGTGTGAGGCGATCTTATTACCAG
CTTTATTGCAATATTTGTGCAAAAATCCAGTTATAAATTTTTAATAAAAAAAAATAA
TAATAATAATATATATATATTTTTTTTTATTTATTTTTTTTTCTCTTTTTTTTAATTTTT
TTTTTTCCCCCTCCTCTTTTTCAAGCTGCCAGATAAAGCCAACTGTGGTGTGTGAAGC
ACAATTCGTGTGCATTGTTGTTATTAGTATTGTGTTGTGTTGTTAGTGAATTAGTTTG
TGCAACAAAGAAACTTCCTGTTGCAAGATTAATCTTGAACACTTGTAGCTATTGTCC
CAGAGGTGAACAGATCACCTCATCGACCAGTCACAGGGTCAGTACTGTGTGGGAAC
AATCAGCCTATATATACCTATTGTCCGTAGGTTGTAAGTCACAGGTACACCTCTCAC
CTGTGCATCAACAGCAGGTTAACCACTCACACATAAGCAAACTGACCCACAGGTAG
CTCTTTCCTTGCTACCAGAGCGAGAATGAGCCTCAACGAAACCTGTGAGACAGGAG
CGGAAGCTTCTAACGAGTCGACCCCGAGAGATACGGTGGCCGGACTCATGTCTCTCT
CTCCCGAAGAGAAAAGCTATTTAGATAGTTATAGTATGGAGACATGCCAAAAGAAG
ATTGAGGAGCTTATAAGCTACATTGATAACCTGGCGGGGAAGCCAACGCGCACAAG
AGCTGATGTTCTGGGTCAGCTCACAATGTTGTATCAGCGTGAAAGGGAATTGCAGGA
CAGGGAGAATTCCAAGCTTCAACTGGAGAATTCCCAGTTGCAGCTGGAGAAATCCC
AGCTTCAGCTGGAGAGTTCCCAGCTGCAGCTGGAGAATACCCAGCTTCAGCGGGAA
AGTTCCCAGTTGCTGCAAGGAGTTTCCCAGTTGCAGCAGGAGAAAAGGGGGTTGAC
TGCTCCTGCACCAGCGCAAGCAGTATAGGAACAGTCAGCCAGTGAGGACTCTGAGC
GTGATTCGGACCAACCAATGGCTGATGAGGGAATACAGACGGGGCAGGCCCATAGA
CAAAGATTCTCAGAAACGAGCCTGACAAGGCAAGGGGGGGGGGGCAATCCGCCCA
CAGATCGCCCATTGATCTTGATTCACCAGATCACGCTGGGCCGCCTGCCCGGAACCG
CAAACCAAATAGACGAGCGCCGCCCCCAGGTAACCCTCCTCAGCCACGGTCGAGCC

ACAGCCCATCACGCTCCGACGACCCTCCAGCACCAGCCTATGGGGATAGGGGCAGA
TATCACCAGTCCCAACCTAAAACCTTCGGGTTTGGACTCCCGCCTCACGGTGAGCCT
TCCCTACAGTATTCCCCACACTATGATTCCCCATACTATGGCAGCTACAGAGACACA
TTCTCTGATGATTCCGATGAACCTTATTCCCCTCGGGAATACAACATACGGACCCGT
CAGATAGAGTCCCTCGCCAAGGACATCGAGCGCTTTGACCCTGGTAATAAGGACTCT
AACGTAGACGATTATATCCGAGAACTCAAGCGTGTTCTCCTCGACCTGCCCAACCCA
ACCACCCGAGAAAAACTTAAGCTCCTCTGGAAGACCACTTCCAGAAGTGTCCATGTT
TTCATGGAACAACTCCCGTCCAGAGTCAGGGACCGTTTCTCTGCATCCTGCAAATCT
TTACAGGAAGAGTACTCTCTGTACACTAACCCTACCTCTTCTGCCCTATGTGCGTTCG
CCGTCATACAAAAGAAGCAGGAAACCCCCAAAGAGTACTACCGACGCCTGAGGGCA
GCGTACTTTCAAGGCCGTAATACACCTGGGTTAGAGGAGGAGCCTTACTTCAAATCG
CTCTTTATCCATAACCTTCACGAGAGCATACGCTCTCAAGTGGAACTGCATTGTCGA
ACGAGACAGCTCAATATGTCTGAACTCAAGAGGTACTCGATGATGGTGTGGGAGGT
TCAAGCCCGCCCAAATCGAAAGCCCGAGAGTCAGGACATGAGAGTTCTGGGGATCC
AAGCCTCTGATGAAGCAGATCTACAACTGGAAGGCAATGAGGTGCCTCGACCTAGG
TTCGCTCCCCACAGAGGACCCCCAGAACAACCGTCATTCCGGCAGAATGAACATTC
GTCATTCCGGCCTAATGAACATTCGTCATTCCGGCCTAATGAACATTCGTCATTCCG
GCCTAATGAACATTCGTCATTCCGACACAATGAACAACCGCATCAGGGGCTCAGAA
GCTGGTTCAGGCCAGCACGTCAAGACTTTGCTGGCCCCAAGAACCGGAGTGATCGG
AGAACCAAGAATGGTAAGTATCAGGCGAAACCTTCTTACCAAGCCCAAAAGGAGAG
TCGCAAGCCAGACACGGACGACATGCTCCAGCAATGTGTTAGAGACGCATTGAAAC
AATTCAAAGACGATCTTCCTCCCACAGATTCCAAAAAGGGGCCAGATCACGAACCC
CCGCCACAATGACTAGGCGTAACTGCTTCCTCCCCTACCACCAAGGTTTACCTTGTG
GACTGGGAGAGAGGAGCAGGTAAAGCCGCAAGTCTCCCCGAAAAGGCCACTCCAG
CCCCCGCCGGGTCCCCTTTTCTCACGTTCTTGGGCGACCTTGTCCGCCAAGGCAACG
CCAAACGGATCTACACCTCGATTACGATAGGACAGTGTATCAGAACTCTCGCCTTGC
TCGATTCTGGATCAGAAATCAGTTTGATGAACTCTGACATGTTCGAAAAGGTTGCCA
GCGCAATGCTATCTCTTGGGAAACCTCTTCAGACCGAAGCGTGCAATCTGAACATCA
CGTCCTACACTCAGGACAAGTCACCCGTCACTCGACGTGCGTGGGTTGACATAAACT
TTCAAGAGATGTCACTGATACACCCTATCTATATCTGTTCGCTAGATACCGAACGGC
TGCTCATTGGCCAAGACCTTCTGGATCGCTTGGCTCCGCTCATCGACTGTCACCGGG
GACACATCTGGGCTCAGGTGGAAACACCTAAACCACTCAACCCTGGCGTCAACTCA
TCAGGCCCTGATAGCAACATAGCTGCTATTGAACTTAGGGAAACCATGTTGACGCCC
CAAGTAGAGTCCAGTCTTGACCCCGCTCCTGCAACACAGAGAGCAGATCCCATTCTG
CCACAAATTGGTGAACAGGGGTCGACCACCCTTAGGGAACACAGGTCGTTCCTCTG
CGCACTGGAAAACTCGGGCTCTACTTTGTACGCTCCCAAAGTCATTGGGGGAGCACA
TCTTGATGGCACCCCGGTCCCGGATGTACTTTTGGCACTTTGGTCCGAAAAGTCCGC
TATTGACAAGGACCTTTTTGAGGACCTGCATCAGAAAAATCCGGCACTAACCTTCGT
GCAAAAACGACATCGTCTCCTTTCAGCAGGGACCCCTCAGACCCCTCTTAAAGCAAC
CGGCGTATGTGCACTCCCGGTACTGGTGGGAACAAAGCGACTGACGCACTACATGA
GCGTAGTCCCCCAACTTGGTTCCCCTTTCTTCATGGGAGCCGATCTTCTTGTTAGGCT
TGGCGTGCAATTAGACACAGTCAACCAAGTCCTGTGGTCCCTGGCAGAGACAGAGG
ACCATTTCTTGACCGGTGAACCGGAACACCTGAACTCAGGTCAAACCGTACCACATG
CAGCTCAAGTCACTAGCGAAATTGATGTGACAATTCCAGCACAAAGCGCAGGGTTT
CCCATCCGTCTCGTCCTTATAAAAGGACAGAAACTTCCTGGAACGCAAGCATTTTTC
CAACCCCTACCTCACTTCGAGGAACTCAACCTGACTGTGTGCGGTACACACCTGCTA

GAGTTGAACAACCGATCAACTTATTTACTGGTACAGAACCTCACCCGTGCTCCGATC
CAAATGGCGGCAGGACAGCCATTAGGGATGCTCATCGACAACTCGTTTTATGACTTT
GAGTTGACGGTGCCCATCATCGGTGAGCTGCCCCCGTCATTGGTCCCAGCTGGTTCT
GTACAAAATGTCCTGTTCACCTTCCCGACAAAAATGATCACTATTGCCCGTCACGAA
CTCCTGCAACAGGAAGCGATCTGTGGCGCAAGGCTGGACACTGAGGGCGAGATGGT
AGTCTATTCGCTCGTGGCCCAGTCAGAGATCACGAGCCCAACGGGCTCCAAAACCC
AGCATTTGCCAGAGGAGGCGTACCCTGGTTTTGAGGCAGAAATCGAGCAACAACTT
GCGAAAGCTGACGCTTTAACAACCGAGGATCAGAGAGATGCCCTCCGGGAACTGTT
CCGAGAGTTCCAGGCCATCTTCTCGAGAGATTCTAATGACTGTGGAGTCACAGATCT
CCACACGGTGCGCATCCCCACGGATCCGCAAGCGCCCCCGACCTTTGTACGTCAGTA
CAGAATCCCTTTAGCAGCGTATGAGTCCATTCAAGATATCCTGGATAAACTGCTGCA
AAAGCAAATTATCAGGGAATGCAACTCGACTTACAACTCGCCCATCTGGCCAGTGTT
GAAACCGACTGGTAAGTGGCGTCTCACCATTGACTACCGTCCCCTGAACAAACAGG
TTCCGCTTTCCCGGTGGCCAATGATCCACTTGGATCAGGAGCTGGCAAAGGTTAAAG
GAGCCTGTTTCTTCTCAACCGTTGACATAGCAAATGGTTTCTGGACCATGAAGGTGG
ACCCCGCAGACCAGTATAAACTAGCCTTCTCATTCTGCAACCGCCAGTATACCTGGA
ATCGTTGCCCATTTGGGTACTCAAACTCACCAGCGGAGTTCAACATTTTTCTCCACA
AAGCAATGAGTGACGCGGCAGCGCGGCAACCTCATATACGTGGATGACATCCTC
ATGAGGAGTTGTACTTTTGAAACCCACCTGGCAGAAATCCGACACGTGCTCAACCAA
CTGACCCGAGCAGGGGCCAAACTGGCCATTGCTAAGGGCCAATGGTGCCGCACGAA
AGTCGAGTACGTTGGCCTAACGGTAGGGGGGAGTGGAATCGAACCTCAAACTGGGA
GGATTCGGGCCATTCAAGACATCAAAGCTCCAACCAATGTGTCTGAGCTGCGGAGCT
TCCTCGGGGTCTGTAACTATTCACGGCAATTCATTGAAGATTATGCTGAGATAGCAA
GACCACTCACGGAGCTTCTTTGTAAAGACAAGCTATTCCAATGGGAGGAAGCTCAG
GAGCTTTCTTTCCAACAAATGAAAAGGAAACTCAGCTCAGCACCTTGCCTGGCTTAC
CCAGACAAGGACCGAGAGTTTCTCCTGGAGGCAGGTTTCTCACAAAAATGCCTCAGT
GCAGCCCTAGCTCAGAAATACGACTCTGACAAAAGAGTAGTGGCCTACGCCAGTCG
ACCACTGAGCAGCGTAGAACTAAAGTTCTCTGACTGTGAGAAAGCCCTCCTAGCTAC
CATCTGGGCAGTGGAACACTTTCGGAGCTATGTTGGGGGCCAGAAAGTAGTGATCG
AGACCTGCCATCAGCCAGTCACCTTCTTGAACAGTCAGCGTCTCAGAGAAGGACGT
GTATCTAACAGCCGCATCGCCTCCTGGATGATGGCATTACAGGGGTACGATGTGGAG
GTGAAATACGCACAGAACCAGAAAATGGCTCTGGGACAAGGTTTGGCTGGCTGCCA
ACACTGCGACTGCTCCGATCAGCAGAACCCCCTCCCTCTTACAGTGTCTACACCAAC
ACTGCCTTCCAATCATCATTATTATGATGAAAACATCTGCCAGGGCCTGCCAAAGGT
CTATGTAGATGGCTGCTCGTTCCACCATGAGAGCCAACTCAGAGCCGGTGTCGGAGT
CGTCTGGGACCATAAAAATGATGAACCAGATCATTATCACTTGGGTCCTAAATTGAG
TCAATATGCAGAAATTGCAGCGGTGCTGATTGTACTCCAACAAGCGGCCAAGCTTGA
TGTGGTGCAGTTAGTAATCTGCTCTGACTCAAATTACGCTCGCCACAGTTTCATCTCA
CACTTCCCCACCTGGAAAGTGAACGACATGAAGAATGCTCGAAATAAACCCGTCAA
ACATGCCGAACTCTTTCAGGCATGTGATCTGCTGGTCACAGATAAGGGAATAATGGT
GTACTGGAAGAAGGTCAAGGGACACTCTCAGATCCCTGGCCCTGAAAAGGACGGGA
ATGATGAGGCGGACCGCTTGGCAAAGTTAGGTGCAGAGAAGGGTACACCGTGGGAG
TTTCAGACCGACTGGCTCCCTACCCCCCAGACAGGCGTGGTCAATGCCATCACTCGC
AGACAAGCACGAGAGAAACTAGACAACTCTCAAACTGATGGTCAGACACTACACCT
GGGTCGAAAACCCGATGATATAGATCTGGTCACCATGCAGGAACATGACCCTAATC
TTCAAATCGTTCGACAATGGGTCATAACAAGCCCTGCGAGAGGAGCTCCTGAACCC

CTCTCAGTCAGCTCCCGAGAACTGCAGGCTTTCCATCGCGATTTTCCAGACTTAAAG
CTGGAGAAGGACTTGCTGGTATACAGCCAGGGTGGTCAAGGACCGACCCGTTGGGT
TGTGCCCACCGACCATAGGGGGGTGATGCTCTCCCATGCTCACGACACCCCTATTGG
AGGACACCGAAGCTACAAAGCCACGCTTAAAACCCTCCAGCAAGTAGCCTATTGGC
CATTTATGGCCCGCGACACAAAACTTTATGTTCAGGGCTGCCTGATCTGCTGTCAAT
TCCAGCCCTCTCGACCACTTAACCGTGCCCCACTTCAGGAAAGGGGAATCACATTTC
CTTGGTCTCACCTCCAGGTTGACTGGATAGGACCTGTCCCCAAGTCCTCCCGAGGTA
ACAAGTACCTCCTTACGGTCACCTGCTCCTTCACCAAATGGGTTGAGTGCCTCCCGG
CGCCCAACGACACTGCCGTCACAACAGCGGTGCTGCTGATGAACCATGTTTTCAGCC
GCTGGGGCCTCCCACTCTCCATTGACTCAGATAGGGGTACTCACTTCACTTCGAGCG
TAACGACCGCGTTGTACGACATCTTGGGGGTGGAAGCCAAATTCCATATTGCGTATC
ACCCACAGTCATCAGGTCAGGTTGAGCGCGCCAACCGTACCATTGTCAACATGCTGA
AGAAGTATGTTGGGAGTAATGGCAAGGATTGGGACATGAAACTCCCTTTGGTGTTA
ATGGCAATTCGGTCAACCCCTCATCGCTCTACGGGAGTCACACCATTCGAGATGATG
ACTGGCAGACAAATGACCCTGCCAATACATCTCCTGTACCATCCTGAGGATGCCAGC
ATCGCAACCGCCTACACTGCACATCAGTACGTGGCGGACTTGCGAGAACACCTTCA
GGCTACCTTCGCCTGGGCCCAGGTAAACCTGGAGGCCAGCGTAAAAGGGGCAAAAG
CCTACTACGACAAAACTACGTCCAACCGCGAGTACGAAATAGGTAACAAAGTTCTG
TATTTCAGGTTCGCACAACCGACGGGGATTTGTAAAAAATTCCTTCCCCGCTGGTCG
GGCCCCTTCGAGGTTGTCGGTAAGCTCTCACCAGTAGCCTACCGAATAAGAGTCACG
AAACCAAAACAACCATCGACCTACAAATGGGTCCATGCCAACCAAATCAAACTTTT
CAAACCATTTACTCCCCCGGGTCAAATCACTAACCCCACTCCTGACCGGGAGACATC
ATCTCCTGACCAAGGAAACTCAGAGGAGTGTTAGTGGCCAGAGTAGAACAGGGACA
TGCATGCATCACACAGTTAGCATGCACTAACCCTAAAAACCCCCCCAAATGCTCATT
AGTAACAACCATCTGATCTGTTTTCAGGATGGAACTCGCTTGGATCCTTGGCATCAT
CTTTGGCCTACGGGTCATCCTAGCAACTGAGATGACTGAACCAGGCCCTCCAACTGG
GATTGTCCTACAGGAAACACCCGGGCTGCTCATCACCCATTGTGACCTCTACACCCA
GAGATTCTATGTACGCTTAGATCCTTGGGATGTGTATCGTAAGCACATCCGCCTGCC
TCCTAAACTGACAGAGGGGAGATTAAGTGGGGATCAAACCCACAACACCATTGAAC
AAGCCAAACAGAGAACTATCCATGTTCTTGAGCAATTAGAAAAATTCTTAACAACTG
AGAGAGACCTCACTGACAAAAAGAGACCCAAACGATTCCTGGGTGCCCTACTCGCG
GCAGCAACTGGGATCGGGTCCTTATTCTCTATCGGCCTCTCAGCTGCCAACTCTGTC
AGCCTCGGGGCACTACAAAGGCACATGGGTGAACTTGAGGAAGAAATGCCAGAGAT
CCAACAGAGACTCTCCCTCCAACGTGAACAATTACAGGATCTGGGTGAAACCCTCA
AGGGTACCATAGTCACAGTAAATCTACATTCTGCCCTAATCAATAACACGGTGCATG
CCTTAGAAACATTGTCAGAAGTAGTGAAGGAAGACATCACCTATGTGCGGGTAGTG
AGGGACCTAATGCAGGACCTCATAAGAGAAGTCAGCTCCTCAGTCAGCAGCCTCAG
TGCTGGTAAGATACCAACCTACCTGGTACCCTTAGACCTGGTGGAGAATATTCTCAA
AGACGCTACTACCTCAGTTGTGCAACCTTTACAAGTACACCTGGCATACAATCTTGG
CAGCGCTATTCCAATTTCTGTAAACCCTCGGAACTTGGAGATTGGTTTCATATTGAA
CCTACCAATAATCGGTAAACAGAATGTGTACCAACTTAAATCTGTGTCAAACGTTGG
ATTTTGGAGAGATAACACACACATACACCTTAGAACCCCACCAACACTTGCATACCA
CAACGATAACCCCTCACTCTACCTCGTTCCTAATCTTAGTATGTGTACCAGGACAAA
GGACATTCACTGGGTCTGTCCAAGTAACCCCTTTATCCGCGATGTGACTAACTACCT
CTGCGGTTTGAGGACCGATGCACCTGAGCAAAAGTGCCAAGGTAGACTGTCCCTCA
AGGATGAAGAAACAGAGACCCGAGTAGAAAGAGCGGGCAGTAGATGGCTCGTGAG

CACACCTGCCACGGAAGCCATAGTGTCATATGACCGCCATGATACAGCCACTAAGC
TAAACCTGCCTAACCAGACAATGTTCATAACAGTTCCTCAAGGTGCCACAGTCCACA
TCCAAGACATAATCCTCCATCATCTCGATCCAGACAGACATGACGTAGAGATCGAG
ATCATGGACGCCTTTGTCGGCCAGAACCTCACCATTGAAGACGATATTCAGCAACAA
CTCTTAGCTGAAGGCACCAAAACAGTAAAATTCAAGGTGAGGCCAAATGGACTTAC
AACAACAATATCTGGACAGACAGGCAGGTCCTCTTACCAGGCGCACACCATTAGCG
TAACCGCTATTGCGCTGCTGCTCAGTGGATGGATCATCACGGCAATAATCGCACACA
TAATGTACAAATACATTCAAAAGCTCCAGACAAAACTAGACTCTGCCCTCCTAGTAC
CCCCACGTTTCCAACAACAGATTGACACCATGCCTCTAGCCGCTCTCGGTCCTTCAG
GAGGACCGAGGAAAGCGTACTAACCGTTTGCTAACAACCACATTTCCTCTTATTTCA
CCTATCTTTCTTTTCTTCTTTGTTTCCAATTGTGTTTTTGGTGCTTTTAAGAGAAATGC
TGACGTGTTGTTTGTAAAAATGTTTAGATAAGCTAGGGACATTTAAGAACCTTCTGC
CCAAACTGTGCCTCTCTGTCAATGCCCAGACATTAGCCTCTGAACTGTTTTGAACTG
AATTGAACTGAACTGAATTGTGAAACCATAATCACTCTCTCGTCTCAAGGACATGTG
GATTATACAACTGCCCATGTTTTCGGGGCTCAAGGACAGTTAGGCCCTTGGGTTGCT
CCGAAGGCCGTGGCCAGCACCCCGCCCTTCGGACCAAAGGGGGGACTGTTGGGCCT
CATGCCTCTATACNATAATATTGTAAAAAAAAAAAAAAATAACATGTAATTATAAAC
AGTGTTTTAATATGGATATAAAAAAAATAAACGCAGATAGTGCAAACGCACTGTTT
GGGCCTCAATGCCTCTATACCACATGCCCTCAAACAAGAGAAAAGGCCTGAAGGTT
GAAAGCCCAAAAGCCCTAAACTACTAACTCCCCCGCCCAACCACGGGACTCAGATA
CAAAGTGACCATAAAAGGGCTCCCAAGTGTGGGAAAACCGCACGTCACCGCTGAAC
CCTTGCCCAGAAGACGTCACTGGTCCCTATTGTGCTGTAGCCAGGAAGTTCCCGAGA
TACACGGTGACGTCAACCTAGGAGATTAAAAGGTCGCACCCCCGAAGTACCGGTGT
ATTATTTCCCGACGACACGTGCGTGTCAACGTCTGCAGCTGAAAATGTCTTAAGAAA
CACCTGAGAAACGTTTTCCCTCCCTTTTTTTTTATTGAGCTAAAGTTCAGTTCCCTCCT
CGGTGGACGAGATAGAGACTGTTTTGGCTTCACTTTCGTTTTGTTCGAACATTCAGTA
TTATTTTGGATCCCGAGAGGAACGCTGCACAACCCAGCCGTTCCCTCACTACCTGCA
GAAGGAAGAAAATAATCGTCAGAAGGGACGAGACCGTCCCTCCCTTCGTGTCGACC
GATCGCTGTGTTACAGCCTGCCGTCAGCTGTGAGGACGCTCCAGTCTGATCCCGCGC
TCGCCTCCTGAAGGAGTTGGAGTATACGAGTTAGCCGGGCATCCCACGCTACCGCAT
CCCGCGGCATCCGAAAGTAAGAGGTTCTTTGTTTGGGCAGAAACAGATAGTATATA
GCGTGATTGTTTGTTTTTGTTGTGAAAGGACCGCCGAGTCCGCTTATGGTTACCTGTA
CTGTGTTTTCACTGTGTTTGATACGCTCGCTAGGAGCTCTTTCAGGCAACCCAGCCTG
GGTAGTCTTGCTGCCGTGTGTGACCACAATACGTCGTATTCCGGCTCGATCAGCGGA
CGGCAAACGGACCTCGATTCGCTTGCCGACCACTGTTCGATGGAGAGACGACCGCT
GTGTACCACATGGTCGTTGGGTTGTTTCTGTGTTGTCTGCTAAGCCTCGCTTAGCTCT
TTCCCTCTTTTACTAACACACATACACACACACACGCACACACACATACACGTTGAG
AGACGTATACTGTTTTGGGCGATCTGTTGTAACACACCGTTAACAACTCCGCTATAA
CCCAAGACGTGTGCTTATCATCCCTCCTTCATTGACTATAAAGACGGACTTCTACTA
ATTCGGAATTGTGGGTATCATCCCTGTAAATTCGACCCTCGTGTGGCCAAACCGCAC
AGGAGATCAAAGACGGCGCAACCGCTGCCCTTGTTCAAATCTCGCGCGCACATGAA
GTGACCACCCCTGGTCGTTACAGTCACTGCCGCCATTTTTGAATTCCAGGTACATGA
ATCAGCACATCTGCATTCATCCCCTCGGAATTCACATTCCTGCGTGACAGAGTCACA
TCCGGTGTTTTCTGGAATCTCGTCGGCGAGATTCCTTAGATCACGCCCACTAATTTGT
AGTCCGGTGAACGTGACTTCCCAAGTCACGTTCGGCCATCTTTGTAGTCTTTGTTGGA
GACTTACCTATACACACACACCCTCACACACACATACATACTTACCTACATACGCGT

ACACCTACACACACACACACACACACACACACACATACATATTTACCTACACGCGCGT
ACACCGACACATACACACACACACATCAACACTTACCTACACAGCTACACCGACAA
TACACACACACACACAATACACATTACCTACACACGTACTCTACACACACTTTGGTG
CTTAATGTTTTATCTCACTGTCCAATAAGTTAAAGGAAAGTTGTTTGGTTAAAGGTG
ATGCAAGATTGTTTGTGGGACAGACTTTAATACATTTGACTATTGAAATTTGCATCT
GTGAATTATTTGTACTAATTTCATGTATTGTAATAACTGTGGCTTGAGATGTCATGCT
CGATTCAGCATCTCTCACT
> Loki-Sau consensus sequence *gag* gene
CTGCCCAACCCAACCACCCGAGAAAAACTTAAGCTCCTCTGGAAGACCACTTCCAG
AAGTGTCCATGTTTTCATGGAACAACTCCCGTCCAGAGTCAGGGACCGTTTCTCTGC
ATCCTGCAAATCTTTACAGGAAGAGTACTCTCTGTACACTAACCCTACCTCTTCTGCC
CTATGTGCGTTCGCCGTCATACAAAAGAAGCAGGAAACCCCCAAAGAGTACTACCG
ACGCCTGAGGGCAGCGTACTTTCAAGGCCGTAATACACCTGGGTTAGAGGAGGAGC
CTTACTTCAAATCGCTCTTTATCCATAACCTTCACGAGAGCATACGCTCTCAAGTGG
AACTGCATTGTCGAACGAGACAGCTCAATATGTCTGAACTCAAGAGGTACTCGATG
ATGGTGTGGGAGGTTCAAGCCCGCCCAAATCGAAAGCCCGAGAGTCAGGACATGAG
AGTTCTGGGGATCCAAGCCTCTGATGAAGCAGATCTACAACTGGAAGGCAATGAGG
TGCCTCGACCTAGGTTCGCTCCCCACAGAGGACCCCCAGAACAACCGTCATTCCGGC
AGAATGAACATTCGTCATTCCGGCCTAATGAACATTCGTCATTCCGGCCTAATGAAC
ATTCGTCATTCCGGCCTAATGAACATTCGTCATTCCGACACAATGAACAACCGCATC
AGGGGCTCAGAAGCTGGTTCAGGCCAGCACGTCAAGACTTTGCTGGCCCCAAGAAC
CGGAGTGATCGGAGAACCAAGAATGGTAAGTATCAGGCGAAACCTTCTTACCAAGC
CCAAAAGGAGAGTCGCAAGCCAGACACGGACGACATGCTCCAGCAATGTGTTAGAG
ACGCATTGAAACAATTCAAAGACGATCTTCCTCCCACAGATTCCAAAAAGGGGCCA
GATCACGAACCCCCGCCACAATGA
> Loki-Sau consensus sequence Gag protein
LPNPTTREKLKLLWKTTSRSVHVFMEQLPSRVRDRFSASCKSLQEEYSLYTNPTSSALCA
FAVIQKKQETPKEYYRRLRAAYFQGRNTPGLEEEPYFKSLFIHNLHESIRSQVELHCRTR
QLNMSELKRYSMMVWEVQARPNRKPESQDMRVLGIQASDEADLQLEGNEVPRPRFAP
HRGPPEQPSFRQNEHSSFRPNEHSSFRPNEHSSFRPNEHSSFRHNEQPHQGLRSWFRPAR
QDFAGPKNRSDRRTKNGKYQAKPSYQAQKESRKPDTDDMLQQCVRDALKQFKDDLPP
TDSKKGPDHEPPPQ*
> Loki-Sau consensus sequence *pol* gene
TTGGGCGACCTTGTCCGCCAAGGCAACGCCAAACGGATCTACACCTCGATTACGATA
GGACAGTGTATCAGAACTCTCGCCTTGCTCGATTCTGGATCAGAAATCAGTTTGATG
AACTCTGACATGTTCGAAAAGGTTGCCAGCGCAATGCTATCTCTTGGGAAACCTCTT
CAGACCGAAGCGTGCAATCTGAACATCACGTCCTACACTCAGGACAAGTCACCCGT
CACTCGACGTGCGTGGGTTGACATAAACTTTCAAGAGATGTCACTGATACACCCTAT
CTATATCTGTTCGCTAGATACCGAACGGCTGCTCATTGGCCAAGACCTTCTGGATCG
CTTGGCTCCGCTCATCGACTGTCACCGGGGACACATCTGGGCTCAGGTGGAAACACC
TAAACCACTCAACCCTGGCGTCAACTCATCAGGCCCTGATAGCAACATAGCTGCTAT
TGAACTTAGGGAAACCATGTTGACGCCCCAAGTAGAGTCCAGTCTTGACCCCGCTCC
TGCAACACAGAGAGCAGATCCCATTCTGCCACAAATTGGTGAACAGGGGTCGACCA
CCCTTAGGGAACACAGGTCGTTCCTCTGCGCACTGGAAAACTCGGGCTCTACTTTGT
ACGCTCCCAAAGTCATTGGGGGAGCACATCTTGATGGCACCCCGGTCCCGGATGTAC
TTTTGGCACTTTGGTCCGAAAAGTCCGCTATTGACAAGGACCTTTTTGAGGACCTGC

ATCAGAAAAATCCGGCACTAACCTTCGTGCAAAAACGACATCGTCTCCTTTCAGCAG
GGACCCCTCAGACCCCTCTTAAAGCAACCGGCGTATGTGCACTCCCGGTACTGGTGG
GAACAAAGCGACTGACGCACTACATGAGCGTAGTCCCCCAACTTGGTTCCCCTTTCT
TCATGGGAGCCGATCTTCTTGTTAGGCTTGGCGTGCAATTAGACACAGTCAACCAAG
TCCTGTGGTCCCTGGCAGAGACAGAGGACCATTTCTTGACCGGTGAACCGGAACAC
CTGAACTCAGGTCAAACCGTACCACATGCAGCTCAAGTCACTAGCGAAATTGATGTG
ACAATTCCAGCACAAAGCGCAGGGTTTCCCATCCGTCTCGTCCTTATAAAAGGACAG
AAACTTCCTGGAACGCAAGCATTTTTCCAACCCCTACCTCACTTCGAGGAACTCAAC
CTGACTGTGTGCGGTACACACCTGCTAGAGTTGAACAACCGATCAACTTATTTACTG
GTACAGAACCTCACCCGTGCTCCGATCCAAATGGCGGCAGGACAGCCATTAGGGAT
GCTCATCGACAACTCGTTTTATGACTTTGAGTTGACGGTGCCCATCATCGGTGAGCT
GCCCCCGTCATTGGTCCCAGCTGGTTCTGTACAAAATGTCCTGTTCACCTTCCCGACA
AAAATGATCACTATTGCCCGTCACGAACTCCTGCAACAGGAAGCGATCTGTGGCGC
AAGGCTGGACACTGAGGGCGAGATGGTAGTCTATTCGCTCGTGGCCCAGTCAGAGA
TCACGAGCCCAACGGGCTCCAAAACCCAGCATTTGCCAGAGGAGGCGTACCCTGGT
TTTGAGGCAGAAATCGAGCAACAACTTGCGAAAGCTGACGCTTTAACAACCGAGGA
TCAGAGAGATGCCCTCCGGGAACTGTTCCGAGAGTTCCAGGCCATCTTCTCGAGAGA
TTCTAATGACTGTGGAGTCACAGATCTCCACACGGTGCGCATCCCCACGGATCCGCA
AGCGCCCCCGACCTTTGTACGTCAGTACAGAATCCCTTTAGCAGCGTATGAGTCCAT
TCAAGATATCCTGGATAAACTGCTGCAAAAGCAAATTATCAGGGAATGCAACTCGA
CTTACAACTCGCCCATCTGGCCAGTGTTGAAACCGACTGGTAAGTGGCGTCTCACCA
TTGACTACCGTCCCCTGAACAAACAGGTTCCGCTTTCCCGGTGGCCAATGATCCACT
TGGATCAGGAGCTGGCAAAGGTTAAAGGAGCCTGTTTCTTCTCAACCGTTGACATAG
CAAATGGTTTCTGGACCATGAAGGTGGACCCCGCAGACCAGTATAAACTAGCCTTCT
CATTCTGCAACCGCCAGTATACCTGGAATCGTTGCCCATTTGGGTACTCAAACTCAC
CAGCGGAGTTCAACATTTTTCTCCACAAAGCAATGAGTGACGCGGCAGCGCGCGGC
AACCTCATATACGTGGATGACATCCTCATGAGGAGTTGTACTTTTGAAACCCACCTG
GCAGAAATCCGACACGTGCTCAACCAACTGACCCGAGCAGGGGCCAAACTGGCCAT
TGCTAAGGGCCAATGGTGCCGCACGAAAGTCGAGTACGTTGGCCTAACGGTAGGGG
GGAGTGGAATCGAACCTCAAACTGGGAGGATTCGGGCCATTCAAGACATCAAAGCT
CCAACCAATGTGTCTGAGCTGCGGAGCTTCCTCGGGGTCTGTAACTATTCACGGCAA
TTCATTGAAGATTATGCTGAGATAGCAAGACCACTCACGGAGCTTCTTTGTAAAGAC
AAGCTATTCCAATGGGAGGAAGCTCAGGAGCTTTCTTTCCAACAAATGAAAAGGAA
ACTCAGCTCAGCACCTTGCCTGGCTTACCCAGACAAGGACCGAGAGTTTCTCCTGGA
GGCAGGTTTCTCACAAAAATGCCTCAGTGCAGCCCTAGCTCAGAAATACGACTCTGA
CAAAAGAGTAGTGGCCTACGCCAGTCGACCACTGAGCAGCGTAGAACTAAAGTTCT
CTGACTGTGAGAAAGCCCTCCTAGCTACCATCTGGGCAGTGGAACACTTTCGGAGCT
ATGTTGGGGGCCAGAAAGTAGTGATCGAGACCTGCCATCAGCCAGTCACCTTCTTGA
ACAGTCAGCGTCTCAGAGAAGGACGTGTATCTAACAGCCGCATCGCCTCCTGGATG
ATGGCATTACAGGGGTACGATGTGGAGGTGAAATACGCACAGAACCAGAAAATGGC
TCTGGGACAAGGTTTGGCTGGCTGCCAACACTGCGACTGCTCCGATCAGCAGAACCC
CCTCCCTCTTACAGTGTCTACACCAACACTGCCTTCCAATCATCATTATTATGATGAA
AACATCTGCCAGGGCCTGCCAAAGGTCTATGTAGATGGCTGCTCGTTCCACCATGAG
AGCCAACTCAGAGCCGGTGTCGGAGTCGTCTGGGACCATAAAAATGATGAACCAGA
TCATTATCACTTGGGTCCTAAATTGAGTCAATATGCAGAAATTGCAGCGGTGCTGAT
TGTACTCCAACAAGCGGCCAAGCTTGATGTGGTGCAGTTAGTAATCTGCTCTGACTC

AAATTACGCTCGCCACAGTTTCATCTCACACTTCCCCACCTGGAAAGTGAACGACAT
GAAGAATGCTCGAAATAAACCCGTCAAACATGCCGAACTCTTTCAGGCATGTGATCT
GCTGGTCACAGATAAGGGAATAATGGTGTACTGGAAGAAGGTCAAGGGACACTCTC
AGATCCCTGGCCCTGAAAAGGACGGGAATGATGAGGCGGACCGCTTGGCAAAGTTA
GGTGCAGAGAAGGGTACACCGTGGGAGTTTCAGACCGACTGGCTCCCTACCCCCCA
GACAGGCGTGGTCAATGCCATCACTCGCAGACAAGCACGAGAGAAACTAGACAACT
CTCAAACTGATGGTCAGACACTACACCTGGGTCGAAAACCCGATGATATAGATCTG
GTCACCATGCAGGAACATGACCCTAATCTTCAAATCGTTCGACAATGGGTCATAACA
AGCCCTGCGAGAGGAGCTCCTGAACCCCTCTCAGTCAGCTCCCGAGAACTGCAGGC
TTTCCATCGCGATTTTCCAGACTTAAAGCTGGAGAAGGACTTGCTGGTATACAGCCA
GGGTGGTCAAGGACCGACCCGTTGGGTTGTGCCCACCGACCATAGGGGGGTGATGC
TCTCCCATGCTCACGACACCCCTATTGGAGGACACCGAAGCTACAAAGCCACGCTTA
AAACCCTCCAGCAAGTAGCCTATTGGCCATTTATGGCCCGCGACACAAAACTTTATG
TTCAGGGCTGCCTGATCTGCTGTCAATTCCAGCCCTCTCGACCACTTAACCGTGCCCC
ACTTCAGGAAAGGGGAATCACATTTCCTTGGTCTCACCTCCAGGTTGACTGGATAGG
ACCTGTCCCCAAGTCCTCCCGAGGTAACAAGTACCTCCTTACGGTCACCTGCTCCTT
CACCAAATGGGTTGAGTGCCTCCCGGCGCCCAACGACACTGCCGTCACAACAGCGG
TGCTGCTGATGAACCATGTTTTCAGCCGCTGGGGCCTCCCACTCTCCATTGACTCAG
ATAGGGGTACTCACTTCACTTCGAGCGTAACGACCGCGTTGTACGACATCTTGGGGG
TGGAAGCCAAATTCCATATTGCGTATCACCCACAGTCATCAGGTCAGGTTGAGCGCG
CCAACCGTACCATTGTCAACATGCTGAAGAAGTATGTTGGGAGTAATGGCAAGGAT
TGGGACATGAAACTCCCTTTGGTGTTAATGGCAATTCGGTCAACCCCTCATCGCTCT
ACGGGAGTCACACCATTCGAGATGATGACTGGCAGACAAATGACCCTGCCAATACA
TCTCCTGTACCATCCTGAGGATGCCAGCATCGCAACCGCCTACACTGCACATCAGTA
CGTGGCGGACTTGCGAGAACACCTTCAGGCTACCTTCGCCTGGGCCCAGGTAAACCT
GGAGGCCAGCGTAAAAGGGGCAAAAGCCTACTACGACAAAACTACGTCCAACCGC
GAGTACGAAATAGGTAACAAAGTTCTGTATTTCAGGTTCGCACAACCGACGGGGAT
TTGTAAAAAATTCCTTCCCCGCTGGTCGGGCCCCTTCGAGGTTGTCGGTAAGCTCTC
ACCAGTAGCCTACCGAATAAGAGTCACGAAACCAAAACAACCATCGACCTACAAAT
GGGTCCATGCCAACCAAATCAAACTTTTCAAACCATTTACTCCCCCGGGTCAAATCA
CTAACCCCACTCCTGACCGGGAGACATCATCTCCTGACCAAGGAAACTCAGAGGAG
TGTTAG
> Loki-Sau consensus sequence Pol protein
LGDLVRQGNAKRIYTSITIGQCIRTLALLDSGSEISLMNSDMFEKVASAMLSLGKPLQTE
ACNLNITSYTQDKSPVTRRAWVDINFQEMSLIHPIYICSLDTERLLIGQDLLDRLAPLIDC
HRGHIWAQVETPKPLNPGVNSSGPDSNIAAIELRETMLTPQVESSLDPAPATQRADPILPQ
IGEQGSTTLREHRSFLCALENSGSTLYAPKVIGGAHLDGTPVPDVLLALWSEKSAIDKDL
FEDLHQKNPALTFVQKRHRLLSAGTPQTPLKATGVCALPVLVGTKRLTHYMSVVPQLG
SPFFMGADLLVRLGVQLDTVNQVLWSLAETEDHFLTGEPEHLNSGQTVPHAAQVTSEID
VTIPAQSAGFPIRLVLIKGQKLPGTQAFFQPLPHFEELNLTVCGTHLLELNNRSTYLLVQN
LTRAPIQMAAGQPLGMLIDNSFYDFELTVPIIGELPPSLVPAGSVQNVLFTFPTKMITIARH
ELLQQEAICGARLDTEGEMVVYSLVAQSEITSPTGSKTQHLPEEAYPGFEAEIEQQLAKA
DALTTEDQRDALRELFREFQAIFSRDSNDCGVTDLHTVRIPTDPQAPPTFVRQYRIPLAA
YESIQDILDKLLQKQIIRECNSTYNSPIWPVLKPTGKWRLTIDYRPLNKQVPLSRWPMIHL
DQELAKVKGACFFSTVDIANGFWTMKVDPADQYKLAFSFCNRQYTWNRCPFGYSNSP
AEFNIFLHKAMSDAAARGNLIYVDDILMRSCTFETHLAEIRHVLNQLTRAGAKLAIAKG

QWCRTKVEYVGLTVGGSGIEPQTGRIRAIQDIKAPTNVSELRSFLGVCNYSRQFIEDYAEI
ARPLTELLCKDKLFQWEEAQELSFQQMKRKLSSAPCLAYPDKDREFLLEAGFSQKCLSA
ALAQKYDSDKRVVAYASRPLSSVELKFSDCEKALLATIWAVEHFRSYVGGQKVVIETC
HQPVTFLNSQRLREGRVSNSRIASWMMALQGYDVEVKYAQNQKMALGQGLAGCQHC
DCSDQQNPLPLTVSTPTLPSNHHYYDENICQGLPKVYVDGCSFHHESQLRAGVGVVWD
HKNDEPDHYHLGPKLSQYAEIAAVLIVLQQAAKLDVVQLVICSDSNYARHSFISHFPTW
KVNDMKNARNKPVKHAELFQACDLLVTDKGIMVYWKKVKGHSQIPGPEKDGNDEAD
RLAKLGAEKGTPWEFQTDWLPTPQTGVVNAITRRQAREKLDNSQTDGQTLHLGRKPDD
IDLVTMQEHDPNLQIVRQWVITSPARGAPEPLSVSSRELQAFHRDFPDLKLEKDLLVYSQ
GGQGPTRWVVPTDHRGVMLSHAHDTPIGGHRSYKATLKTLQQVAYWPFMARDTKLY
VQGCLICCQFQPSRPLNRAPLQERGITFPWSHLQVDWIGPVPKSSRGNKYLLTVTCSFTK
WVECLPAPNDTAVTTAVLLMNHVFSRWGLPLSIDSDRGTHFTSSVTTALYDILGVEAKF
HIAYHPQSSGQVERANRTIVNMLKKYVGSNGKDWDMKLPLVLMAIRSTPHRSTGVTPF
EMMTGRQMTLPIHLLYHPEDASIATAYTAHQYVADLREHLQATFAWAQVNLEASVKG
AKAYYDKTTSNREYEIGNKVLYFRFAQPTGICKKFLPRWSGPFEVVGKLSPVAYRIRVT
KPKQPSTYKWVHANQIKLFKPFTPPGQITNPTPDRETSSPDQGNSEEC*
> Loki-Sau consensus sequence *env* gene
ATGGAACTCGCTTGGATCCTTGGCATCATCTTTGGCCTACGGGTCATCCTAGCAACT
GAGATGACTGAACCAGGCCCTCCAACTGGGATTGTCCTACAGGAAACACCCGGGCT
GCTCATCACCCATTGTGACCTCTACACCCAGAGATTCTATGTACGCTTAGATCCTTGG
GATGTGTATCGTAAGCACATCCGCCTGCCTCCTAAACTGACAGAGGGGAGATTAAGT
GGGGATCAAACCCACAACACCATTGAACAAGCCAAACAGAGAACTATCCATGTTCT
TGAGCAATTAGAAAAATTCTTAACAACTGAGAGAGACCTCACTGACAAAAAGAGAC
CCAAACGATTCCTGGGTGCCCTACTCGCGGCAGCAACTGGGATCGGGTCCTTATTCT
CTATCGGCCTCTCAGCTGCCAACTCTGTCAGCCTCGGGGCACTACAAAGGCACATGG
GTGAACTTGAGGAAGAAATGCCAGAGATCCAACAGAGACTCTCCCTCCAACGTGAA
CAATTACAGGATCTGGGTGAAACCCTCAAGGGTACCATAGTCACAGTAAATCTACAT
TCTGCCCTAATCAATAACACGGTGCATGCCTTAGAAACATTGTCAGAAGTAGTGAAG
GAAGACATCACCTATGTGCGGGTAGTGAGGGACCTAATGCAGGACCTCATAAGAGA
AGTCAGCTCCTCAGTCAGCAGCCTCAGTGCTGGTAAGATACCAACCTACCTGGTACC
CTTAGACCTGGTGGAGAATATTCTCAAAGACGCTACTACCTCAGTTGTGCAACCTTT
ACAAGTACACCTGGCATACAATCTTGGCAGCGCTATTCCAATTTCTGTAAACCCTCG
GAACTTGGAGATTGGTTTCATATTGAACCTACCAATAATCGGTAAACAGAATGTGTA
CCAACTTAAATCTGTGTCAAACGTTGGATTTTGGAGAGATAACACACACATACACCT
TAGAACCCCACCAACACTTGCATACCACAACGATAACCCCTCACTCTACCTCGTTCC
TAATCTTAGTATGTGTACCAGGACAAAGGACATTCACTGGGTCTGTCCAAGTAACCC
CTTTATCCGCGATGTGACTAACTACCTCTGCGGTTTGAGGACCGATGCACCTGAGCA
AAAGTGCCAAGGTAGACTGTCCCTCAAGGATGAAGAAACAGAGACCCGAGTAGAA
AGAGCGGGCAGTAGATGGCTCGTGAGCACACCTGCCACGGAAGCCATAGTGTCATA
TGACCGCCATGATACAGCCACTAAGCTAAACCTGCCTAACCAGACAATGTTCATAAC
AGTTCCTCAAGGTGCCACAGTCCACATCCAAGACATAATCCTCCATCATCTCGATCC
AGACAGACATGACGTAGAGATCGAGATCATGGACGCCTTTGTCGGCCAGAACCTCA
CCATTGAAGACGATATTCAGCAACAACTCTTAGCTGAAGGCACCAAAACAGTAAAA
TTCAAGGTGAGGCCAAATGGACTTACAACAACAATATCTGGACAGACAGGCAGGTC
CTCTTACCAGGCGCACACCATTAGCGTAACCGCTATTGCGCTGCTGCTCAGTGGATG
GATCATCACGGCAATAATCGCACACATAATGTACAAATACATTCAAAAGCTCCAGA

CAAAACTAGACTCTGCCCTCCTAGTACCCCCACGTTTCCAACAACAGATTGACACCA
TGCCTCTAGCCGCTCTCGGTCCTTCAGGAGGACCGAGGAAAGCGTACTAA
> Loki-Sau consensus sequence Env protein
MELAWILGIIFGLRVILATEMTEPGPPTGIVLQETPGLLITHCDLYTQRFYVRLDPWDVYR
KHIRLPPKLTEGRLSGDQTHNTIEQAKQRTIHVLEQLEKFLTTERDLTDKKRPKRFLGAL
LAAATGIGSLFSIGLSAANSVSLGALQRHMGELEEEMPEIQQRLSLQREQLQDLGETLKG
TIVTVNLHSALINNTVHALETLSEVVKEDITYVRVVRDLMQDLIREVSSSVSSLSAGKIPT
YLVPLDLVENILKDATTSVVQPLQVHLAYNLGSAIPISVNPRNLEIGFILNLPIIGKQNVYQ
LKSVSNVGFWRDNTHIHLRTPPTLAYHNDNPSLYLVPNLSMCTRTKDIHWVCPSNPFIR
DVTNYLCGLRTDAPEQKCQGRLSLKDEETETRVERAGSRWLVSTPATEAIVSYDRHDTA
TKLNLPNQTMFITVPQGATVHIQDIILHHLDPDRHDVEIEIMDAFVGQNLTIEDDIQQQLL
AEGTKTVKFKVRPNGLTTTISGQTGRSSYQAHTISVTAIALLLSGWIITAIIAHIMYKYIQK
LQTKLDSALLVPPRFQQQIDTMPLAALGPSGGPRKAY*
>Loki-Rca consensus sequence
GTTGTGAATTCATGAAGAAATACATTTTATATTTACTGGCAGAAATATGAGAGTTGC
TTAGGCCTCAAAAAGTAGATGAACTGTGATGTTGCAATAATTGAACTGTTGCCCTTT
TGGAGACCTTTAGGACAAAATAGATGTTCTCATGTTTTTACAAAGTTTTATRCATCAT
CCAATATTGAAGACTTTTACAAGCATGTATTGTGAGGAAAACTCGATGGAACAATG
GCTGCTCTCGTCACAAAACAAGGAGATATTTACATGGTATTAGTGAATAAACAAAA
ACTACCCTTACAGCCATAGGAAACAAGGTTATACAGACTATTCTCACATAGAAATAA
CCTCACCTGGCCTAAACCTATTGTGTTCAAGTGAGGGACTGTTTAAATAGCTAACAC
CTTGGCTAACAAAACCCATTGTGTGAATTTTTAGGAACTGTTTAGACAAACAAGTCA
TTCTAAACATCCAGTGGGAATGTTCAATTCAACACATGGTGGGAAAGAGTTGAATTT
TGTGCCAGTCAAAAATATTTTAAAAGTATCTATAAGGTTATTATAAGTCATTTCAAA
CTATCTATGACATTGTAGGTCACAGAGACATATTATAATTTTGGCATCTGCGGACCC
AACAACTTTGGTATGAAGAAAATATATATGAAAGTGTGATTTTATAAGTAGAAGTGT
TTTAAGAAACTAATTAGTATTAGGAAAAACATATAATTATGATTATAACTATATAAG
ACTGTAGCCTATGCTTAGATAAACAAATATTCATTGGAGATACCAGAATTATACAGG
TATCCATATAGATAATCATTTTTATGTAGTATTGATTAATATATATAACAACGTGTGT
ATTTCATAGATAGACCAGTTTTTATAGATATACATTTATAGAGATACATATATATAA
CCAAGTTATGTAATAGTGATTAATCAAAATTATACTGAGTGTATTTTACATTAGACC
GATTACCAGAATATTTAAAGGTATATACTATCCTTACCATTATACACATTATGAAAA
TAGAACGTATTAACTTAAAAGGATAGACTCAAAACACATGTGAGACACCTGGTGGT
TGAAGGTGGTATTATTTGAACTCACTAAATTTCCAGAAAAGGTTAATAATTAAGTCT
CTTAACCAATGGAAAGTGAGCCACACCTTTTGGGCTGAACTTATGATAATTTTATGA
GCTTATTGTCCCAAATTGGGATAAAAGGCTGAGAGGGCCAAAAGGAAGGGATCACA
GACCCCACACCAGATGCACCCCTGACATCAAGAAGTCTGTTGATATTTAACCACTCC
CCATCCTGCTGGACTTTATGACGGAAGCCATTTGATGTGCTTAGATTTTGGCGGAAA
TCTTCCCTTTTAAGCTGATTGGTAAGATACAGACCACACAAATTTTATTTAATTTCTG
TCCTTAAAATGGTAGGGATTATATTTTAGAGGTTAAGTATTGATGTTTATTGACTGTC
TGGGGTGTATTATTATAACCGTGTTTTAGGGAATTCAAGCCAATATTTACATCAATTT
AAATTTCTGTGACCCCAGGGAAGTTTTATTACTCCATTTAATTTCTTAGCTCTGTTGG
TTTAATAAGTAATTCAAATGGGCAATATTGTTCTGTTGACTCGGTGAGATATACTTTG
GATTTCTTCCGGCCGGAGAGATTTGAAGGACTGCTCCCCTGAGCAACAAGGGATATT
GCGTTCTTGAATATAAAGGGTAAATTTCCACCCATATTTTTCCAACAGACCAATTTTG
ACTCCCCTGGGTGATTATAWTAATTTTCCTCACTACAATWCTGATGTTTTGATAATG

TATGTGTGAAGTCTCAATCACCAGAAAATTTAGTTCTGTATGAAATTGTATGAATAT
ACTCGCTAATTAATAAAGTCAAATTATTTACYTATTAAGTTTGTTGGGAGATATTTtA
TACTTTGGTATGAAAACCCCATAAATCAACAAACTCAGGAAAGAATTGTGGTGTGC
ACTGAcAGTAGGAACCAAATTTAGAAAATTACCGTAAGGTTGGAGGCACTGAAAAT
AACACTGAGATCTTAATTAAGATTCCATACCACACTCACAACATAAAATTGGAGGCA
CTGCTGAGATATTTTGTAAGAGGGAATCTTACCACATTACAGAACATACTGGTGATT
GTGTTTGCATATACATTTGAAAAATATCGGTTGTATATTAACATGGCTGGTAGTGAT
AATGAATCTTGTTCATCTAGAAACAGCCAGGATGgTGATACCGCCCCAGCGGTAGCT
ACCTCAAGCCAGAAAAAGGCCACTGGGTCTCGGGAAAACTTAAGTAGGGTCCTTGA
CTATGGTGTGACCGAATATATTCGCAAAAGATTAGACATAAaAGAAAATTTAGAGAA
ATTTATTAAAGAGGAGARAGAACAGGCCCAGAGGTTAAGTGATGATTTTTTTATTGA
AGATGGTCAGGTTTTCCAAATGATTATGGCTTTTGCTAGCGCAAAAGACACAAAAAG
CCGAGACCGAATCATTATAGAGTCTTGGCCGGTCTTaTGGTTTTTGTGTAATGAAGTT
GCCGAACAAATGCCAGCCTTAAAAAGAGAATTGTTGGCCACCCGCACCACCTTAAA
TgACAAAACAGAAAAAATAGAATCCCTGTCAACCAAAATCCAACAGGCTGAGAGTA
GACTCAGACAATCTAAAGAAGGTCAGGAAGAGTTAATGGGCCAGATCCAAAATGCT
CAGACCTTATATGAAGACTACTCAAAAAACTGCAAGAGAAATATCAAAAATGTAAA
CAAGAAATAGAATATTACAAACAAGAAAGAAAAGGCCCTATTGAAGCAGACTCTGC
TGACAGAGATAAAGAGAGTGAGGCAGAAATTGAGAAACTYAGATCTGAGCAGAGA
GCAATGGAGGGACRGATAGAAAGATTAACGGAACAGTTAGACTATGAATCCATGAG
CCGAAGTCATTTAGAGGAAATGGAAAAACGGAAATATTCATATTTCCGGGGGACAC
AGGTAAGTTACAATGAGACAGGGTCAGAAGACAGGTCACCCCCTAAGCTGCAAAGA
TACTCCCCAGAATGTACAGAGGAGGAAGTAGAAAGAGCTTCATATGGGCAGAAAAG
GAAGCCTAGGACCCTAATAAAAAAATATGACACCCCCAAGATGATAAAATTCTTAA
AAGAAACTGTTCCCGTTTTCTCCAAGAAGGGTTCCCAAACCATCACTGACCATATAG
GAACCTACGAACGTACCCTGAAAATCTTAAACATTCAGGACGATCGTACCAAAATT
GAGTTACTGCCATGGATATTTGAAGAAAAGTACAGACAATTCTTTGATTCATTGCAG
GCTCTATACGGTCAGTCATGGAGGGAAATTAAACATAGGTGTCAAATGGAGTTTGG
GCCCTACAAAACTGTTACCGCTGCCAAAGCCGCGGCCTACAATTTAAAATGTGGCCC
TAATCAGAGTCCCACAGAATTCCTAGCCGTATTGAGAAGCGCATACTTTCTGGCTGA
AAAGGACCCCAATCCGGACCAGCCAGAATTCCAGACTCTATTTTTTGAAGCACTACC
TGCTTATATCAAATTAGGACTGGCAAGGGATTTTAAGAAAGGACGTTCCCTGGATTG
GTTAGTCAGRGAGAGTGATCACCTGTACTCAATCCAGCAGTCCAGCGAAGGTGGGG
TGAAAAGAGAAGCCAAACCTAACCCTAAAACCATAGCAGCCACTACTGTGACCCCC
CAACCTCATGGTCTACAAACAGAGAGGAAATCATATGCTGAGGCAGTCCGAGATGG
GAAACCCATGGGACCTCCCGGAGTTTCACCCCAGCCACAGCAGATGATGGAGGACA
GACAGAGGTTGCCAGGTTACGGTAGTGAGGGTGAAAGGGGCCCCAGGTACAGAAAT
AGCCAGAGGGGATATCTGCAGGGGTCCCAGAACAACCAACAATCAAATTGGAATCG
AGATCCCAATTCTAATAGGTACCCCTACTATAACAGGAACCCTAACTTTGTCAGAAA
CCCCAATTTTCAAGGGTACCCACGCTTTAACAGAAATCCAAACTTCCACGGCCACCC
TGGTTACTACAGAGACCCCAACTTCATAAGAGACCCTAATTTTAATGAGAATCCTGA
TAATCGCAGGACACCTCCCAATCAGGAACAGGTCAATCACCCCAGGGTTGGTGTAG
CTCCCCAAATTCCCCAAAATCAAAGGGAACCACAGGAACCCCCGAGAGAAAGGCCA
CAGCACCAAGACCAGAGGAGTGGTGATGAATCTGATGCATCACTGCAGTCCCGTCC
TGAACGCAACCGGGTCAGGCAAAATAGGGCAAATCGCCAAAATACCATGGGCCGGC
GAATCGAGCAGGTAGAGGCACAAATTAATCAGATGTGCGAACAAATGACACACCTG

ATCAATTACCAATCTACCAACCCGTCTTTTTTAGGGGCCGCTCCTGTGACCAACCCA
GGGTCACAGAATGCTTAGAACCCCAGACAGTCCTCCAGCAGGGGACAACCCTGGGT
GCGATGACAGGAGACACACAACACACCCCTGACATGCCCCTAAACAGGGGGGATCA
GGGTGGGATTGTTAATTCTGCTCCTGTTTTTCAGTATCCTTCTAACCCCTCTAATGTA
GCTGAGCCACTGGGAGCCATTGTAAATGACCAATCGGGCGGGATGCCTCCTACTCC
GTGTGCTTCCGTCTCaAGTAGGCCCCGCAGGCCATCCCGCAGGAATCATGAGTAGGG
ATGAGCCGAACACCCCCCTGTTTGGTTCGCACCAGAACTCGCGAACACCGTTAAAGT
CTATGGGACACGAACATGAATAATCAAAAGTGCTAATTTTAAAGGCTTATATGCAA
GATATTGTCATAAAAAGTGTTTGGGGACCTGGGTCCTGCCCCAGGGGACATGGATCA
ATGCAAAAAAAAAGTTTTAAAAACGGCCATTTTTTCAGGAGCAGTGATTTTAATAATG
CTTAAAGTGAAACAATAAAAGTGTAATATCCCTTTAAATTTCGTACCTGGGGGTGTC
TATAGTATGCCTGTAAAGGGGCGCATGTTTCCCATGTTTAGAACAGTCTGACAGCAA
AATGACATTTCAAAGGAAAAAAAGTCATTTAAAACTACTCGCGGCTATTAATGCATT
GCCGGTCCGACAATACACATAGAAGTTCATCGAAATTTAAAGGGATATTACACTTTT
ATTGTTTCACTTTAAGCATTATTAAAATCACTGCTCCCGAAAAAACGGCCGTTTTTAA
AACTTTTTTTGCATTGATCCTTGTCCCCTGGGGCAGGACCCAGGTCCCCAAACACTTT
TTATGACAATACCATGAATATAAGCCTTTAAAATAAGCACTTTTGATTACTCCCATA
GACTTTTAAAGGGTGTTCCGCGGCATTCGAATTTGCTGCGAACACCCCAAATTGTTC
GCTGTTCGGCGAACTTGCAAACGGCCAATGTTCGAGTCGAACATGAGTTCGACTCGA
ACTCGAAGCTCATCCCTAATCATAAGCACAAAGAAGCTGTGACATGCGGTGAGATA
AATGCCAACTGTAAATTATTGGAAAGCCCACAGATCCTCAACCACGCTAATTTCATG
TGTGAGTTGACTGAATCTGATGGAAGATATTTCACTACTGTAAAGGTACATGAGGCC
GGTGATCACCCCATTTCAGGTTTAATTGACACGGGATCCCAAGCCACCATCCTGTCG
CAGGAACAATTTAATGAAGCAATAAATTCAATGCCCAACAAACCAAAACTTAGGAG
CTTTGATTGCTCACTAATCGGGGTGGGGGGAGAGCCACTGCAAGTCATGGGTACTGC
CTGGCTTAAACTCAAATTAGGTAATAAGGAAATTAGGCACCCTATTATTATTGCAAA
TCTCCCACACCATCACATGATAATTGGAATGGATCTCCTTAAGAGGCTCAACACAAT
TGTTGACTGTGTGAATCAAGTAATGTGGTCACAGGTCAACTCCCCGATTAAATATGA
AAGACTTAAAACTCCCCATTCCCGGCGACACTGCCACCTGGTGGAAGAAAGGCCTG
GTTCCCTAGAAATCCACTTCAGGAATAGTCAGTTCCCCGATGTTACGGTTATCCAAC
CTGACAATCATGGCCTGGAAGTAGGAGATGAGGACCATATTTTTAAAATTCCCCGTG
ATAAAATAAAAGAGATCTCCCTCCAGGATAACGTACTGACCATATCCTTCCATCAGG
AAAATAAAACTAAGAAAGGGGGAAATCACGCCCAGTTCCTGACAGAAATAGAAAG
GGTTAGCAACGACCTATTTATCCCTATCCGGATAAATGGTCTTAATAGGTCAAAATA
TGCCAAAATTGATCTAAAGCAAGAGCATAGCTTTGTGAGCCATGGCCTTGTTAGCCA
CCTTTCTGACATGGAAATGATCCGTTTCTCCACTCCTCAGCGACATTGGATGTCAGAT
GCCATTGATGACTCTGAGGACCATAACTCAATTGCTAGATGCCTTTTAAATATCACT
ATAGGTAATAAGACAACCACCCACGTGTTCATAGTGCTGGATAATCCCCAATCCCAA
ATTTACCTCGGCAACGACTTACTGCATCGTTTTGCCACCCAAATTGACCTGATAAAT
GACTGCTTATGGAGTCGACTAAAGGGGGACCCCGAGGGATACCAGGATATGACCTC
AGCCCTGAAGTTCCAGCAACAGATGCCTTATGCTGTTGAGATTCAGGTAGCTCGAGA
AATTATCATCCCTAGGGACATAAAAAATTTCCTACTCCCCATAAATGTAAAAAAGGG
ACAAAAACTGAAGGGATCAGATGCATTAGTATGCCTGTCCCAGAGAATACAGGAGG
AGGGAATAAAAATTACcCACGCACCTTTGATCAATACACATCAGTCTATAATATCCA
TCTACATCCACAACCTCTCAAATCAGGACCTTAACCTACCCAAGGGCACCCTTATCG
GCCTTGCGGTAGAGACGGGGTATTATACGTTTGGAGTCACCAACCAAGTGATTGGAC

TCATACCAGATGAATATTTATCTGAGGAGAAATTAATTGAACAATCGTTCAATTCTT
TACCCGAGGGTATRTTCACCATAGAGTCCATTTACCCCTTCGATTCTAAGGATGGTA
CCTGTAGAATTGAAGAGACATCTCTCGTCTTTGATCAACCAATCAAAGAACAAGAGT
ACCCCAAGGGACCTGACTGGGAGAAAAATACACCTCGCACACCCAATGACCTGACC
ACCAGGCTTGAAGAGGCCTATGAGATAGGGAAACCTGAAAGCTTTCCGGGATTCAA
AGAGATAGTCGAAGAGATCATAGCCCTAGCCGATGCATGCACCAATGATCTTGAGC
GACAATTGCTCCGCGAACTTCTTTTTGAATTCCAGGAAATCTTTGCTCAGGATTCCTA
CGATTGTGGGAATACTGATCTCCACATCGCCAAAATCCAGACCGACCCTCAGGCCCC
ACCAGTATTTGTAAAGCAGTATAGACTTCCGCTAGCCTCCTATGAGTCTCTCGCCGA
GATCATTAGGAAGTTACACGAACGGGGTATAATTAGGCCCATCCACAGCTCGTACA
ATCACCCCATCCTGGGGATTTTGAAACCCAACGGACAGTGGCGCTTGTGCGCTGATC
TCCGACAACTCAATAAGCGTGTCTACATGTCAGGATGGCCCGTACCCTACATAGACC
AGTGCCTTGCCCAACTCCAGGGRTCCCAGGTATTCTCGGCCCTCGACTGTGCCCAGG
GATACTGGACCATAAAAGTCAGTGAAGAAGATCAGTATAAACTGGCCTTCAGTTTTG
GGAACCAACAGTACTGCTGGGTGCGTATGCCTTTCGGCTATATTAATTCTGGCCATG
AATTTGCAGTCTTCATGCACAAGGCTATGCCAGATGCCTTGGAGCGAGGAACATTGG
CATACGTCGATGACATCCTCATCAAAAGTACTGACATGACCAAACACCTCACAGAA
CTCCGACACATCTTGGGCCAGCTAAGGACRGCCGGAGTGAAGTTATCCATTCAGAA
AGCACAATGGTGCCGCACCCGAGTAAATTTCTTAGGCCATGAAGTTACCTCAGAAG
GTTTAAACCCCCAAAAGAAAAAGGTTGAGGCAGTATTTAACTGTAAGGTCCCTACTA
ATCTTAAAGAATTAAGATCGTTTCTAGGAATGGTTAACTATTCGCGTAAGTTCATCG
ACAATTACGCCGAGATTAGCAAACCGTTACTTCATCTGCTGAAGAAAGACACTGAAT
GGTCATGGGGCAAGGAACAAGCAGACGCTGTGAAGGAACTGAAGGTACGACTCACC
CAAGCCCCCTGCCTTGCGTATCCAGAAGGAGGTAAACCCTTTTTTCTTGAGACAGGG
TACACGCAGGTGAGTATGAGTTCCGTTCTTTATCAAAAACAAGAGAATGTCAGTAAG
ATTATCGCCTATGCCAGTAAAACATTGTCCTCAGTAGAGATTAAGTTTTCTGACTGC
GAAAAAGCTTTATTGTCCACTGTCTGGGCATTACAGAATTTCCGCAGTTACATTCAA
GGTGAAAGAATTATTGTTGAGACCGCACACCAGCCTTTACAGTATTTACAGAGTGAA
CGTATCAGGGATGGAAATGTGTCTAATAGCCGCATCACTGCTTGGACTTTGTCCTTG
CAAGGGTGGCCCCTGGAAATTCGATACAAACAGAACAAGAAAAGTCCAGTGGCCCA
GGGCCTTGCTGAGCTGCATGATTGTTCCGCCCCTGATCCCTGCGCGGGGACCACGGA
CACTGACTTTTTAGAAGAACAACTGTTGTCTCCCTACAAGGTTTATGATGATGAATA
CTGTAAACAGTTGCCGTGTGTATATGTAGATGGCTGTGCCTATCATCATAAGGACAA
AGACCAGCTCCAACTAGTGGCTGGCATCGGGATCATATGGACCCAAGGTGTAGCCC
CAATCTCTGTGGGATACAAGCTTGGGCCAAGAAGCAGCCAAGTTGCTGAACTCGCT
GCAATCCTGAAAGCCGTTGAGATGGCCATAGATTACCACCTTAAAGAATTTGCCATC
ATCACCGATTCAAATTATGTCAGGAATAGTTTTGTTGAGTATCTTCCCCGATGGAAA
AGGAACAACATGGTAAAAAGTAACCATAAACCTGTCAAACATGGTAAGTTGTTTTG
TCAAATCGACCAATTGGTAAGGGAACACGAGATCACTCTGTATTTTAAGAAAACCA
AGGGCCACTCCAAAGTCCTAGGTGAAGAAAAGGAAGGGAATGATCAGGCTGATACC
CTAGCCAAAGAAGGGGCCCTAACAGGTGAGTTGCTAGATATAGATCATCTCTTAGGT
TCAGTACAGTTAGAAGCTGTCACTAGACAACAAGCTAGGATGACTGATGATTCAAA
TGTATTGCAATGGTGTCATGAAAACCCCAGTGAAGATTTAATCACAGCCCAAAAAG
AAGACCCAATAATTGGCTTATTTTATAACCATTTAGAACGCCCGGACCAGAATCCCC
TTAGTCAAGAAAATTATCCTGGCAAGGAAGACTTACGTATATTAATGAAATCCAAAT
CCCAGTTTACGCTGCATCAGGGATTGTTGATCCGCACCTCAAAAGGGGGTATCCAAC

AATGGGTAGTCCCACTTAAATTCAGAGGTTTAATGCTACAGCATGCCCATGACGCCC
CCACGTCTGGGCACCGCGGAGAAAAAATCACCTATGAGATCCTCCGGGATTACGCC
TACTGGCCTCATATGCTGCGTGATGTGAAGGCATACTGCCAAGGATGCTTAATATGT
CCTCAATTTCAACCAGCTTCGCCCACCCATAGGGCACCCCTTCAGAAAAGGGGGATA
TGCATGCCTTGGTCAGACCTTCAAATGGATTTCATTGGGCCTGTAACTAAATCTTCAC
GAGGGAATCGATATATGTTAACCGTGACATGTTTATTTACCAAATATGTTGARTGTA
TTCCAGCACCAAATAACAGCTCCGACACATGTGCCGCCCTTTTGATAAATCAcATCTT
TTCACGCTTCGGCCTTCCCCAGAGAATAGAGTCTGATAGGGGTACACATTTCACTAG
TGAAGTAATGACTAAACTGTGGAAGATATTGGGCGTAAAAAGGAAATTACACATCG
CCTATCGGCCCGCCTCTAGTGGTGGAGTGGAACGGTACAATCAGACAATTGTAAAC
ATACTCAAGAAATTTGTCAAGGAAACAGGCAAGGACTGGGATGTAAAGCTTCCCTT
AGTCCTGATGGCCTTAAGGGCAACCCCGAATACGGTCACAAAAATGTCACCCTTTGA
ACTCATGACCGGAAGGCGGATGGTCCTGCCCCAACACCTCTTGTACCGCACCACTGA
CCATAACTTGATCAATGCCACCACAACTCACCAATACATAGAAGATTTACGGAARC
ATTTACAATATGCGTTCTCATTTGCCCAAAGAAATATTGAAAGAGCCGCAACTAGCG
CTAAAACATACTATGACCTTAAAACCACCAAGAAGGAGTATGAAATTGGGGACAAG
GTATATCTCTACAATTTCACGAGAAACCAGGTAAAGGAACGGAAATTCCTGCCGTCC
TGGAGAGGTCCATTTAACATTGTGGACAAGATCTCCCCCGTGGCATATAAACTGCAT
ATTCCACGGAATGGTGAAATCAGTGAGAGATGGGTGCACATCAATCAATTACGTGC
TTGCCATCCCAAGTCCCAATTACAGTTGCTGGAGGACTAACCCTCGGCAACATATAC
CTTCAAGGTCAATCAAAAGGGTTAAATCACCTGAGAATAAAATTTTTTTATCACAT
CTCCACATGATACTCAAACTGTCTATTTGTTTTCTTTTGACAGTGTATACATCAGAGG
AATGATATCCCTGCCACCATCCAAACTCGCGATCAAAGATGATAGCGACCGGAAAT
ACCAAAGACGGAACCCGGCCATGAGACTACCCCAGGAAACTACCACATGGAAAGTC
CTTGGTATTCTGCTATGGATGTCCATACCAGGCCGGACCAGCCCCATCGCCCGACCA
GGACCTTCCTCTGGGATATTAGTCCAAGAAGCCCCCGGATTTATCTTAACAGACAAG
AGAATCACCACCCAAAAAGTATGTATAAGCGTGGATCCAAAGACCTCTATAGAAAG
ACTTTTTAACCCACAGAGGATCCAATCCCCTGAGGTCAAGGAATGGTACCCCATGCA
CCTGAAAGCAGCAAGAACTCGGGTGGACGAGATACTAACACAAGCTCGTAAGCCAT
TTATCCCGGCAACTATAACCCGTGAGAAGAGACCCAAGAGGTTCGTGGGAATCCTG
GTAGCTGGACTCATCTTCACCGTCCTCGGAGCTGTAGTATCCACCACCATATCTGCC
GCTAATTCCGTTTCCGTAAAGACTCTTACAACGGAAATTGAGAATTTACAAATGAAT
ATTCAAGCCATACACCAGGAACTTGAAACACAAAAAACTGAGTTCACCGATTTCCG
CACCATCATCAATGACACGATGGTCACTGTTAACTTACATTCAGAACTGATCACCCG
GTCGATCGACCTCCATCAGAGTCACGAACTATTCAAAAGGGAACTATTATTCATCCA
GGAACCAATTCTTTTCCACACATTAGGATATCTAAATGAAATACAGAATGGAATGAA
TGATCTCTCCCACGGACATATTCCCCTGTACTTTGTGTCCACTGAGATGATAAAAAC
CCTACTCACAAGgGTCAATGGGGGCGAGGTAGAAGATTTCCAAGTCAACCTGGCCTT
CGAAACAGGAAGcGCTATGCCGCTATATATCAATCCCGAAGATCTTGAAGTATGTTT
CTTGTTAGCCATACCCTTCGTCTCCTCTAAAAATATTTTCCAGATGAAAACAGTACAT
AATGTCGGAACATGGAGGGGTGATATACACGTCCAGATCAAAACCCCAGACCTGAT
TGCCTATCAACCATGGATGCCGGATTTATACGCTATACCCCTTTTAAAGGAATGCAC
ACGATTCAAGGACATCAATTTCCAATGTAACGGAAACCCATTCGTGTATGACACAAC
CCACGCCCTATGTGGCATGGAATATGAAAAACATGGTTCCGAAaAATGCGAAGTACA
AATGCGGAAGACTACAGATgCCTCAGAAGTACAGGAaATACTCGTGAAGGACGAAT
GGCTGGTGAGCACTcCGTTAAAAAACATGAGCATATTTATGAAAAGCAaATCATGA

GCCGTGTCATAGCTaTACCCGCTACCGTATCTATTATCAAAGTTCCAACCAATGCCAG
AATCACGATAGGGAAAACCACATTATATCCGATAGGAACCGAAGTCTGGGAAGGGG
AAGTTGAAAGTGTCGATGCCTTCCAACATCAAGACATACCAATGAACCCGGACCTC
AAATTCCTGTTAGAACATAAGCCTAACCACACCCTCCAGATGTCTATAAAAAAGAAC
AATCTGTCAATTGAAGCTTTTGAATATTCCGAGAAGGACCACATCTTATTAGAAAGC
AGATCATATTTTCCACTGACAGATATAATAATTTTAATCTGTATCGGACTGTTAATAC
TATGGTTAGTGGTTCTCACCATCcAATTGACAAAAGTATCCCGAGAAACCCAcGaACA
aTTACCAACGGTCTTGAGGCTCATTAGCGGAGGTTACCAACAACCTTAACTGTCAAT
ACCCTACACACCAAGTTACAAAAAAAAAAAAAAACTACAAACCAACTTCTGACTTTC
GCCACGTTATGAAGTGAATTCCCGGCACGTTTCCATGACATACCCGGAATTCAAAAG
GGGGATTTTGTCATAACGGAGATACCTCAAAATATCATATTTTGATTTTAAACCGTG
AAAAACCTGAATTAACACTCCCAAAGAACACTTCCCCTCCCAACAaAATTAAGTAAC
TAATTTATACaACCCCTGTTGTGAAGCAAGGAGGAAGATAGGAGAAAGCTCAGGAA
GAGCTGAGGGGGGCTCAGGGTACCCACCCTGATGGACACTCAGGGACACACCCTGA
GGGGGGACACTGGGAGATACACACACTCCCAGACAGACTGACACTCATGCACCATG
CATGAATACATATTTTGACATTCATGAAGATTCCTGAACCCTGATCCTTGGAACACG
GAGGTCGCAAGAAGTAAATCCTCTTTAAGAGTATCCAGTTTGAAAGCCACGGCAGC
CATCTTAACTCTGGTAACCAACAATCCAGACTAGTAGCCATCTTGGAATGGGACTAA
GATGTCCATGAGCCGGTCAACAGAACCGAGAGTCCAACGTGTGTGAAACATCCAGA
GAAGGAACCGTATGGGGGATTCAGGgACACCACTGATGGCCAGTCTACACGGACAA
ACCTACCAGGACGGCCAAGACCCAAAGGTGCCCGGAGTGATCCATTTCCCGCCAAC
CAACGGATGACGGCCCTGAGATGCACACATCAATCACCAACCCATTTTTTTTCCTCA
AAGAATTACACCCATATGAAGGTGTCTACACCCTTTTAAACATAGATCCACAATACG
GATCCATCTCCACAGCTAGATTTACAAGGCCCGAATGAAGCCATATTTCTTCAACAT
TTTGCATTGGTGAAGTCTTTCAAACGACTTCCGGCATTTGAACACTTCACAAGGGGG
GCTTTGTTGTGAATTCATGAAGAAATACATTTTATATTTACTGGCAGAAATATGAGA
GTTGCTTAGGCCTCAAAAAGTAGATGAACTGTGATGTTGCAATAATTGAACTGTTGC
CCTTTTGGAGACCTTTAGTACAAAATAGATGTTCTCATGTTTTTcCAAAGCTTTcTaCA
TCATCCAAATATTGAAGACTTTTACAAGCATGTATTGTGAGAGAAACTCGATGGAAC
AATGGCTGCTCTCGTCACAAAACAAGGAGATATTTACATGGTATTAGTGAATAAACA
AAACTACCCtTACAGCCATAGGAAACAAGGTTATACAGACTGTTATCACA
>Loki-Rca consensus sequence 5’LTR
GTTGTGAATTCATGAAGAAATACATTTTATATTTACTGGCAGAAATATGAGAGTTGC
TTAGGCCTCAAAAAGTAGATGAACTGTGATGTTGCAATAATTGAACTGTTGCCCTTT
TGGAGACCTTTAGGACAAAATAGATGTTCTCATGTTTTTACAAAGTTTTATRCATCAT
CCAATATTGAAGACTTTTACAAGCATGTATTGTGAGGAAAACTCGATGGAACAATG
GCTGCTCTCGTCACAAAACAAGGAGATATTTACATGGTATTAGTGAATAAACAAAA
ACTACCCTTACAGCCATAGGAAACAAGGTTATACAGACTATTCTCACA
>Loki-Rca consensus sequence 3’LTR
GTTGTGAATTCATGAAGAAATACATTTTATATTTACTGGCAGAAATATGAGAGTTGC
TTAGGCCTCAAAAAGTAGATGAACTGTGATGTTGCAATAATTGAACTGTTGCCCTTT
TGGAGACCTTTAGTACAAAATAGATGTTCTCATGTTTTTcCAAAGCTTTcTaCATCATC
CAAATATTGAAGACTTTTACAAGCATGTATTGTGAGAGAAACTCGATGGAACAATG
GCTGCTCTCGTCACAAAACAAGGAGATATTTACATGGTATTAGTGAATAAACAAAA
CTACCCtTACAGCCATAGGAAACAAGGTTATACAGACTGTTATCACA
>Loki-Rca consensus sequence *pol* gene

TTGCTGCGAACACCCCAAATTGTTCGCTGTTCGGCGAACTTGCAAACGGCCAATGTT
CGAGTCGAACATGAGTTCGACTCGAACTCGAAGCTCATCCCTAATCATAAGCACAA
AGAAGCTGTGACATGCGGTGAGATAAATGCCAACTGTAAATTATTGGAAAGCCCAC
AGATCCTCAACCACGCTAATTTCATGTGTGAGTTGACTGAATCTGATGGAAGATATT
TCACTACTGTAAAGGTACATGAGGCCGGTGATCACCCCATTTCAGGTTTAATTGACA
CGGGATCCCAAGCCACCATCCTGTCGCAGGAACAATTTAATGAAGCAATAAATTCA
ATGCCCAACAAACCAAAACTTAGGAGCTTTGATTGCTCACTAATCGGGGTGGGGGG
AGAGCCACTGCAAGTCATGGGTACTGCCTGGCTTAAACTCAAATTAGGTAATAAGG
AAATTAGGCACCCTATTATTATTGCAAATCTCCCACACCATCACATGATAATTGGAA
TGGATCTCCTTAAGAGGCTCAACACAATTGTTGACTGTGTGAATCAAGTAATGTGGT
CACAGGTCAACTCCCCGATTAAATATGAAAGACTTAAAACTCCCCATTCCCGGCGAC
ACTGCCACCTGGTGGAAGAAAGGCCTGGTTCCCTAGAAATCCACTTCAGGAATAGTC
AGTTCCCCGATGTTACGGTTATCCAACCTGACAATCATGGCCTGGAAGTAGGAGATG
AGGACCATATTTTTAAAATTCCCCGTGATAAAATAAAAGAGATCTCCCTCCAGGATA
ACGTACTGACCATATCCTTCCATCAGGAAAATAAAACTAAGAAAGGGGGAAATCAC
GCCCAGTTCCTGACAGAAATAGAAAGGGTTAGCAACGACCTATTTATCCCTATCCGG
ATAAATGGTCTTAATAGGTCAAAATATGCCAAAATTGATCTAAAGCAAGAGCATAG
CTTTGTGAGCCATGGCCTTGTTAGCCACCTTTCTGACATGGAAATGATCCGTTTCTCC
ACTCCTCAGCGACATTGGATGTCAGATGCCATTGATGACTCTGAGGACCATAACTCA
ATTGCTAGATGCCTTTTAAATATCACTATAGGTAATAAGACAACCACCCACGTGTTC
ATAGTGCTGGATAATCCCCAATCCCAAATTTACCTCGGCAACGACTTACTGCATCGT
TTTGCCACCCAAATTGACCTGATAAATGACTGCTTATGGAGTCGACTAAAGGGGGAC
CCCGAGGGATACCAGGATATGACCTCAGCCCTGAAGTTCCAGCAACAGATGCCTTA
TGCTGTTGAGATTCAGGTAGCTCGAGAAATTATCATCCCTAGGGACATAAAAAATTT
CCTACTCCCCATAAATGTAAAAAAGGGACAAAAACTGAAGGGATCAGATGCATTAG
TATGCCTGTCCCAGAGAATACAGGAGGAGGGAATAAAAATTACCCACGCACCTTTG
ATCAATACACATCAGTCTATAATATCCATCTACATCCACAACCTCTCAAATCAGGAC
CTTAACCTACCCAAGGGCACCCTTATCGGCCTTGCGGTAGAGACGGGGTATTATACG
TTTGGAGTCACCAACCAAGTGATTGGACTCATACCAGATGAATATTTATCTGAGGAG
AAATTAATTGAACAATCGTTCAATTCTTTACCCGAGGGTATRTTCACCATAGAGTCC
ATTTACCCCTTCGATTCTAAGGATGGTACCTGTAGAATTGAAGAGACATCTCTCGTC
TTTGATCAACCAATCAAAGAACAAGAGTACCCCAAGGGACCTGACTGGGAGAAAAA
TACACCTCGCACACCCAATGACCTGACCACCAGGCTTGAAGAGGCCTATGAGATAG
GGAAACCTGAAAGCTTTCCGGGATTCAAAGAGATAGTCGAAGAGATCATAGCCCTA
GCCGATGCATGCACCAATGATCTTGAGCGACAATTGCTCCGCGAACTTCTTTTTGAA
TTCCAGGAAATCTTTGCTCAGGATTCCTACGATTGTGGGAATACTGATCTCCACATC
GCCAAAATCCAGACCGACCCTCAGGCCCCACCAGTATTTGTAAAGCAGTATAGACTT
CCGCTAGCCTCCTATGAGTCTCTCGCCGAGATCATTAGGAAGTTACACGAACGGGGT
ATAATTAGGCCCATCCACAGCTCGTACAATCACCCCATCCTGGGGATTTGAAACCC
AACGGACAGTGGCGCTTGTGCGCTGATCTCCGACAACTCAATAAGCGTGTCTACATG
TCAGGATGGCCCGTACCCTACATAGACCAGTGCCTTGCCCAACTCCAGGGRTCCCAG
GTATTCTCGGCCCTCGACTGTGCCCAGGGATACTGGACCATAAAAGTCAGTGAAGA
AGATCAGTATAAACTGGCCTTCAGTTTTGGGAACCAACAGTACTGCTGGGTGCGTAT
GCCTTTCGGCTATATTAATTCTGGCCATGAATTTGCAGTCTTCATGCACAAGGCTATG
CCAGATGCCTTGGAGCGAGGAACATTGGCATACGTCGATGACATCCTCATCAAAAG
TACTGACATGACCAAACACCTCACAGAACTCCGACACATCTTGGGCCAGCTAAGGA

CRGCCGGAGTGAAGTTATCCATTCAGAAAGCACAATGGTGCCGCACCCGAGTAAAT
TTCTTAGGCCATGAAGTTACCTCAGAAGGTTTAAACCCCCAAAAGAAAAAGGTTGA
GGCAGTATTTAACTGTAAGGTCCCTACTAATCTTAAAGAATTAAGATCGTTTCTAGG
AATGGTTAACTATTCGCGTAAGTTCATCGACAATTACGCCGAGATTAGCAAACCGTT
ACTTCATCTGCTGAAGAAAGACACTGAATGGTCATGGGGCAAGGAACAAGCAGACG
CTGTGAAGGAACTGAAGGTACGACTCACCCAAGCCCCCTGCCTTGCGTATCCAGAA
GGAGGTAAACCCTTTTTTCTTGAGACAGGGTACACGCAGGTGAGTATGAGTTCCGTT
CTTTATCAAAAACAAGAGAATGTCAGTAAGATTATCGCCTATGCCAGTAAAACATTG
TCCTCAGTAGAGATTAAGTTTTCTGACTGCGAAAAAGCTTTATTGTCCACTGTCTGG
GCATTACAGAATTTCCGCAGTTACATTCAAGGTGAAAGAATTATTGTTGAGACCGCA
CACCAGCCTTTACAGTATTTACAGAGTGAACGTATCAGGGATGGAAATGTGTCTAAT
AGCCGCATCACTGCTTGGACTTTGTCCTTGCAAGGGTGGCCCCTGGAAATTCGATAC
AAACAGAACAAGAAAAGTCCAGTGGCCCAGGGCCTTGCTGAGCTGCATGATTGTTC
CGCCCCTGATCCCTGCGCGGGGACCACGGACACTGACTTTTTAGAAGAACAACTGTT
GTCTCCCTACAAGGTTTATGATGATGAATACTGTAAACAGTTGCCGTGTGTATATGT
AGATGGCTGTGCCTATCATCATAAGGACAAAGACCAGCTCCAACTAGTGGCTGGCA
TCGGGATCATATGGACCCAAGGTGTAGCCCCAATCTCTGTGGGATACAAGCTTGGGC
CAAGAAGCAGCCAAGTTGCTGAACTCGCTGCAATCCTGAAAGCCGTTGAGATGGCC
ATAGATTACCACCTTAAAGAATTTGCCATCATCACCGATTCAAATTATGTCAGGAAT
AGTTTTGTTGAGTATCTTCCCCGATGGAAAAGGAACAACATGGTAAAAAGTAACCAT
AAACCTGTCAAACATGGTAAGTTGTTTTGTCAAATCGACCAATTGGTAAGGGAACAC
GAGATCACTCTGTATTTTAAGAAAACCAAGGGCCACTCCAAAGTCCTAGGTGAAGA
AAAGGAAGGGAATGATCAGGCTGATACCCTAGCCAAAGAAGGGGCCCTAACAGGT
GAGTTGCTAGATATAGATCATCTCTTAGGTTCAGTACAGTTAGAAGCTGTCACTAGA
CAACAAGCTAGGATGACTGATGATTCAAATGTATTGCAATGGTGTCATGAAAACCCC
AGTGAAGATTTAATCACAGCCCAAAAAGAAGACCCAATAATTGGCTTATTTTATAAC
CATTTAGAACGCCCGGACCAGAATCCCCTTAGTCAAGAAAATTATCCTGGCAAGGA
AGACTTACGTATATTAATGAAATCCAAATCCCAGTTTACGCTGCATCAGGGATTGTT
GATCCGCACCTCAAAAGGGGGTATCCAACAATGGGTAGTCCCACTTAAATTCAGAG
GTTTAATGCTACAGCATGCCCATGACGCCCCCACGTCTGGGCACCGCGGAGAAAAA
ATCACCTATGAGATCCTCCGGGATTACGCCTACTGGCCTCATATGCTGCGTGATGTG
AAGGCATACTGCCAAGGATGCTTAATATGTCCTCAATTTCAACCAGCTTCGCCCACC
CATAGGGCACCCCTTCAGAAAAGGGGGATATGCATGCCTTGGTCAGACCTTCAAAT
GGATTTCATTGGGCCTGTAACTAAATCTTCACGAGGGAATCGATATATGTTAACCGT
GACATGTTTATTTACCAAATATGTTGARTGTATTCCAGCACCAAATAACAGCTCCGA
CACATGTGCCGCCCTTTTGATAAATCACATCTTTTCACGCTTCGGCCTTCCCCAGAGA
ATAGAGTCTGATAGGGGTACACATTTCACTAGTGAAGTAATGACTAAACTGTGGAA
GATATTGGGCGTAAAAAGGAAATTACACATCGCCTATCGGCCCGCCTCTAGTGGTGG
AGTGGAACGGTACAATCAGACAATTGTAAACATACTCAAGAAATTTGTCAAGGAAA
CAGGCAAGGACTGGGATGTAAAGCTTCCCTTAGTCCTGATGGCCTTAAGGGCAACCC
CGAATACGGTCACAAAAATGTCACCCTTTGAACTCATGACCGGAAGGCGGATGGTC
CTGCCCCAACACCTCTTGTACCGCACCACTGACCATAACTTGATCAATGCCACCACA
ACTCACCAATACATAGAAGATTTACGGAARCATTTACAATATGCGTTCTCATTTGCC
CAAAGAAATATTGAAAGAGCCGCAACTAGCGCTAAAACATACTATGACCTTAAAAC
CACCAAGAAGGAGTATGAAATTGGGGACAAGGTATATCTCTACAATTTCACGAGAA
ACCAGGTAAAGGAACGGAAATTCCTGCCGTCCTGGAGAGGTCCATTTAACATTGTG

GACAAGATCTCCCCCGTGGCATATAAACTGCATATTCCACGGAATGGTGAAATCAGT
GAGAGATGGGTGCACATCAATCAATTACGTGCTTGCCATCCCAAGTCCCAATTACAG
TTGCTGGAGGACTAA
>Loki-Rca consensus sequence Pol protein
LLRTPQIVRCSANLQTANVRVEHEFDSNSKLIPNHKHKEAVTCGEINANCKLLESPQILN
HANFMCELTESDGRYFTTVKVHEAGDHPISGLIDTGSQATILSQEQFNEAINSMPNKPKL
RSFDCSLIGVGGEPLQVMGTAWLKLKLGNKEIRHPIIIANLPHHHMIIGMDLLKRLNTIVD
CVNQVMWSQVNSPIKYERLKTPHSRRHCHLVEERPGSLEIHFRNSQFPDVTVIQPDNHG
LEVGDEDHIFKIPRDKIKEISLQDNVLTISFHQENKTKKGGNHAQFLTEIERVSNDLFIPIRI
NGLNRSKYAKIDLKQEHSFVSHGLVSHLSDMEMIRFSTPQRHWMSDAIDDSEDHNSIAR
CLLNITIGNKTTTHVFIVLDNPQSQIYLGNDLLHRFATQIDLINDCLWSRLKGDPEGYQD
MTSALKFQQQMPYAVEIQVAREIIIPRDIKNFLLPINVKKGQKLKGSDALVCLSQRIQEEG
IKITHAPLINTHQSIISIYIHNLSNQDLNLPKGTLIGLAVETGYYTFGVTNQVIGLIPDEYLS
EEKLIEQSFNSLPEG?FTIESIYPFDSKDGTCRIEETSLVFDQPIKEQEYPKGPDWEKNTPRT
PNDLTTRLEEAYEIGKPESFPGFKEIVEEIIALADACTNDLERQLLRELLFEFQEIFAQDSY
DCGNTDLHIAKIQTDPQAPPVFVKQYRLPLASYESLAEIIRKLHERGIIRPIHSSYNHPILGI
LKPNGQWRLCADLRQLNKRVYMSGWPVPYIDQCLAQLQ?SQVFSALDCAQGYWTIKV
SEEDQYKLAFSFGNQQYCWVRMPFGYINSGHEFAVFMHKAMPDALERGTLAYVDDILI
KSTDMTKHLTELRHILGQLR?AGVKLSIQKAQWCRTRVNFLGHEVTSEGLNPQKKKVE
AVFNCKVPTNLKELRSFLGMVNYSRKFIDNYAEISKPLLHLLKKDTEWSWGKEQADAV
KELKVRLTQAPCLAYPEGGKPFFLETGYTQVSMSSVLYQKQENVSKIIAYASKTLSSVEI
KFSDCEKALLSTVWALQNFRSYIQGERIIVETAHQPLQYLQSERIRDGNVSNSRITAWTLS
LQGWPLEIRYKQNKKSPVAQGLAELHDCSAPDPCAGTTDTDFLEEQLLSPYKVYDDEY
CKQLPCVYVDGCAYHHKDKDQLQLVAGIGIIWTQGVAPISVGYKLGPRSSQVAELAAIL
KAVEMAIDYHLKEFAIITDSNYVRNSFVEYLPRWKRNNMVKSNHKPVKHGKLFCQIDQ
LVREHEITLYFKKTKGHSKVLGEEKEGNDQADTLAKEGALTGELLDIDHLLGSVQLEAV
TRQQARMTDDSNVLQWCHENPSEDLITAQKEDPIIGLFYNHLERPDQNPLSQENYPGKE
DLRILMKSKSQFTLHQGLLIRTSKGGIQQWVVPLKFRGLMLQHAHDAPTSGHRGEKITY
EILRDYAYWPHMLRDVKAYCQGCLICPQFQPASPTHRAPLQKRGICMPWSDLQMDFIGP
VTKSSRGNRYMLTVTCLFTKYV?CIPAPNNSSDTCAALLINHIFSRFGLPQRIESDRGTHF
TSEVMTKLWKILGVKRKLHIAYRPASSGGVERYNQTIVNILKKFVKETGKDWDVKLPL
VLMALRATPNTVTKMSPFELMTGRRMVLPQHLLYRTTDHNLINATTTHQYIEDLR?HLQ
YAFSFAQRNIERAATSAKTYYDLKTTKKEYEIGDKVYLYNFTRNQVKERKFLPSWRGPF
NIVDKISPVAYKLHIPRNGEISERWVHINQLRACHPKSQLQLLED*
>Loki-Rca consensus sequence *env* gene
TTGTGGACAAGATCTCCCCCGTGGCATATAAACTGCATATTCCACGGAATGGTGAAA
TCAGTGAGAGATGGGTGCACATCAATCAATTACGTGCTTGCCATCCCAAGTCCCAAT
TACAGTTGCTGGAGGACTAACCCTCGGCAACATATACCTTCAAGGTCAATCAAAAG
GGTTAAATCACCTGAGAATAAAATTTTTTTATCACATCTCCACATGATACTCAAACT
GTCTATTTGTTTTCTTTTGACAGTGTATACATCAGAGGAATGATATCCCTGCCACCAT
CCAAACTCGCGATCAAAGATGATAGCGACCGGAAATACCAAAGACGGAACCCGGCC
ATGAGACTACCCCAGGAAACTACCACATGGAAAGTCCTTGGTATTCTGCTATGGATG
TCCATACCAGGCCGGACCAGCCCCATCGCCCGACCAGGACCTTCCTCTGGGATATTA
GTCCAAGAAGCCCCCGGATTTATCTTAACAGACAAGAGAATCACCACCCAAAAAGT
ATGTATAAGCGTGGATCCAAAGACCTCTATAGAAAGACTTTTTAACCCACAGAGGAT
CCAATCCCCTGAGGTCAAGGAATGGTACCCCATGCACCTGAAAGCAGCAAGAACTC

GGGTGGACGAGATACTAACACAAGCTCGTAAGCCATTTATCCCGGCAACTATAACC
CGTGAGAAGAGACCCAAGAGGTTCGTGGGAATCCTGGTAGCTGGACTCATCTTCAC
CGTCCTCGGAGCTGTAGTATCCACCACCATATCTGCCGCTAATTCCGTTTCCGTAAA
GACTCTTACAACGGAAATTGAGAATTTACAAATGAATATTCAAGCCATACACCAGG
AACTTGAAACACAAAAAACTGAGTTCACCGATTTCCGCACCATCATCAATGACACG
ATGGTCACTGTTAACTTACATTCAGAACTGATCACCCGGTCGATCGACCTCCATCAG
AGTCACGAACTATTCAAAAGGGAACTATTATTCATCCAGGAACCAATTCTTTTCCAC
ACATTAGGATATCTAAATGAAATACAGAATGGAATGAATGATCTCTCCCACGGACA
TATTCCCCTGTACTTTGTGTCCACTGAGATGATAAAAACCCTACTCACAAGGGTCAA
TGGGGGCGAGGTAGAAGATTTCCAAGTCAACCTGGCCTTCGAAACAGGAAGCGCTA
TGCCGCTATATATCAATCCCGAAGATCTTGAAGTATGTTTCTTGTTAGCCATACCCTT
CGTCTCCTCTAAAAATATTTTCCAGATGAAAACAGTACATAATGTCGGAACATGGAG
GGGTGATATACACGTCCAGATCAAAACCCCAGACCTGATTGCCTATCAACCATGGAT
GCCGGATTTATACGCTATACCCCTTTTAAAGGAATGCACACGATTCAAGGACATCAA
TTTCCAATGTAACGGAAACCCATTCGTGTATGACACAACCCACGCCCTATGTGGCAT
GGAATATGAAAAACATGGTTCCGAAAAATGCGAAGTACAAATGCGGAAGACTACAG
ATGCCTCAGAAGTACAGGAAATACTCGTGAAGGACGAATGGCTGGTGAGCACTCCG
TTAAAAAACATGAGCATATTTTATGAAAAGCAAATCATGAGCCGTGTCATAGCTATA
CCCGCTACCGTATCTATTATCAAAGTTCCAACCAATGCCAGAATCACGATAGGGAAA
ACCACATTATATCCGATAGGAACCGAAGTCTGGGAAGGGGAAGTTGAAAGTGTCGA
TGCCTTCCAACATCAAGACATACCAATGAACCCGGACCTCAAATTCCTGTTAGAACA
TAAGCCTAACCACACCCTCCAGATGTCTATAAAAAAGAACAATCTGTCAATTGAAGC
TTTTGAATATTCCGAGAAGGACCACATCTTATTAGAAAGCAGATCATATTTTCCACT
GACAGATATAATAATTTTAATCTGTATCGGACTGTTAATACTATGGTTAGTGGTTCTC
ACCATCCAATTGACAAAAGTATCCCGAGAAACCCACGAACAATTACCAACGGTCTT
GAGGCTCATTAGCGGAGGTTACCAACAACCTTAA
>Loki-Rca consensus sequence Env protein
LWTRSPPWHINCIFHGMVKSVRDGCTSINYVLAIPSPNYSCWRTNPRQHIPSRSIKRVKSP
ENKIFFITSPHDTQTVYLFSFDSVYIRGMISLPPSKLAIKDDSDRKYQRRNPAMRLPQETT
TWKVLGILLWMSIPGRTSPIARPGPSSGILVQEAPGFILTDKRITTQKVCISVDPKTSIERLF
NPQRIQSPEVKEWYPMHLKAARTRVDEILTQARKPFIPATITREKRPKRFVGILVAGLIFT
VLGAVVSTTISAANSVSVKTLTTEIENLQMNIQAIHQELETQKTEFTDFRTIINDTMVTVN
LHSELITRSIDLHQSHELFKRELLFIQEPILFHTLGYLNEIQNGMNDLSHGHIPLYFVSTEMI
KTLLTRVNGGEVEDFQVNLAFETGSAMPLYINPEDLEVCFLLAIPFVSSKNIFQMKTVHN
VGTWRGDIHVQIKTPDLIAYQPWMPDLYAIPLLKECTRFKDINFQCNGNPFVYDTTHAL
CGMEYEKHGSEKCEVQMRKTTDASEVQEILVKDEWLVSTPLKNMSIFYEKQIMSRVIAI
PATVSIIKVPTNARITIGKTTLYPIGTEVWEGEVESVDAFQHQDIPMNPDLKFLLEHKPNH
TLQMSIKKNNLSIEAFEYSEKDHILLESRSYFPLTDIIILICIGLLILWLVVLTIQLTKVSRET
HEQLPTVLRLISGGYQQP*
>Loki-Npa consensus sequence
TCAGATCGTAAATTGTGGGGTCTCAGACGCATATCAGTTTAGATGGTGCCAGTCTGT
GTAGGGCTTGGTCGTTTAAATTCTTTACCTCCCTATGAGAGATTAGAATTTCTAAAG
AGTTTTTTTTCTTTTCTTCTCTTTCTTTAATGAATAGACAAAGGGGGTTATTTTAAAG
TTCCTTATGAAACATTAATATTCATTGTGGTGATAATATTAGTAACAATGTTATGATA
ATAATATTATATCATATGTAGCATCTATCTGCCTACTTAATTGACTGATTTTTAAGTA
GAATTTTATACGTGTACTTTTATATATCCAACTATATTCATTATATATTGGTTGAGTG

```
TACCCATTGAAATATATTCATGTGTATTACTAATAATGGGGTTTTATATTTAAGTGC
AAAGAATTAGGTGGTACATTCTAATGTATATTCTTTTCTTATAAGGTATGAAAAACC
TCGGTCTTAAAACTTCTACCTCACTTTTAAGGAATTGAGAAAGAATGTTAATTGCTT
AAGTTAGACTGTTTTTCCAAGCCGTAAATTGCAGGACCAGACCGATAGTTTGGCATG
GTAGGATTAAGTATTGCCAAAAGTTTATATTTATGGGTCTTGTCCAGGGTCTTCGGTC
TAAGGTCATATCACTGCCTTATCTTCAATAATTAGTAAAATTAAATTAAATGCTCCC
AATTTTTACATAGAACAATAGCGCCCCCTGATGGTGATTGTGAGGGAAAGTCTATAA
CCATATGGGGCATTTGACGTCATGTTACACGTCAGTAATCACGTGTTAGATCAAACC
AAGCCCACTCCGCAGGGGTATATAGATCAGGAGCACAAGTGAAGAGGGAGAGTAA
AATTCAATCAAGAGAGAGAATCGAACATCAAGGCAGTAAGCAAATACATTCCAGAG
GGTATAGGTCAGAAGCTGAGGACGGCTGTCCTAGGAAACTGAATTAAAGAGAACCT
GACTATTCCTGGTGGCCGAGCATTGACTGCTGACTTCTTCCAAAACAACTAAATTTA
TATTGCTGGTAAAGTATACTTTATATCTGTATATTAGGCAATTCATCAGCCACATATA
TATATATTCGGGTAAACCATTAAGTTCTTGTAAATCGTGGCCAGAATTTCTTTATGTA
ATATGTGAATTGGCGGAAAAATTAAAAAAAGCAAAATATGAAACCAGGCGAAGGG
CTGATAATGTTAAAAAGTTACGAGAAGCCCTGCGTAAGATTGAAATTGAATCTGAA
AAGGTTTCTAATGAGATGGCACAATTGTCTGCCCAAATGTCGGGTTATAGTGATAGA
GATATTCAGAATAGAGATTTGATCAGTACTCTTAAAAGGGATCTGGATCAAATCTCT
AAAGAATCATCAGAACATGAAAGGTcAAAATTGGCTCTAGAAGAAGAATGTGAAAA
GTTGTTTACTGCCAATAGAAAACTACAGGAGGAGTTGGATGGCCAGAATCTGGCAT
ATCTGCAGTCTGTGCAGGAAAAACAGGCCAGCAGCGCAAAGGTTAAACCTTCTGGT
TGGGAGGGGGTAAATGTCACTGAAACTGTTAGCagcTCACACCCAGACTTTTGGTCTG
TTAAACCTAACATTGTTAGACGaACCATAAAAGAGAGGTTTTCAGATGTTGAAAGTG
ATTTCAGTGTCCCCTCAGAGGAAGAAGAACATTGTTTTACTAAGGAACGTTCTCTTC
CTAAATTAAAACCCATCTCACCTCATAAGCACTCCAAAAGTAAGTCCCATAGGAGTG
AGAGGAGAGTCTATCAGCCTGTGTTTGATACAACAAAAGTCATTAAGTTTTAAAAG
AAACTGTGGGTCTATTTACTAAGAAAGGATCGCCTTCTGTAATTAATCATATAGAAA
CTTATGAGGAGGCGATGATGACCCTAAACTTAAGAAGTGACAAACAGAAAATTGAA
TTCATTTCATGGGTGTTTGAGCCAAAGCACAGGTACTTCTTTACTTCCCTACGTGATA
TGGGTATAAATACCTGGTCTGAGGTCGTCTCAGAAATAAAGACAGAATTTGGTCCCT
ATCGGTCAGCTACTGCTGCCAGGAGAGCCATTTTTACTTTGAAATGTCGGCCGAATC
AGAGTCCCAGGGAATTTTTGTCAGTTCTCAAGAATGCCTATAGCCTGGCTGAGAAAG
ATCCCAATTTTGAGAGCCATGAATTCATCCAACTGTTTTATGATGCTTTGCCAAATAA
AATAAAAATTAGTCTTGCCCGGGATTACCATCCCAGGAAAACTCTGGATTGGCTACT
GACAGAGAGTACCACCCTTTACGCTGTTTTACAAAATAGCGATGAATCCCAGGATAG
GAGATTTAGGAAGCCAGCTGACGAGATAGCTGAGACTCATaTTAAAAGGGATCAGG
GACAGTTTGAAGCCCCAAAAAGGAGTTATGCTGAAGTGGTTAAGAACCCCAGTACT
CCTAAATTGAACAAGGCCAAACCCAATCCAGATAATAAGGTACCAAGTGTTGATGA
AGACCCCGGTGAAAAGCATCATTATGTAAAAAGGCAGTTCCCTCCTTTCAGGAGATA
TCAGAGAGGTTATTTCAGGAAGTTCCAGCATGGCCCAGATAAaGTCCAAAATGGTTC
TGAATCAGAAGGTTCTGGTAAATCGCAGGCTTCTAATTCCAAAAACCCACCTCCCAA
GGGTCATTGGAAAAATAGAAGCAATTTATATGATCGAATGGACCAAATGCAGAACC
AATTTGGTGCTTTGGCAGAACAGCTGGAAAAGATATCCAACGTCCTCTCTGCTAAAC
CCAAAAATCCTTTTTTAGGGGAGGCTACTCCGGGCGAAAGCCCTCCACAATAnCAAA
TGGGGATAAACAGGTTAAAGTGTCCAATCATGCAAATACCAATGATTTACCAAGAT
GCAAAAGTGAGCAAATTAACCCAAGTGTGTATGCCTGAAAATAATTTAACCCTTCCA
```

ATGCCAAATGGCAAGGAGGAGGGGCGACATGCTTGTGATTTTATCTTGCAGCCCACC
GAGTCTTCAGCCGTGCAAAGCCTGACAGGCACTCTTGCCTGTGAAGAAAAAGAAGG
GGGGTCGCGGGTCAAGATACACAGAGAGCAAAcAAGCAGCTTCACTTGTCATCCAA
AATATCTCTGTaCCTTACATGAGATAGAACACAGGTATTTCATTGAtGTGGAATTGCA
TAAGAATTTACCACAAATAATGGCTTTATTGGACACGGGTAGCCAAGTCACCATTTT
ATCTCAGAAATATTTTCAACAAATCCAGGATTTATCTTCTGAGAAGTATAAGCTGAA
AAGTTTTACAGCTAACCTTGTAAGTGTTTCAGGCACTGCCCTTAAGGTCATAGGAAA
AGCcTGGCTGAACTTCAACTTGGGGAATAAAGTAATTTGCCATCCCACACTTATAGT
TGATTTGCCCTATGATCGCTTAATTATTGGCATAGATCTCCTTAAGAGATTAAATCCT
ATTGTTGATTTTATCAATAAGGTTATCTGGTCCCAAGTGAGAATGCCCATTAACTAT
GAGCATTCCCAATCCTATCACAGTCGGCGTAACTGCCATGTGATTGAGGAAAGGCCG
GACTCTATTGAGATCCATTTCAGGAATAGCCAAGTTCCAGATGTCTCCTTCTTACAG
GCAACACCGACCTCCACCGTAGGTGGACAGGAGGATAACACCATATGTGTGTCCCC
TGAAAGAGTTAGAGATGTTTGCTTGGAAGGTGATGTTCTCACCATTCACCTCCATCC
AGAAAATACTGACTCTACTGGGTTCAGTGATCATGCgCAATTCCTAGCAGGGATTGA
AAAGGTGAATGATGATGTGTTCTTTCCCTTACAGGTTAATGACCTGGGAAGAACTAA
AAGGGCTAAACTGGATTTACGCCAGGAGAACAGTTACATTAGTAGGGACCTACTAA
AACAAATGGCCAATCCTAAAGTAATTAGGTATCCAAACCCTCAGGATAACTGGGTC
CATGATTTGGATGATGACTCaAACAGTCATAATATAATTGCTAAATGCATGTTATCCA
TTTCTATTGGGGGTAAATCAGCCACTCATTTATTCCTGGTACTCAATGAGCCACGGT
GTCATATTTACATTGGCAACGATATATTACATCGCTTTGCTATACATGTGGATTTGGT
GAATTCCACATTGTGGACTAGATTAAAAGGTAATCCATGTAGTTTCATGAGTGAACA
TGAAGCCCTCAAATCATCaCAGATACTGCCATATGCAGTAAGTGTATTAGTACCACA
AAATGTTATCATCCCTCCAGGTACTAATAATTTCCTGTTACCCATACAGGTATCAAA
AGGCCAAAAATTAAGGAATCCTGATGCATTAATATGCCTCTCAGACAGGATGCAAC
AATTAGGAATTTCAGTAAATCCTACACCCATGGTaAACATACATCCATTCAAAACCC
AAATAATTGTTAATAACaGTACGCCAAATGAAGTCCAATTGTCTAAAGACACCATTA
TTGGTTTGGCATTAGATTCAGATTATTATACTTTTGGGTTTCAAAATGTAATTATTGG
CCTAATACCTGAACCATATCTCACCGAGGAGCAAATGGTTGAACAGACATTTTATTC
TTCTCCTGAAGGATTGTTCACCATTAACTCAGTTTATCCTTTTAGTCCTGAAGAAGGT
ACCTGCAAGGTgGAGGAATCCTCCCTCACCTTTGATCATACACTTAATGAGCAGGAA
GCCGAGGAATATCGTGAGTGTACCGAAAAGGCTAGTAAGTCAAATCATGACCTCAC
TGACAGGCTTGAGGAAGCCTATGAAATAGGTCAACCTGAAGTTTTTCCAGGATTTCT
GGAAATGGTCCAGCAGCAAATATCTTTAGCCGATGGATGTTCTAGTGACCTAGAGCG
CCAACAATTGAAAGAAGTCCTACTGGAGTTCAAGGATATCTTTGCTAAAGATTCCTA
TGATTGTGGTCTAACAGACATACACATTGCTAGAATTCAGACTGATCCAAAGGCACC
TCCTGTTTATATCAAGCAATATCGTCTTCCACTAGCATGTTACGATTCGCTAGCGGAA
ATTATAAGGAATCTCAAGGAAAGAGGCGTTATAAGATCGGTGCACAGTTCTTATAAT
AATCCTATTCTTGGTATCCTTAAACCGAATGGTCAATGGCGTTTATGTGCTGATCTCA
GACAGCTCAATAAACGAGTATATATGTCCGGCTGGCCCGTGCCATATATTGACCAAT
GCTTGGTACAAATGCAAGGCGCCAGGATCTTTACCACATTGGATTGTGCTCAAGGGT
ATTGGACCATCAAGGTAAGTCCTGAAGACCAATATAAATTGGCATTTACATTTGGTA
AGGAGCAGTACACCTGGACCAGAATGCCGTTTGGTTACATAAACTCAGGTCATGAA
TTTGCTGTGTTCTTACATAAGGCTATGCCTGATGCAACTGAAAGAGGAAATTTAACA
TACGTGGATGATATTTTGTTGAAAAGTACCTCGTTTGAAAAACATGTaCAAGAGATT
AGGCATGTGCTTAATCAGTTGAGAGAGGCAGGTGTTAAAATCTCTTTACAGAAAGCC

CAATGGTGTCGTACCAAGGTCAACTTCTTGGGACATGAGGTTACTTCAGAAGGCCTC
AATCCCCAAAAGAAGAAAGTGGAGGCTGTGATGAAGTCTAAGACACCCACCAATGT
CAGTGAACTGAGGTCATTTTTGGGTATGATGAATTATTCACGTAAGTTTATTGATAA
TTATGCGGAAATTAGTAAACCATTGTTGACATTGTTAAAGAAGGATGTACCATGGAT
GTGGGGTGAGGATCAGGAACAAGCCATGAATGAGCTTAAAAGAAAACTCACCCAAG
CCCCATGcCTAACcTATCCAGAGGGGGGTAAACCCTTTTATCTGGAGACTGGTTTTAC
AAACCTCAGTATCAGTGCGGTACTGTGTCAAAAACAAGACAAATTGAACAAAGTTA
TTGCTTATGCAAGTAAaTCCCTGTCACCTGTAGAGGTAAAATTTAACGACTGTGATA
AAGCTTTATTAGCCACAGTCTGGGCCCTTCAAAACTTTAGAAGTTACATTCAGGGTG
AGAAAATCATCATTGAAACGGCCCACCAACCGCTACAGTATTTGCAAAGTGACAAA
ATCAGGGATGGTAATTTGTCCAACAGCAGAATAACTGCTTGGACGTTATCCTTACAA
GGATGGCCATTGGAAGTCCGGTATAAACAAAATAAAAAGAGTCCAGTAGCTCAAGG
ACTTGCCGAACTTCATGATTGTGCAGCTCCTTCTCcTAcAGAAGAGCTATTAGATGAT
TTCTTAGAGGAACAAATCTCTTCCCCTTACAAAGCATATGATGAGGAGTACTGTCCT
CTACTCCCCTGTGTATACATTGATGGTTGTGCTTATCACACCACCGTAGGGGATGAG
AGAAAACTAGTGGCTGGTATTGGAATCACTTGGACAAGTGAATTCCCAGATGTATCA
GTAGGATACAGTATTGGCGCTAAAAGTAGCCAAGTGGCAGAATTAGCAGCTGTATA
CAAGACTGTTCAAATTGCTCTAGAGCATAATATTACTGAATTTGTTATAGTCACAGA
TTCCAATTATGTACGGAACAGTTTTGTCGAGTATCTGCCTACTTGGAAAAGAAATAA
CATGTTACGGTCCAATAACAAACCTGTGAAACATGCCAAATTGTTTGGTGCCATAGA
TGAGTTAGTCAAAAGTAATGATCTAACCATTTATTGGAAGAAAACTAAAGGTCATTC
TAAAACCCAGAACATGGACAAGGAAGGTAATGATCTAGCTGACAAGCTTGCCAAAC
AAGGTGCCATAAATGGAGAAAGTGTGGACATAGATCATCTACTAGGAATGATACCC
ATTGAAGCTGTCACTAGACACCAAGCTAAGGTACAGACTACAAATAATGTGGCTCA
ATGGAGTCAAGAGTCTACAAGTGAGGATCTAATTAAAAGCCAAAAGGAAGATCCTA
TACTTGGGATCTTCTATAAACACATTGAAGATCCTAGTAACAATCCCATATCGGAGG
AGGATTATGCCCAAAAAGAAGACCTCCGATTACTAGTGAAAACCAAATCTCAATTT
AGTATtCGAGATGGTTTATTAATGAGAACTTCAAAGAATGGGATTCAACAGTGGGTA
GTTCCCACTATATTTAGAGGCCTAATGTTACAACATGCTCATGATGCACCTACAGCA
GGTCACCGTGGTGAGAAGATTACCTACGAATTGCTACGTGATTATGCTTATTGGCCT
CACATGTTGAAAGATGTCCAAACCTACTGTCAAGGTTGTCTGGTTTGTGCTCAGTTC
CAACCTCAAGGTCCTACACATCGTGCCCCACTTCAGAAAAAGGGTTTCTCATTGCCA
TGGTCCGATATACAAGTTGACTTTATAGGTCCTGTCACAAGGTCATCCAGAGGAAAC
AAATACATGTTAACGGTTACTTGTGTTTTTACCAAGTGGGTAGAGTGTTTGCCGGCT
CATAATAACACAGCCGAAACATGCGCCACTTTGCTTATTAACCATGTATTTTCCAGA
TACGGACTACCGTTAAGGATTGATACGGACCGAGGATCGCATTTCATAAGTGAGGT
AATGTCTAAGATGTGGAAgATCTTGGGTGTAAAACGGAAATTACACATTGCTTATCG
TCCAGCATCAAGTGGATCAGTAGAGCGCTATAACCAGTCAATTGTTAACATCCTTAA
GAAATTCGTTAAGGAATCAGGTAAAGACTGGGACATGAAGTTGCCTTTAGTGCTCAT
GGCTATCAGGGCCACTCCAAGTGCTGCTACTAAGATGTCACCCTTTGAGCTGATGAC
AGGTAGGAAAATGGTACTTCCACAACATCTGTTGTATCGCACATCGGATCATAACCT
GATCAATGCAGCCACGACACATCAGTATGTGGAGAACCTTCGTCAACATCTTCAACA
TGCTTTTGCTTTCGCACAGAAGAATCTCGAGAAAGCAGCCGTAAGTAATAAAACTTA
TTATGATCTCAAAACCACCAAGAAGGAGTATGAGATcGGTGAcAAGGTGTATCTTTA
CAATTTTGCTCGGGACCGAGTAAAAGAAAAGAAGTTTCTTCCGTCATGGAAAGGTCC
TTACACCATCATCGACAAAATCTCACCGGTGGTTTACAAAATAAAAATTGTTAAGAA

TGATGGTTTTACAGAAAAATGGGTTCATGTCAATCAATTGCGTGCTTGTCATCCTAG
GTCTCAATTACGAATACTGGATGGTGATGCCGAAGAAGAGACTTAATAAGTATCTCA
AGTGAATGTGATGATTGAATTGACATTGTGATATTTTCTCTCAGCTTACAGATCTTCA
TAGCCTGAGCTTCTCCGCATTCATTGGTCCTGCCAATCATAGCAGCATACCATGATC
CTGGAGATTGTTCATTAATGTTTAGCTAACTGAAGGATAATGAACCAGTCGCTATAT
CATATATATATAGCATAGCTAGAGAATGGAAATATCAATTGAAGAGAATATGTACT
CATATTATTTGTTTTTCTTAAATAGAAAGTGATTGACAGTCCAAACCAGGTCGTTGAT
TCGCAAAGACGATGATGACTACCAAAGACGCGGAGATCGTTTGGATGATCTGATCC
TTGGTGTCCAGTTCAAGACGGAACCCGTGTATCGAACCAGGTCCCTCAACTGGAAGT
CATGCTGCAGGATGATCCAGGATTCCTACTGACGGGAGAATATATCCTAGGAGATCC
TTGCGAGTTCGGATCCAAAGATATCCAGCGAAAGAAAATTCTACATCCCATTTTCAA
CTGGGAAAGGCATGGATGTTATGTTGTCCATGTGAAGACAAAGTTTTATATCCGGAG
ACCACTGATGAGTCCAATGATGTTCTGAAATGAACGATCAGTTTCATCAAGTGTCCT
AACTGCAAGCTTTGCTTTGATTGTGTTTGCTACACCCACTTTGGCTATTCAAGTCCTG
TTCATAAAgACATTATGAGCCAAGGGGGGAAATTCAACGGAAAACCTTTGAGgTTTT
ATAAAAAAAAAAAAAGAGGCCACATCTTTTGGCGAATATGACATTTCTTCTGAACTT
TTTCCCAGATATGTCTAAGGTTGTACACCTGAATACCTGTGAAGAAGGGGGGGgATT
aTGTCAGAACTATTGGGCCaAAAGGAATCAATTTCATAGGCCTAAACATGTTATTTTA
CAGATGTGTCTATGGGCATACTAGTAGCCATTTTGTCTATGTGCTAaAAGGTGAATAT
AAGTTAATTCTTCTTTAATAAGAGACTTTCACAGAGAAAGGTTAGGAATGTTAGTAT
AAACATTCTTCAGATCGTAAATTGTGGGGGTCTCAGACGCATATCAGTTAAAGATAG
TGCCAGCCTGTTTGGGCCTGGTTGATTACaTTCTTTGCCTCCCTACAAGAGATTGGAA
TTTCTAAAGAGTTTTTTTTTCCTCTTTSTTCAATGAATAGAaAAAGGGGCTTATTTAAA
AGTTCCTTATGAAACCTTACTATTCATTGTAgTAATAATTAtTAAcAATGTTGTtATTAT
ATTATTATATTATATGTAGCATCTATATGCCTACTTAATTGACTGATTTTAAGTAAAA
TTTTATACATGTACTTTTATATATAtCCAAACTATATTCATTATATATTGTTTGAGTGT
ACCCATTGAAATATATTGTTGTGTATTACTAATAATGGGGTTTTATATTTTAGTGCAA
AGAATTAGGTGGTAtATGCCAATGTATATTCTTTTCTTATAtGgTATGAAAAACCTCTG
GGTcTAAACTTaaAACTAAtTacCTCaCTTTAAaGgAaTtgGAaGAATGTTAATTGCTTTcGA
ATAGATTGTTTCAGagGTAAATtaCGAAGAGCAGACCTAAGGTCTGgcAGGgTGTAATt
AGGATTGTCCAAAATTTTATATTtATGGTggTGTCCAAGGTCTcCGGTCAAAGGTCATA
TCACTGCCTTATCTGTAGAAATTAGTGAAATTAACTGAAATGTTCCCAATTTTTACAT
AGAACAATAGTGCCCCCTGATGGTGATTGTAGGGAAAAGTCTATGACCATATAGGG
CATcTGACGTCATGCTACACGTCAgTAATCACGTGaTAGATCAAACCAAGCCCACTCC
GCAGGGGGgaGATATAGgccaGGagCCcGGgAGaAGCTAGGGAGAGAGGAAAGAggAgA
AATTCCAgAAgcaAGAGgGAtTCCAACAtCAAGGCCGTAAGCaAAtACaTTCCAAGGGTA
TAGGTCAGAAGCTGAGGACGGCTGTCCcAGGAAACTGAACTAAAGAGAACCTGACT
ATCCCTGGGGCTGAGCATTGACTGCTGACTTCTTCCAAAACAACTAAATTTATATTG
CTGGTAAAGTATACTTTcTATCTGTATATTAGGCAATTCATTAGCCACATATATATAT
ATTCGGGTAAACCATTGAGTTCTTGTA
>Loki-Npa consensus sequence 5’LTR
TCAGATCGTAAATTGTGGGGTCTCAGACGCATATCAGTTTAGATGGTGCCAGTCTGT
GTAGGGCTTGGTCGTTTAAATTCTTTACCTCCCTATGAGAGATTAGAATTTCTAAAG
AGTTTTTTTTTCTTTTCTTCTCTTTCTTTAATGAATAGACAAAGGGGGTTATTTTAAAG
TTCCTTATGAAACATTAATATTCATTGTGGTGATAATATTAGTAACAATGTTATGATA
ATAATATTATATCATATGTAGCATCTATCTGCCTACTTAATTGACTGATTTTTAAGTA

GAATTTTATACGTGTACTTTTATATATCCAACTATATTCATTATATATTGGTTGAGTG
TACCCATTGAAATATATTCATGTGTATTACTAATAATGGGGTTTTATATTTAAGTGC
AAAGAATTAGGTGGTACATTCTAATGTATATTCTTTTCTTATAAGGTATGAAAAACC
TCGGTCTTAAAACTTCTACCTCACTTTTAAGGAATTGAGAAAGAATGTTAATTGCTT
AAGTTAGACTGTTTTTCCAAGCCGTAAATTGCAGGACCAGACCGATAGTTTGGCATG
GTAGGATTAAGTATTGCCAAAAGTTTATATTTATGGGTCTTGTCCAGGGTCTTCGGTC
TAAGGTCATATCACTGCCTTATCTTCAATAATTAGTAAAATTAAATTAAATGCTCCC
AATTTTTACATAGAACAATAGCGCCCCCTGATGGTGATTGTGAGGGAAAGTCTATAA
CCATATGGGGCATTTGACGTCATGTTACACGTCAGTAATCACGTGTTAGATCAAACC
AAGCCCACTCCGCAGGGGTATATAGATCAGGAGCACAAGTGAAGAGGGAGAGTAA
AATTCAATCAAGAGAGAGAATCGAACATCAAGGCAGTAAGCAAATACATTCCAGAG
GGTATAGGTCAGAAGCTGAGGACGGCTGTCCTAGGAAACTGAATTAAAGAGAACCT
GACTATTCCTGGTGGCCGAGCATTGACTGCTGACTTCTTCCAAAACAACTAAATTTA
TATTGCTGGTAAAGTATACTTTATATCTGTATATTAGGCAATTCATCAGCCACATATA
TATATATTCGGGTAAACCATTAAGTTCTTGTA
>Loki-Npa consensus sequence 3’LTR
TCAGATCGTAAATTGTGGGGGTCTCAGACGCATATCAGTTAAAGATAGTGCCAGCCT
GTTTGGGCCTGGTTGATTACaTTCTTTGCCTCCCTACAAGAGATTGGAATTTCTAAAG
AGTTTTTTTTTCCTCTTTSTTCAATGAATAGAaAAAGGGGCTTATTTAAAAGTTCCTTA
TGAAACCTTACTATTCATTGTAgTAATAATTAtTAAcAATGTTGTtATTATATTATTATA
TTATATGTAGCATCTATATGCCTACTTAATTGACTGATTTTAAGTAAAATTTTATACA
TGTACTTTTATATATAtCCAAACTATATTCATTATATATTGTTTGAGTGTACCCATTGA
AATATATTGTTGTGTATTACTAATAATGGGGTTTTATATTTTAGTGCAAAGAATTAGG
TGGTAtATGCCAATGTATATTCTTTTCTTATAtGgTATGAAAAACCTCTGGGTcTAAAC
TTaaAACTAAtTacCTCaCTTTAAaGgAaTtgGAaGAATGTTAATTGCTTTcGAATAGATTG
TTTCAGagGTAAATtaCGAAGAGCAGACCTAAGGTCTGgcAGGgTGTAATtAGGATTGTC
CAAAATTTTATATTtATGGTggTGTCCAAGGTCTcCGGTCAAAGGTCATATCACTGCCT
TATCTGTAGAAATTAGTGAAATTAACTGAAATGTTCCCAATTTTTACATAGAACAAT
AGTGCCCCCTGATGGTGATTGTAGGGAAAAGTCTATGACCATATAGGGCATcTGACG
TCATGCTACACGTCAgTAATCACGTGaTAGATCAAACCAAGCCCACTCCGCAGGGGG
gaGATATAGgccaGGagCCcGGgAGaAGCTAGGGAGAGAGGAAAGAggAgAAATTCCAgA
AgcaAGAGgGAtTCCAACAtCAAGGCCGTAAGCaAAtACaTTCCAAGGGTATAGGTCAG
AAGCTGAGGACGGCTGTCCcAGGAAACTGAACTAAAGAGAACCTGACTATCCCTGG
GGCTGAGCATTGACTGCTGACTTCTTCCAAAACAACTAAATTTATATTGCTGGTAAA
GTATACTTTcTATCTGTATATTAGGCAATTCATTAGCCACATATATATATATTCGGGT
AAACCATTGAGTTCTTGTA
>Loki-Npa consensus sequence *smc-gag-pol* gene
TGTGAATTGGCGGAAAAATTAAAAAAAGCAAAATATGAAACCAGGCGAAGGGCTG
ATAATGTTAAAAAGTTACGAGAAGCCCTGCGTAAGATTGAAATTGAATCTGAAAAG
GTTTCTAATGAGATGGCACAATTGTCTGCCCAAATGTCGGGTTATAGTGATAGAGAT
ATTCAGAATAGAGATTTGATCAGTACTCTTAAAAGGGATCTGGATCAAATCTCTAAA
GAATCATCAGAACATGAAAGGTCAAAATTGGCTCTAGAAGAAGAATGTGAAAAGTT
GTTTACTGCCAATAGAAAACTACAGGAGGAGTTGGATGGCCAGAATCTGGCATATC
TGCAGTCTGTGCAGGAAAAACAGGCCAGCAGCGCAAAGGTTAAACCTTCTGGTTGG
GAGGGGGTAAATGTCACTGAAACTGTTAGCAGCTCACACCCAGACTTTTGGTCTGTT
AAACCTAACATTGTTAGACGAACCATAAAAGAGAGGTTTTCAGATGTTGAAAGTGA

TTTCAGTGTCCCCTCAGAGGAAGAAGAACATTGTTTTACTAAGGAACGTTCTCTTCC
TAAATTAAAACCCATCTCACCTCATAAGCACTCCAAAAGTAAGTCCCATAGGAGTGA
GAGGAGAGTCTATCAGCCTGTGTTTGATACAACAAAAGTCATTAAGTTTTTAAAAGA
AACTGTGGGTCTATTTACTAAGAAAGGATCGCCTTCTGTAATTAATCATATAGAAAC
TTATGAGGAGGCGATGATGACCCTAAACTTAAGAAGTGACAAACAGAAAATTGAAT
TCATTTCATGGGTGTTTGAGCCAAAGCACAGGTACTTCTTTACTTCCCTACGTGATAT
GGGTATAAATACCTGGTCTGAGGTCGTCTCAGAAATAAAGACAGAATTTGGTCCCTA
TCGGTCAGCTACTGCTGCCAGGAGAGCCATTTTTACTTTGAAATGTCGGCCGAATCA
GAGTCCCAGGGAATTTTTGTCAGTTCTCAAGAATGCCTATAGCCTGGCTGAGAAAGA
TCCCAATTTTGAGAGCCATGAATTCATCCAACTGTTTTATGATGCTTTGCCAAATAAA
ATAAAAATTAGTCTTGCCCGGGATTACCATCCCAGGAAAACTCTGGATTGGCTACTG
ACAGAGAGTACCACCCTTTACGCTGTTTTACAAAATAGCGATGAATCCCAGGATAGG
AGATTTAGGAAGCCAGCTGACGAGATAGCTGAGACTCATATTAAAAGGGATCAGGG
ACAGTTTGAAGCCCCAAAAAGGAGTTATGCTGAAGTGGTTAAGAACCCCAGTACTC
CTAAATTGAACAAGGCCAAACCCAATCCAGATAATAAGGTACCAAGTGTTGATGAA
GACCCCGGTGAAAAGCATCATTATGTAAAAAGGCAGTTCCCTCCTTTCAGGAGATAT
CAGAGAGGTTATTTCAGGAAGTTCCAGCATGGCCCAGATAAAGTCCAAAATGGTTCT
GAATCAGAAGGTTCTGGTAAATCGCAGGCTTCTAATTCCAAAAACCCACCTCCCAAG
GGTCATTGGAAAAATAGAAGCAATTTATATGATCGAATGGACCAAATGCAGAACCA
ATTTGGTGCTTTGGCAGAACAGCTGGAAAAGATATCCAACGTCCTCTCTGCTAAACC
CAAAAATCCTTTTTTAGGGGAGGCTACTCCGGGCGAAAGCCCTCCACAATANCAAAT
GGGGATAAACAGGTTAAAGTGTCCAATCATGCAAATACCAATGATTTACCAAGATG
CAAAAGTGAGCAAATTAACCCAAGTGTGTATGCCTGAAAATAATTTAACCCTTCCAA
TGCCAAATGGCAAGGAGGAGGGGCGACATGCTTGTGATTTTATCTTGCAGCCCACCG
AGTCTTCAGCCGTGCAAAGCCTGACAGGCACTCTTGCCTGTGAAGAAAAAGAAGGG
GGGTCGCGGGTCAAGATACACAGAGAGCAAACAAGCAGCTTCACTTGTCATCCAAA
ATATCTCTGTACCTTACATGAGATAGAACACAGGTATTTCATTGATGTGGAATTGCA
TAAGAATTTACCACAAATAATGGCTTTATTGGACACGGGTAGCCAAGTCACCATTTT
ATCTCAGAAATATTTTCAACAAATCCAGGATTTATCTTCTGAGAAGTATAAGCTGAA
AAGTTTTACAGCTAACCTTGTAAGTGTTTCAGGCACTGCCCTTAAGGTCATAGGAAA
AGCCTGGCTGAACTTCAACTTGGGGAATAAAGTAATTTGCCATCCCACACTTATAGT
TGATTTGCCCTATGATCGCTTAATTATTGGCATAGATCTCCTTAAGAGATTAAATCCT
ATTGTTGATTTTATCAATAAGGTTATCTGGTCCCAAGTGAGAATGCCCATTAACTAT
GAGCATTCCCAATCCTATCACAGTCGGCGTAACTGCCATGTGATTGAGGAAAGGCCG
GACTCTATTGAGATCCATTTCAGGAATAGCCAAGTTCCAGATGTCTCCTTCTTACAG
GCAACACCGACCTCCACCGTAGGTGGACAGGAGGATAACACCATATGTGTGTCCCC
TGAAAGAGTTAGAGATGTTTGCTTGGAAGGTGATGTTCTCACCATTCACCTCCATCC
AGAAAATACTGACTCTACTGGGTTCAGTGATCATGCGCAATTCCTAGCAGGGATTGA
AAAGGTGAATGATGATGTGTTCTTTCCCTTACAGGTTAATGACCTGGGAAGAACTAA
AAGGGCTAAACTGGATTTACGCCAGGAGAACAGTTACATTAGTAGGGACCTACTAA
AACAAATGGCCAATCCTAAAGTAATTAGGTATCCAAACCCTCAGGATAACTGGGTC
CATGATTTGGATGATGACTCAAACAGTCATAATATAATTGCTAAATGCATGTTATCC
ATTTCTATTGGGGGTAAATCAGCCACTCATTTATTCCTGGTACTCAATGAGCCACGG
TGTCATATTTACATTGGCAACGATATATTACATCGCTTTGCTATACATGTGGATTTGG
TGAATTCCACATTGTGGACTAGATTAAAAGGTAATCCATGTAGTTTCATGAGTGAAC
ATGAAGCCCTCAAATCATCACAGATACTGCCATATGCAGTAAGTGTATTAGTACCAC

AAAATGTTATCATCCCTCCAGGTACTAATAATTTCCTGTTACCCATACAGGTATCAA
AAGGCCAAAAATTAAGGAATCCTGATGCATTAATATGCCTCTCAGACAGGATGCAA
CAATTAGGAATTTCAGTAAATCCTACACCCATGGTAAACATACATCCATTCAAAACC
CAAATAATTGTTAATAACAGTACGCCAAATGAAGTCCAATTGTCTAAAGACACCATT
ATTGGTTTGGCATTAGATTCAGATTATTATACTTTTGGGTTTCAAAATGTAATTATTG
GCCTAATACCTGAACCATATCTCACCGAGGAGCAAATGGTTGAACAGACATTTTATT
CTTCTCCTGAAGGATTGTTCACCATTAACTCAGTTTATCCTTTTAGTCCTGAAGAAGG
TACCTGCAAGGTGGAGGAATCCTCCCTCACCTTTGATCATACACTTAATGAGCAGGA
AGCCGAGGAATATCGTGAGTGTACCGAAAAGGCTAGTAAGTCAAATCATGACCTCA
CTGACAGGCTTGAGGAAGCCTATGAAATAGGTCAACCTGAAGTTTTTCCAGGATTTC
TGGAAATGGTCCAGCAGCAAATATCTTTAGCCGATGGATGTTCTAGTGACCTAGAGC
GCCAACAATTGAAAGAAGTCCTACTGGAGTTCAAGGATATCTTTGCTAAAGATTCCT
ATGATTGTGGTCTAACAGACATACACATTGCTAGAATTCAGACTGATCCAAAGGCAC
CTCCTGTTTATATCAAGCAATATCGTCTTCCACTAGCATGTTACGATTCGCTAGCGGA
AATTATAAGGAATCTCAAGGAAAGAGGCGTTATAAGATCGGTGCACAGTTCTTATA
ATAATCCTATTCTTGGTATCCTTAAACCGAATGGTCAATGGCGTTTATGTGCTGATCT
CAGACAGCTCAATAAACGAGTATATATGTCCGGCTGGCCCGTGCCATATATTGACCA
ATGCTTGGTACAAATGCAAGGCGCCAGGATCTTTACCACATTGGATTGTGCTCAAGG
GTATTGGACCATCAAGGTAAGTCCTGAAGACCAATATAAATTGGCATTTACATTTGG
TAAGGAGCAGTACACCTGGACCAGAATGCCGTTTGGTTACATAAACTCAGGTCATG
AATTTGCTGTGTTCTTACATAAGGCTATGCCTGATGCAACTGAAAGAGGAAATTTAA
CATACGTGGATGATATTTTGTTGAAAAGTACCTCGTTTGAAAAACATGTACAAGAGA
TTAGGCATGTGCTTAATCAGTTGAGAGAGGCAGGTGTTAAAATCTCTTTACAGAAAG
CCCAATGGTGTCGTACCAAGGTCAACTTCTTGGGACATGAGGTTACTTCAGAAGGCC
TCAATCCCCAAAAGAAGAAAGTGGAGGCTGTGATGAAGTCTAAGACACCCACCAAT
GTCAGTGAACTGAGGTCATTTTTGGGTATGATGAATTATTCACGTAAGTTTATTGAT
AATTATGCGGAAATTAGTAAACCATTGTTGACATTGTTAAAGAAGGATGTACCATGG
ATGTGGGGTGAGGATCAGGAACAAGCCATGAATGAGCTTAAAAGAAAACTCACCCA
AGCCCCATGCCTAACCTATCCAGAGGGGGGTAAACCCTTTTATCTGGAGACTGGTTT
TACAAACCTCAGTATCAGTGCGGTACTGTGTCAAAAACAAGACAAATTGAACAAAG
TTATTGCTTATGCAAGTAAATCCCTGTCACCTGTAGAGGTAAAATTTAACGACTGTG
ATAAAGCTTTATTAGCCACAGTCTGGGCCCTTCAAAACTTTAGAAGTTACATTCAGG
GTGAGAAAATCATCATTGAAACGGCCCACCAACCGCTACAGTATTTGCAAAGTGAC
AAAATCAGGGATGGTAATTTGTCCAACAGCAGAATAACTGCTTGGACGTTATCCTTA
CAAGGATGGCCATTGGAAGTCCGGTATAAACAAAATAAAAAGAGTCCAGTAGCTCA
AGGACTTGCCGAACTTCATGATTGTGCAGCTCCTTCTCCTACAGAAGAGCTATTAGA
TGATTTCTTAGAGGAACAAATCTCTTCCCCTTACAAAGCATATGATGAGGAGTACTG
TCCTCTACTCCCCTGTGTATACATTGATGGTTGTGCTTATCACACCACCGTAGGGGAT
GAGAGAAAACTAGTGGCTGGTATTGGAATCACTTGGACAAGTGAATTCCCAGATGT
ATCAGTAGGATACAGTATTGGCGCTAAAAGTAGCCAAGTGGCAGAATTAGCAGCTG
TATACAAGACTGTTCAAATTGCTCTAGAGCATAATATTACTGAATTTGTTATAGTCA
CAGATTCCAATTATGTACGGAACAGTTTTGTCGAGTATCTGCCTACTTGGAAAAGAA
ATAACATGTTACGGTCCAATAACAAACCTGTGAAACATGCCAAATTGTTTGGTGCCA
TAGATGAGTTAGTCAAAAGTAATGATCTAACCATTTATTGGAAGAAAACTAAAGGT
CATTCTAAAACCCAGAACATGGACAAGGAAGGTAATGATCTAGCTGACAAGCTTGC
CAAACAAGGTGCCATAAATGGAGAAAGTGTGGACATAGATCATCTACTAGGAATGA

TACCCATTGAAGCTGTCACTAGACACCAAGCTAAGGTACAGACTACAAATAATGTG
GCTCAATGGAGTCAAGAGTCTACAAGTGAGGATCTAATTAAAAGCCAAAAGGAAGA
TCCTATACTTGGGATCTTCTATAAACACATTGAAGATCCTAGTAACAATCCCATATC
GGAGGAGGATTATGCCCAAAAAGAAGACCTCCGATTACTAGTGAAAACCAAATCTC
AATTTAGTATTCGAGATGGTTTATTAATGAGAACTTCAAAGAATGGGATTCAACAGT
GGGTAGTTCCCACTATATTTAGAGGCCTAATGTTACAACATGCTCATGATGCACCTA
CAGCAGGTCACCGTGGTGAGAAGATTACCTACGAATTGCTACGTGATTATGCTTATT
GGCCTCACATGTTGAAAGATGTCCAAACCTACTGTCAAGGTTGTCTGGTTTGTGCTC
AGTTCCAACCTCAAGGTCCTACACATCGTGCCCCACTTCAGAAAAAGGGTTTCTCAT
TGCCATGGTCCGATATACAAGTTGACTTTATAGGTCCTGTCACAAGGTCATCCAGAG
GAAACAAATACATGTTAACGGTTACTTGTGTTTTTACCAAGTGGGTAGAGTGTTTGC
CGGCTCATAATAACACAGCCGAAACATGCGCCACTTTGCTTATTAACCATGTATTTT
CCAGATACGGACTACCGTTAAGGATTGATACGGACCGAGGATCGCATTTCATAAGT
GAGGTAATGTCTAAGATGTGGAAGATCTTGGGTGTAAAACGGAAATTACACATTGC
TTATCGTCCAGCATCAAGTGGATCAGTAGAGCGCTATAACCAGTCAATTGTTAACAT
CCTTAAGAAATTCGTTAAGGAATCAGGTAAAGACTGGGACATGAAGTTGCCTTTAGT
GCTCATGGCTATCAGGGCCACTCCAAGTGCTGCTACTAAGATGTCACCCTTTGAGCT
GATGACAGGTAGGAAAATGGTACTTCCACAACATCTGTTGTATCGCACATCGGATCA
TAACCTGATCAATGCAGCCACGACACATCAGTATGTGGAGAACCTTCGTCAACATCT
TCAACATGCTTTTGCTTTCGCACAGAAGAATCTCGAGAAAGCAGCCGTAAGTAATAA
AACTTATTATGATCTCAAAACCACCAAGAAGGAGTATGAGATCGGTGACAAGGTGT
ATCTTTACAATTTTGCTCGGGACCGAGTAAAAGAAAAGAAGTTTCTTCCGTCATGGA
AAGGTCCTTACACCATCATCGACAAAATCTCACCGGTGGTTTACAAAATAAAAATTG
TTAAGAATGATGGTTTTACAGAAAAATGGGTTCATGTCAATCAATTGCGTGCTTGTC
ATCCTAGGTCTCAATTACGAATACTGGATGGTGATGCCGAAGAAGAGACTTAA
>Loki-Npa consensus sequence Smc-Gag-Pol polyprotein
CKSWPEFLYVICELAEKLKKAKYETRRRADNVKKLREALRKIEIESEKVSNEMAQLSAQ
MSGYSDRDIQNRDLISTLKRDLDQISKESSEHERSKLALEEECEKLFTANRKLQEELDGQ
NLAYLQSVQEKQASSAKVKPSGWEGVNVTETVSSSHPDFWSVKPNIVRRTIKERFSDVE
SDFSVPSEEEEHCFTKERSLPKLKPISPHKHSKSKSHRSERRVYQPVFDTTKVIKFLKETV
GLFTKKGSPSVINHIETYEEAMMTLNLRSDKQKIEFISWVFEPKHRYFFTSLRDMGINTW
SEVVSEIKTEFGPYRSATAARRAIFTLKCRPNQSPREFLSVLKNAYSLAEKDPNFESHEFI
QLFYDALPNKIKISLARDYHPRKTLDWLLTESTTLYAVLQNSDESQDRRFRKPADEIAET
HIKRDQGQFEAPKRSYAEVVKNPSTPKLNKAKPNPDNKVPSVDEDPGEKHHYVKRQFP
PFRRYQRGYFRKFQHGPDKVQNGSESEGSGKSQASNSKNPPPKGHWKNRSNLYDRMD
QMQNQFGALAEQLEKISNVLSAKPKNPFLGEATPGESPPQ?QMGINRLKCPIMQIPMIYQ
DAKVSKLTQVCMPENNLTLPMPNGKEEGRHACDFILQPTESSAVQSLTGTLACEEKEGG
SRVKIHREQTSSFTCHPKYLCTLHEIEHRYFIDVELHKNLPQIMALLDTGSQVTILSQKYF
QQIQDLSSEKYKLKSFTANLVSVSGTALKVIGKAWLNFNLGNKVICHPTLIVDLPYDRLII
GIDLLKRLNPIVDFINKVIWSQVRMPINYEHSQSYHSRRNCHVIEERPDSIEIHFRNSQVPD
VSFLQATPTSTVGGQEDNTICVSPERVRDVCLEGDVLTIHLHPENTDSTGFSDHAQFLAG
IEKVNDDVFFPLQVNDLGRTKRAKLDLRQENSYISRDLLKQMANPKVIRYPNPQDNWV
HDLDDDSNSHNIIAKCMLSISIGGKSATHLFLVLNEPRCHIYIGNDILHRFAIHVDLVNSTL
WTRLKGNPCSFMSEHEALKSSQILPYAVSVLVPQNVIIPPGTNNFLLPIQVSKGQKLRNP
DALICLSDRMQQLGISVNPTPMVNIHPFKTQIIVNNSTPNEVQLSKDTIIGLALDSDYYTF
GFQNVIIGLIPEPYLTEEQMVEQTFYSSPEGLFTINSVYPFSPEEGTCKVEESSLTFDHTLN

EQEAEEYRECTEKASKSNHDLTDRLEEAYEIGQPEVFPGFLEMVQQQISLADGCSSDLER
QQLKEVLLEFKDIFAKDSYDCGLTDIHIARIQTDPKAPPVYIKQYRLPLACYDSLAEIIRNL
KERGVIRSVHSSYNNPILGILKPNGQWRLCADLRQLNKRVYMSGWPVPYIDQCLVQMQ
GARIFTTLDCAQGYWTIKVSPEDQYKLAFTFGKEQYTWTRMPFGYINSGHEFAVFLHKA
MPDATERGNLTYVDDILLKSTSFEKHVQEIRHVLNQLREAGVKISLQKAQWCRTKVNFL
GHEVTSEGLNPQKKKVEAVMKSKTPTNVSELRSFLGMMNYSRKFIDNYAEISKPLLTLL
KKDVPWMWGEDQEQAMNELKRKLTQAPCLTYPEGGKPFYLETGFTNLSISAVLCQKQ
DKLNKVIAYASKSLSPVEVKFNDCDKALLATVWALQNFRSYIQGEKIIIETAHQPLQYLQ
SDKIRDGNLSNSRITAWTLSLQGWPLEVRYKQNKKSPVAQGLAELHDCAAPSPTEELLD
DFLEEQISSPYKAYDEEYCPLLPCVYIDGCAYHTTVGDERKLVAGIGITWTSEFPDVSVG
YSIGAKSSQVAELAAVYKTVQIALEHNITEFVIVTDSNYVRNSFVEYLPTWKRNNMLRS
NNKPVKHAKLFGAIDELVKSNDLTIYWKKTKGHSKTQNMDKEGNDLADKLAKQGAIN
GESVDIDHLLGMIPIEAVTRHQAKVQTTNNVAQWSQESTSEDLIKSQKEDPILGIFYKHIE
DPSNNPISEEDYAQKEDLRLLVKTKSQFSIRDGLLMRTSKNGIQQWVVPTIFRGLMLQH
AHDAPTAGHRGEKITYELLRDYAYWPHMLKDVQTYCQGCLVCAQFQPQGPTHRAPLQ
KKGFSLPWSDIQVDFIGPVTRSSRGNKYMLTVTCVFTKWVECLPAHNNTAETCATLLIN
HVFSRYGLPLRIDTDRGSHFISEVMSKMWKILGVKRKLHIAYRPASSGSVERYNQSIVNI
LKKFVKESGKDWDMKLPLVLMAIRATPSAATKMSPFELMTGRKMVLPQHLLYRTSDH
NLINAATTHQYVENLRQHLQHAFAFAQKNLEKAAVSNKTYYDLKTTKKEYEIGDKVYL
YNFARDRVKEKKFLPSWKGPYTIIDKISPVVYKIKIVKNDGFTEKWVHVNQLRACHPRS
QLRILDGDAEEET*
>Loki-Xla consensus sequence
ATGAGTGTCACAGGAAGTGATAATGGTAATGACCCCGAGCAGTTGACCCGcAACATT
GTTGAATTTTTGCAGCCACGGGTAAAATTTAGTCAAAGTGTGGACGATATGATTTAC
GATTTGAAAAAGTATGTGAAGGAGGATATTGATGATATCTTTGCAATCAATGGTGTG
GTTGAACGTTATTTAACAAGTATGACCTTAGCAACTAACAGTGCAGACAAAGAGAA
AAGATTAATGTGTGCCCTACCTCCTGTTTTGTATGCTTTAATAGAGGTAGATACCCTA
TCAAAAGCTAGGAAAGTAGAGATAGAAAAACTGAGATTTCAAATACATGAGACAGA
ATCAAATGTCCAATTTTTGACATCTCAATTACATGAGATAGAATCAAATATCCAACA
TATGAAcTCTGAGAGAAACCAATTAGAGATAGATATCTCAAACAAGGAAGCCTTTAT
TAGGGACTTAGAAACTGATGCTGAACATAGTCGCATAGCATACCGTAAGCTCAGAG
AGGATTTTGAAGAATTACGGCAACAGTTGTTTAAAAATCAACAGTCAGGTGTATATG
GTGTGGAGAGATCTCCTGATCCTCCAATAATCACTCCCCAATATAAAAATTGGGCAC
CATATGCTCCTAGAGCCAAACCAATTAgTAATGGTGATTCAGAGGCACAAGaTGTTTT
GaGCCGGTTGaCTGAAGCTTGCCAAGCTGTGACCACCTTATTAGGCAATCACAGAAC
CAGGACCCCTATAGACTCCTGTCAAGACTCATCCCCAGAGAGGAATGAGCGACGTA
AATCCCTTGTGTCCAATGTATCCTATAAAAGTTATGCTTCCAATCTTAATACAAGTGG
GTCTGAGAGTGATGAGAGGAAGCGGGATTCCAGTCCTGTCCCAAGAAGGGGGAAAA
ACACCCATAAAAGGGGTTCAAAAGGTAAAAATCCCCAATATCCAGGTGAAAAACAA
AATTCTGCCAGCATAATTAAATTCTTGAATACTGCTGTGCCAAAATTCTCTAAGAAA
GGTTCTgCCCAAATAGCGACCCATTTAGAAATGTATGAGTCTACTATGGACTCTTTGG
ACTTGTCTGAAGCAGACAGAATCAGGTTTCTTCCATGGGCCTTTGATGACAGATACC
GTCATTACTTTTCCTCCTTTAGAGAGAGAGGAATCACCAAATGGCAGTCGGTCCTAC
ATGAAGTTAAATTAGAATTTGGGCCATACCGAAATACCACTGCCGCAAAGAGAGAA
ATTTATAGACTTACATGTCGGCCTAACCAAAGCCCCCGAGAATTTCTCTCTGTGCTC
AAAAATGCCTATGGTTTGGCCTACAGAAACCCTGATTGGGAATCTATAGATTTTAAA

CAGTTGTTTTATGAGGCCTTACCAACCCAGATCAAATTAAGTCTTGCTCATGATCTG
GATTTTGAAAACCCTCTAGAAAGACTGGTAACATCAGCAACcTTACTGTTTAACATC
AGTGAGGGCCAAAACTTAAATGAAAGGAAAATTAAAAAATTCCCAGAACACAATGT
CGATGAGTCCAGGGTAAAACCTCATTTAAATTTTGAGTCACAGCCAAAGAAAGTAC
CTTCAGCCAATGGGTCTAATCAGCAGAACCGAAGACAACAGCCAAATCCCATACAA
AACAAACCTCAAAATCCTAAGGGATCAGACGGGAACAGCTCAGGTTCTGGAAGCCA
GGACTACCGTCCTCGTTTCAACAAATTGAGCGACATTGTAAGTAAATTGTCTCAAGT
TTCTGCACCAAAACAGTCTGAGGATTTTTTtAGATGTGAAGGCGGGACCCAGTTCCG
CCAATTAAAAGAAACAGGGTCAGaTCCCTGTACGGAAATGGATACGGCAATACCAA
ATATGGACAACGATCATATTTTGCATTGCACCGACCCAGGATTCAGCGGGGTCAAAG
GGAAGTCTAATATTTCATTTATCTTACAGCCAAATAAGGACAATTCTCGTGATCCAA
TGTCACTACAGCCCGTGTGTGAAATACCCTCTGTTCAGTTcGTCCCAgTGGTGGAATC
CATGGGAACCCAGATGACCTGCAGGCAGAACCAGGAGGAGATAAATCCAAATATCC
TGACGCTGCAGAAGGATCATTCTTTAAAACAACAGATGGCAAAACACTCCAATTTTC
TCTGTAATCTCTTACAGTTAGACGGCCGGTACTTTATAGATACCTGTCTCCAAGATG
GATGTGTCGGACCAATCCAAGGTTTATTGGATACCGGGTCTCAGGTCACTATCTTAT
CTTTTAGCTACTTCCAACAGATCTTGGGGTCATCAAAAAAGAAACCGAGATTAACGG
CATTTGATGGCTCTTTGATAAGTGTGGGCGGTGACAGTTTACAGGTGAAAGGCACAT
CCTGGCTGACATACAAGATTGGCAAAAAGATCCTTAGACATCCAACATTGATTGTAG
AcCTcCCATATGATCGTtTGATCATTGGgATcGATtTGCTCAAGcGATTAAATTCCATAA
TTGACTTTATAAATGAGGCAATTTGGACcCAaGTGAAAATTCCAATTGCTTTTGAGCG
CTCCCACAATGCACAATCGCAACATAGTTGCCACATGATTGAAGAAAGGCCTAATTC
TGTTGAAATCCATTTCCGGAACAGTCAAATACCAAATATATCAATCCTGAAAAACGG
TAAACCTGAAGAAAGTGTTAATACTTCCTTCCACATCATTAATATCCAAAAGGACAA
AATTGAAAAGATAATTTTGGAGGATGATGTACTAACCATTTCTCTTAAAGGGGGAAA
TCAGGAAGTTGCAGGCAAATGTGATTGCAAGGTGCATTTTATCCATCTCCCTAGGAA
ACAAATCTACAGAACACGtTTTCAGTGTGCTAAAGGAACCACAATATCAAATGTATT
TAGGAAATtACATTATACATCGATTTGCCATACAGGTTGTCTGATCAATAATGTTCTG
TGGTCCAGATTACCAGGGAATCCAGAGGAATTCCAGGATAAAGAGCAAACCCTGAG
ATGGGGaCAGCGAACACCCTATGCAGTGGATATCTTAGTCGCCGAAAAGGTAAAAA
TCCCTGCGGGATGTAAATCATTTCTTTTACCCATTCGAGTGAAGAAGGGGCAAAAGC
TAAGAAATGCAGATGCATTATTTTGCTTATCCAATCGGATACAGCAATTTGGATTGA
AGGTAAAACCGAATCCCATGATAAACATTCATCAGGATCCATTATACATGGTTATCC
ATAACATGAGTCCTTATAGCATTACGTTACAAGAGGACACAATGATTGGCCTAGCCA
TTGATTCAGAATATTACACTTTTGGGTTTCAAAATGATATAATTGGATTGATACCTGA
TGAATATCTAACAGAAGAACAAGTAGTGGATCAACATTGCTTCTCAACACCTGAGG
GATTGTTCAGTGTTCACTCTGTTTATCCTTTTAGTCCTGAAGAAGGTACATGCTACAT
cGACGAGACATCATTGATCTTCAACCATGAGATCAATAAAGGTGAGTATGAGTCATT
CCAACAGGGGAAGGATCAGATATaTCACACCCCGGATGACCTGACCAGCAAGCTTG
AAGAGGCTTATGAGATAAaTCAACCcGAGATTTTcCCAGAATTTCAGGAgCGGGTGGA
AGAgCAACTCTCTATaGCaGATGCATGTtcTGATGAATCAGAaCGACAACAGcTAAaGG
AACTATTCTgGGAATTCCAGGAtATGTTCGCCAAGGACTCTTATGACTGTGGGAAAC
TAACCTGCATGTTACAAGAATCCAGACCGATCCAGATGCACCCCCTGTGTTTGTTAA
ACAATACCGACTTCCGTTAGCGGCTTATGAGTCACTCTCGGAAATCGTTAAAAATCT
GGAGAAAAAGGGTATCATCCGTCCAGTACACAGTTCTTTTAATCATCCAATACTTGG
AGTCCTTAAACCCAATGGTCAGTTTCGTCTTTGCTCAGACTTAAGACAGTTAAACAA

ACGAGTCTACATGTCTGGTTGGCCCGTACCATACATTGACCAATGCTTGGCGCAAAT
CCAAGGATCTAAAATTTTCACaGCCCTTGATTGCGCACAAGGATATTGGACAATTAA
AATAGATGAAAGGGATCAACACAAACTGGCCTTTACATTTGGAAAACAACAATATG
CCTGGACCCGATTACCTTTCGGGTATATCAATGCAGGCCATGAGTTTGCTGTATTCAT
GCACAAAGCAATGCCTGATGCTGCTGAACGAGGTACATTATCTTATGTGGATGACAT
ATTGATAAAAAGTACAACTTTTGAAGAGCATATTGCAGAACTAAAATATGTATTAAA
ACTTTCTTTACAGAAGGCCCAATGGTGCCGATCTAAGGTAAATTTCCTAGGACATGA
AGTTACTGCcGAAGGAATCAAtCCACAGAAAAAGAAAGTGGAAGCAATAAAGAACA
CGAAATCTCCTACAAATCTCAAGGAGCTGAGATCATTCCTGGGTATGATGAACTATT
CACGTAAGTTCATAGACAACTATGCgGAAATTACAAAACCTCTTTTGCAGTTGCTGA
AGAAAGGAGCAAAATGGGAATGGAATGAGTGTCACGAACAAGCAGTATCTGAACTT
AAAACTAAGCTTATCCAGGCTCCATGCCTTGCTTATCCAGAAGGTGGAAAACCCTTC
TATGTCGAAACAGGTTTTACCGACAAAAGTGTaAGTGCGGTTCTTTTCCAAAAACAA
GAAAAACTGAATAAAATAATTGCTTATGCCAGCAAGTCATTATCaCCAGTTGAAATa
AAATTTAATAACTGTGAAAAAGCGTTATTAGCaACTGTATGGGCTTTGCAATATTTCA
GGAGCTTTATTCAGGGTGAAAAGATTATTGTGGAAACTGCACATCAATCACTACAAT
ATTTACAGAGTGGTCGGTTAAAGGATGGAAACCTGTCAAACAGTAGGATAACTGCA
TGGACAATGTCCCTTATGGGATGGCCATTGGAGATCAGATATAAACAGGATACTAA
AAACccAGTTGCTCAAGGATTAGCTGAACTCCATGATTGtAtTAATGTcGAACATAAaG
ATGAGTCACCAGAAAATGACTTTcTGGAGGAACAATCAtTATCTCCATATAAaTCATA
TGAAGAGGAATACTGCAAATCATTACCATGTGTATATGTTGATGGCTGTTCTTTCCA
CACAGAAGAAGCTGAAAAGATGCTGGTTGCAGGTATCGGGATCGTGTGGAATAATA
TATTCCCAGAGATATCCATAGGATACAAaATTGGTCCCAAAAGTAGCCAGGTCGCAG
AGCTTGCCGCTGTGTACAAAGCAaTACAAATGGCTATTGAATATGATCTTAAAGAAT
TTGTTTTAATTACAGATTCcGATTATGTTCGGAACAGtTTTGTGGAATAcTTTCCCGGA
TGGAAAAGATCCAACATGtTGAGAAGTAACAAGAAACCAGTCAAACATGGTAAGCT
GTTCTGCAAAATCGATGAATTGGTTACTACCTACGGTCTTACTATCTATTGGAAGAA
GGTAAGGGGTCATTCAAAGCATCCAGGCGCAGATAAAGATGGGAACGACCTAGCGG
ATTCCCTCGCCAAGAAAGCAGCCATTGATGGAGAGATCTTAGATATCGATGACCTCA
TGGGAACTATCCAAGTGGATGTCACAACCAGGGGACAGGCCCAAAAGGAGTCTCAA
CCTAATGTGGCAATGTGGAGCCAAGATTCCCCTAGCGAAGATCTCATTGAGAGCCA
GAAAGGGGATCCTATTATTGGTTTATTTTATAAATATATTCAAAATCCACAaGATCAT
CCCATAACTATGGAAGACTGTGCTGATAATGAGGAGTTACGCGTTTTAATGAAAGGT
GAGTCCCAATTCTCATTACAGGATGGATTGTTGATACGAACCTCGAAAAACGGTATC
ACACAATGGGTAGTACCCGCAGCATATCGGGGTTTGATGTTACAACATGCCCATGAT
GCCCCGACAGCTGGTCACCGAGGAGAGAAAATAACATATGAGTTGTTGAGGGACTA
TGCTTACTGGCCACACATGTTACAAGATGTCAGGACATATTGTCAAGGATGTTTGGT
ATGTCCTCAGTTCCAACCACAAGGACCCAATCATCGGGCACCACTGATGAAAAGAG
GAATATCCATGCCATGGTCAGATATTCAGATTGATTTCATTGGACCAGTAACAACTT
CCTCAAAAGGAAACCGATACATGCTAACTGTAACTTGTTTGTTCACAAAATGGGTAG
AATGTCTGCCATGTAAAACATGCAGTTCTAGTGTATGCGCATCACTGCTCATTAACC
ATATATTTTCCAGATTTGGTCTTCCCCAGCGTATTGAATCCGACCGTGGAAGTCACTT
CACCAGTGAAGTgATGACCAAGATGTGGGAAATTCTAGGAGTAAAAAGGAAGCTAC
ACATAGCTTACCGACCAGCTTCTAGTGGTGGAGTAGAGCGTTACAATCAATCCATAG
TAAACATTTTGAAAAAGTTTGTCAAGGAATCTGGTAAAGACTGGGATGTCAAGTTGC
CATTGGTTCTCATGGCTATTAGAGCCACTCCAAGTGCAGCAACCAAACTTTCTCCAT

TTGAGATGATTACAGGTAGAAAAATGGTGTTACCTCAACATTTACTTTACAGGACCA
CAGATCACAATTTGATGAATGCAACaACTGCACATCAATATgTGGAAAAcTTACGcAA
GCATCTACAACATGCTTTTGCATTTGCCCAAAGAAATATTGAaAAAGCTGCAGTAAG
TACGAAAACGTACTATGATCTAAAAACAACTCAAAGGGAATATCAGGTGGATGACA
AAGTTTATCTTTATAACTTTGCCAGgGATCAGGTAAAGGAAAGGAAATTCCTACCTT
CATGGAAAGGTCCATATATAATCACTGATAAATTATCTCCTGTGGTTTATAAAGTTA
AAATTCCTAAGGGGGATGATTTTGTTGAAAAATGGGTGCATATTAACCAATTACGTG
TTTGTCATCCTAGCTCACAGCTTCGGAAAATAGAGCAAGTCGAGAGCTGAGTGGTAA
AGATCTGGAAGTTCCGGTTAAAGTGCAAAGATGCCATCCAAGATGATCGCTATCTTA
TATGCAGTCCTGATCTTGGGGAAAGAGTTCCATGCTGAGCCGATTGTGATGCCAGGT
CCATCATCCGGGATCATGCTACAGGATACGGCAGGATTCATCATGACCAACAAGAG
GATCCTATCCCAGAAGGTTTATATGAGCCTAGATCCACGTTGTGTCATCGAAAGACA
AGTCAACGTATCCAATATTCGATCACCGGAGATTCAAACCTGGTACCAGCTACATAT
CGAATATTCACAAGAGCGGGTTACTCAGATCCTGGAGCAAACCCGGAAAACTTTGA
CCAGAGAGCCCTTTTCTTCAAACCAAAGACCAAAGCGGTTCATTTCTGCCATTGTTG
CTGCTATAATCTTTGCGGTAGTAGGAACTGTGTTTGCAACCGGTGTAACAATCGCTA
ATTCTGTATCCCTTAAAACCATGGAATCTGAGATTAGTTCATTAAAGACTAATTTAC
AACATATTCATGCAGAGATGGAAAGACAAAAGAAAGATTTAACAGACTTGTATGCC
GTTGTGGAAGACACTGTCGTTACCACGAACTTACACTTTAAATTGTTAGACCGCTCA
ATAGATCTACACCAAGGTCATGAGCAGTTTAAGCAAGAACTTATTTTGTTAAAAGAA
CGTATCGACTTTCAACTTCAGGTTTTCCATGACGAGGAGCAAAGTGGGATGCAAGAC
CTAAGGGAAGATGTTACAAAGTGTCGATGGAGAAaAGTAGAGGATATCCAAATGAA
CTTAGCATTTAATATGGGAGTTGCCATACCATTATACATCGATCCCATTAATATGGA
AATTTGCTTTTTACTATCAATCCCAATGGTTAATAGGAACAATATATTTCAAATGAA
GAAAATCCACAATGTAGGTACGTGGAgAGACAATATTCACATTAAAATTCAAACTCC
TGACATTGTAGCATATCAATCTTGGAAACCGGACTGGCTAGCTGTTCCTATTCTGGA
TAATTGTGCAACATTCCGGGATAATAATTTTCAATGTGAAGGCGATCCGTTCAATTA
TGATGCAACTAACGCTTTGTGCGGACTTGAATTTGAGAAACAtGGATCAGAGAAATG
CCAGGaGATCTTGTCTAAAATTGAGGAGAAAGACTGGGTAAATGCAATTTTAATAAA
AGACAAATGGCTCGTTAGTACTCGTCTGCCAGACCTGACTGTTTTTCTAACGACCA
CATTATGGCGAAAGTTATTGCTTTACCTTCCAGTGTATCGATCATCAAGGTACCCCCT
AATTCCAGAGTTGTGATTGGAAATCGCACCCTATATCCTGTAAATATTGAAGACATT
GAAGGAGAGGTGGAAAGCATTGATGCATTCCAGAATCAGGAAATCACTATAAGTCC
GGATTTAAAATTTTTGTTNGACAATAAGCCCAGACATACCGTTCGTATGGTTATTAA
AAAGCATAATGTTACCATGGATAccCCACAATACTCATCAACCGACGACGAGATATT
GTTTCCCTCTGTtTGGTCATGGGGGGAAAATTACATATGGAGCATTTTGTTCGCAGTT
ATAGTAGGGTGGTTAACTGTAATCACCGTCTTGGCAATCCGTTGGATGAAGTTCATC
AAAAGAACTCTAGATAAGCTATGGCTTAAATCTCGTCGATTGAGTCTACATAATGAG
TATGATGAAGGAAGAGTTTGA
>Loki-Xla consensus sequence *smc-gag-pol* gene
ATGAGTGTCACAGGAAGTGATAATGGTAATGACCCCGAGCAGTTGACCCGCAACAT
TGTTGAATTTTTGCAGCCACGGGTAAAATTTAGTCAAAGTGTGGACGATATGATTTA
CGATTTGAAAAAGTATGTGAAGGAGGATATTGATGATATCTTTGCAATCAATGGTGT
GGTTGAACGTTATTTAACAAGTATGACCTTAGCAACTAACAGTGCAGACAAAGAGA
AAAGATTAATGTGTGCCCTACCTCCTGTTTTGTATGCTTTAATAGAGGTAGATACCCT
ATCAAAAGCTAGGAAAGTAGAGATAGAAAAACTGAGATTTCAAATACATGAGACAG

AATCAAATGTCCAATTTTTGACATCTCAATTACATGAGATAGAATCAAATATCCAAC
ATATGAACTCTGAGAGAAACCAATTAGAGATAGATATCTCAAACAAGGAAGCCTTT
ATTAGGGACTTAGAAACTGATGCTGAACATAGTCGCATAGCATACCGTAAGCTCAG
AGAGGATTTTGAAGAATTACGGCAACAGTTGTTTAAAAATCAACAGTCAGGTGTAT
ATGGTGTGGAGAGATCTCCTGATCCTCCAATAATCACTCCCCAATATAAAAATTGGG
CACCATATGCTCCTAGAGCCAAACCAATTAGTAATGGTGATTCAGAGGCACAAGAT
GTTTTGAGCCGGTTGACTGAAGCTTGCCAAGCTGTGACCACCTTATTAGGCAATCAC
AGAACCAGGACCCCTATAGACTCCTGTCAAGACTCATCCCCAGAGAGGAATGAGCG
ACGTAAATCCCTTGTGTCCAATGTATCCTATAAAAGTTATGCTTCCAATCTTAATACA
AGTGGGTCTGAGAGTGATGAGAGGAAGCGGGATTCCAGTCCTGTCCCAAGAAGGGG
GAAAAACACCCATAAAAGGGGTTCAAAAGGTAAAAATCCCCAATATCCAGGTGAAA
AACAAAATTCTGCCAGCATAATTAAATTCTTGAATACTGCTGTGCCAAAATTCTCTA
AGAAAGGTTCTGCCCAAATAGCGACCCATTTAGAAATGTATGAGTCTACTATGGACT
CTTTGGACTTGTCTGAAGCAGACAGAATCAGGTTTCTTCCATGGGCCTTTGATGACA
GATACCGTCATTACTTTTCCTCCTTTAGAGAGAGAGGAATCACCAAATGGCAGTCGG
TCCTACATGAAGTTAAATTAGAATTTGGGCCATACCGAAATACCACTGCCGCAAAGA
GAGAAATTTATAGACTTACATGTCGGCCTAACCAAAGCCCCCGAGAATTTCTCTCTG
TGCTCAAAAATGCCTATGGTTTGGCCTACAGAAACCCTGATTGGGAATCTATAGATT
TTAAACAGTTGTTTTATGAGGCCTTACCAACCCAGATCAAATTAAGTCTTGCTCATG
ATCTGGATTTTGAAAACCCTCTAGAAAGACTGGTAACATCAGCAACCTTACTGTTTA
ACATCAGTGAGGGCCAAAACTTAAATGAAAGGAAAATTAAAAAATTCCCAGAACAC
AATGTCGATGAGTCCAGGGTAAAACCTCATTTAAATTTTGAGTCACAGCCAAAGAA
AGTACCTTCAGCCAATGGGTCTAATCAGCAGAACCGAAGACAACAGCCAAATCCCA
TACAAAACAAACCTCAAAATCCTAAGGGATCAGACGGGAACAGCTCAGGTTCTGGA
AGCCAGGACTACCGTCCTCGTTTCAACAAATTGAGCGACATTGTAAGTAAATTGTCT
CAAGTTTCTGCACCAAAACAGTCTGAGGATTTTTTAGATGTGAAGGCGGGACCCAG
TTCCGCCAATTAAAAGAAACAGGGTCAGATCCCTGTACGGAAATGGATACGGCAAT
ACCAAATATGGACAACGATCATATTTTGCATTGCACCGACCCAGGATTCAGCGGGGT
CAAAGGGAAGTCTAATATTTCATTTATCTTACAGCCAAATAAGGACAATTCTCGTGA
TCCAATGTCACTACAGCCCGTGTGTGAAATACCCTCTGTTCAGTTCGTCCCAGTGGT
GGAATCCATGGGAACCCAGATGACCTGCAGGCAGAACCAGGAGGAGATAAATCCA
AATATCCTGACGCTGCAGAAGGATCATTCTTTAAAACAACAGATGGCAAAACACTC
CAATTTTCTCTGTAATCTCTTACAGTTAGACGGCCGGTACTTTATAGATACCTGTCTC
CAAGATGGATGTGTCGGACCAATCCAAGGTTTATTGGATACCGGGTCTCAGGTCACT
ATCTTATCTTTTAGCTACTTCCAACAGATCTTGGGGTCATCAAAAAAGAAACCGAGA
TTAACGGCATTTGATGGCTCTTTGATAAGTGTGGGCGGTGACAGTTTACAGGTGAAA
GGCACATCCTGGCTGACATACAAGATTGGCAAAAAGATCCTTAGACATCCAACATT
GATTGTAGACCTCCCATATGATCGTTTGATCATTGGGATCGATTTGCTCAAGCGATT
AAATTCCATAATTGACTTTATAAATGAGGCAATTTGGACCCAAGTGAAAATTCCAAT
TGCTTTTGAGCGCTCCCACAATGCACAATCGCAACATAGTTGCCACATGATTGAAGA
AAGGCCTAATTCTGTTGAAATCCATTTCCGGAACAGTCAAATACCAAATATATCAAT
CCTGAAAAACGGTAAACCTGAAGAAAGTGTTAATACTTCCTTCCACATCATTAATAT
CCAAAAGGACAAAATTGAAAAGATAATTTTGGAGGATGATGTACTAACCATTTCTCT
TAAAGGGGGAAATCAGGAAGTTGCAGGCAAATGTGATTGCAAGGTGCATTTTATCC
ATCTCCCTAGGAAACAAATCTACAGAACACGTTTTCAGTGTGCTAAAGGAACCACA
ATATCAAATGTATTTAGGAAATTACATTATACATCGATTTGCCATACAGGTTGTCTG

ATCAATAATGTTCTGTGGTCCAGATTACCAGGGAATCCAGAGGAATTCCAGGATAA
AGAGCAAACCCTGAGATGGGGACAGCGAACACCCTATGCAGTGGATATCTTAGTCG
CCGAAAAGGTAAAAATCCCTGCGGGATGTAAATCATTTCTTTTACCCATTCGAGTGA
AGAAGGGGCAAAAGCTAAGAAATGCAGATGCATTATTTTGCTTATCCAATCGGATA
CAGCAATTTGGATTGAAGGTAAAACCGAATCCCATGATAAACATTCATCAGGATCC
ATTATACATGGTTATCCATAACATGAGTCCTTATAGCATTACGTTACAAGAGGACAC
AATGATTGGCCTAGCCATTGATTCAGAATATTACACTTTTGGGTTTCAAAATGATAT
AATTGGATTGATACCTGATGAATATCTAACAGAAGAACAAGTAGTGGATCAACATT
GCTTCTCAACACCTGAGGGATTGTTCAGTGTTCACTCTGTTTATCCTTTTAGTCCTGA
AGAAGGTACATGCTACATCGACGAGACATCATTGATCTTCAACCATGAGATCAATA
AAGGTGAGTATGAGTCATTCCAACAGGGGAAGGATCAGATATATCACACCCCGGAT
GACCTGACCAGCAAGCTTGAAGAGGCTTATGAGATAAATCAACCCGAGATTTTCCC
AGAATTTCAGGAGCGGGTGGAAGAGCAACTCTCTATAGCAGATGCATGTTCTGATG
AATCAGAACGACAACAGCTAAAGGAACTATTCTGGGAATTCCAGGATATGTTCGCC
AAGGACTCTTATGACTGTGGGGAAACTAACCTGCATGTTACAAGAATCCAGACCGA
TCCAGATGCACCCCCTGTGTTTGTTAAACAATACCGACTTCCGTTAGCGGCTTATGA
GTCACTCTCGGAAATCGTTAAAAATCTGGAGAAAAAGGGTATCATCCGTCCAGTAC
ACAGTTCTTTTAATCATCCAATACTTGGAGTCCTTAAACCCAATGGTCAGTTTCGTCT
TTGCTCAGACTTAAGACAGTTAAACAAACGAGTCTACATGTCTGGTTGGCCCGTACC
ATACATTGACCAATGCTTGGCGCAAATCCAAGGATCTAAAATTTTCACAGCCCTTGA
TTGCGCACAAGGATATTGGACAATTAAAATAGATGAAAGGGATCAACACAAACTGG
CCTTTACATTTGGAAAACAACAATATGCCTGGACCCGATTACCTTTCGGGTATATCA
ATGCAGGCCATGAGTTTGCTGTATTCATGCACAAAGCAATGCCTGATGCTGCTGAAC
GAGGTACATTATCTTATGTGGATGACATATTGATAAAAAGTACAACTTTTGAAGAGC
ATATTGCAGAACTAAAATATGTATTAAAACTTTCTTTACAGAAGGCCCAATGGTGCC
GATCTAAGGTAAATTTCCTAGGACATGAAGTTACTGCCGAAGGAATCAATCCACAG
AAAAAGAAAGTGGAAGCAATAAAGAACACGAAATCTCCTACAAATCTCAAGGAGCT
GAGATCATTCCTGGGTATGATGAACTATTCACGTAAGTTCATAGACAACTATGCGGA
AATTACAAAACCTCTTTTGCAGTTGCTGAAGAAAGGAGCAAAATGGGAATGGAATG
AGTGTCACGAACAAGCAGTATCTGAACTTAAAACTAAGCTTATCCAGGCTCCATGCC
TTGCTTATCCAGAAGGTGGAAAACCCTTCTATGTCGAAACAGGTTTTACCGACAAAA
GTGTAAGTGCGGTTCTTTTCCAAAAACAAGAAAAACTGAATAAAATAATTGCTTATG
CCAGCAAGTCATTATCACCAGTTGAAATAAAATTTAATAACTGTGAAAAAGCGTTAT
TAGCAACTGTATGGGCTTTGCAATATTTCAGGAGCTTTATTCAGGGTGAAAAGATTA
TTGTGGAAACTGCACATCAATCACTACAATATTTACAGAGTGGTCGGTTAAAGGATG
GAAACCTGTCAAACAGTAGGATAACTGCATGGACAATGTCCCTTATGGGATGGCCA
TTGGAGATCAGATATAAACAGGATACTAAAAACCCAGTTGCTCAAGGATTAGCTGA
ACTCCATGATTGTATTAATGTCGAACATAAAGATGAGTCACCAGAAAATGACTTTCT
GGAGGAACAATCATTATCTCCATATAAATCATATGAAGAGGAATACTGCAAATCATT
ACCATGTGTATATGTTGATGGCTGTTCTTTCCACACAGAAGAAGCTGAAAAGATGCT
GGTTGCAGGTATCGGGATCGTGTGGAATAATATATTCCCAGAGATATCCATAGGATA
CAAAATTGGTCCCAAAAGTAGCCAGGTCGCAGAGCTTGCCGCTGTGTACAAAGCAA
TACAAATGGCTATTGAATATGATCTTAAAGAATTTGTTTTAATTACAGATTCCGATTA
TGTTCGGAACAGTTTTGTGGAATACTTTCCCGGATGGAAAAGATCCAACATGTTGAG
AAGTAACAAGAAACCAGTCAAACATGGTAAGCTGTTCTGCAAAATCGATGAATTGG
TTACTACCTACGGTCTTACTATCTATTGGAAGAAGGTAAGGGGTCATTCAAAGCATC

```
CAGGCGCAGATAAAGATGGGAACGACCTAGCGGATTCCCTCGCCAAGAAAGCAGCC
ATTGATGGAGAGATCTTAGATATCGATGACCTCATGGGAACTATCCAAGTGGATGTC
ACAACCAGGGGACAGGCCCAAAAGGAGTCTCAACCTAATGTGGCAATGTGGAGCCA
AGATTCCCCTAGCGAAGATCTCATTGAGAGCCAGAAAGGGGATCCTATTATTGGTTT
ATTTTATAAATATATTCAAAATCCACAAGATCATCCCATAACTATGGAAGACTGTGC
TGATAATGAGGAGTTACGCGTTTTAATGAAAGGTGAGTCCCAATTCTCATTACAGGA
TGGATTGTTGATACGAACCTCGAAAAACGGTATCACACAATGGGTAGTACCCGCAG
CATATCGGGGTTTGATGTTACAACATGCCCATGATGCCCCGACAGCTGGTCACCGAG
GAGAGAAAATAACATATGAGTTGTTGAGGGACTATGCTTACTGGCCACACATGTTAC
AAGATGTCAGGACATATTGTCAAGGATGTTTGGTATGTCCTCAGTTCCAACCACAAG
GACCCAATCATCGGGCACCACTGATGAAAAGAGGAATATCCATGCCATGGTCAGAT
ATTCAGATTGATTTCATTGGACCAGTAACAACTTCCTCAAAAGGAAACCGATACATG
CTAACTGTAACTTGTTTGTTCACAAAATGGGTAGAATGTCTGCCATGTAAAACATGC
AGTTCTAGTGTATGCGCATCACTGCTCATTAACCATATATTTTCCAGATTTGGTCTTC
CCCAGCGTATTGAATCCGACCGTGGAAGTCACTTCACCAGTGAAGTGATGACCAAG
ATGTGGGAAATTCTAGGAGTAAAAAGGAAGCTACACATAGCTTACCGACCAGCTTC
TAGTGGTGGAGTAGAGCGTTACAATCAATCCATAGTAAACATTTTGAAAAAGTTTGT
CAAGGAATCTGGTAAAGACTGGGATGTCAAGTTGCCATTGGTTCTCATGGCTATTAG
AGCCACTCCAAGTGCAGCAACCAAACTTTCTCCATTTGAGATGATTACAGGTAGAAA
AATGGTGTTACCTCAACATTTACTTTACAGGACCACAGATCACAATTTGATGAATGC
AACAACTGCACATCAATATGTGGAAAACTTACGCAAGCATCTACAACATGCTTTTGC
ATTTGCCCAAAGAAATATTGAAAAAGCTGCAGTAAGTACGAAAACGTACTATGATC
TAAAAACAACTCAAAGGGAATATCAGGTGGATGACAAAGTTTATCTTTATAACTTTG
CCAGGGATCAGGTAAAGGAAAGGAAATTCCTACCTTCATGGAAAGGTCCATATATA
ATCACTGATAAATTATCTCCTGTGGTTTATAAAGTTAAAATTCCTAAGGGGGATGAT
TTTGTTGAAAAATGGGTGCATATTAACCAATTACGTGTTTGTCATCCTAGCTCACAG
CTTCGGAAAATAGAGCAAGTCGAGAGCTGA
>Loki-Xla consensus sequence Smc-Gag-Pol protein
MSVTGSDNGNDPEQLTRNIVEFLQPRVKFSQSVDDMIYDLKKYVKEDIDDIFAINGVVE
RYLTSMTLATNSADKEKRLMCALPPVLYALIEVDTLSKARKVEIEKLRFQIHETESNVQF
LTSQLHEIESNIQHMNSERNQLEIDISNKEAFIRDLETDAEHSRIAYRKLREDFEELRQQLF
KNQQSGVYGVERSPDPPIITPQYKNWAPYAPRAKPISNGDSEAQDVLSRLTEACQAVTT
LLGNHRTRTPIDSCQDSSPERNERRKSLVSNVSYKSYASNLNTSGSESDERKRDSSPVPR
RGKNTHKRGSKGKNPQYPGEKQNSASIIKFLNTAVPKFSKKGSAQIATHLEMYESTMDS
LDLSEADRIRFLPWAFDDRYRHYFSSFRERGITKWQSVLHEVKLEFGPYRNTTAAKREIY
RLTCRPNQSPREFLSVLKNAYGLAYRNPDWESIDFKQLFYEALPTQIKLSLAHDLDFENP
LERLVTSATLLFNISEGQNLNERKIKKFPEHNVDESRVKPHLNFESQPKKVPSANGSNQQ
NRRQQPNPIQNKPQNPKGSDGNSSGSGSQDYRPRFNKLSDIVSKLSQVSAPKQSEDFFRC
EGGTQFRQLKETGSDPCTEMDTAIPNMDNDHILHCTDPGFSGVKGKSNISFILQPNKDNS
RDPMSLQPVCEIPSVQFVPVVESMGTQMTCRQNQEEINPNILTLQKDHSLKQQMAKHSN
FLCNLLQLDGRYFIDTCLQDGCVGPIQGLLDTGSQVTILSFSYFQQILGSSKKKPRLTAFD
GSLISVGGDSLQVKGTSWLTYKIGKKILRHPTLIVDLPYDRLIIGIDLLKRLNSIIDFINEAI
WTQVKIPIAFERSHNAQSQHSCHMIEERPNSVEIHFRNSQIPNISILKNGKPEESVNTSFHII
NIQKDKIEKIILEDDVLTISLKGGNQEVAGKCDCKVHFIHLPRKQIYRTRFQCAKGTTISN
VFRKLHYTSICHTGCLINNVLWSRLPGNPEEFQDKEQTLRWGQRTPYAVDILVAEKVKIP
AGCKSFLLPIRVKKGQKLRNADALFCLSNRIQQFGLKVKPNPMINIHQDPLYMVIHNMS
```

PYSITLQEDTMIGLAIDSEYYTFGFQNDIIGLIPDEYLTEEQVVDQHCFSTPEGLFSVHSVY
PFSPEEGTCYIDETSLIFNHEINKGEYESFQQGKDQIYHTPDDLTSKLEEAYEINQPEIFPEF
QERVEEQLSIADACSDESERQQLKELFWEFQDMFAKDSYDCGETNLHVTRIQTDPDAPP
VFVKQYRLPLAAYESLSEIVKNLEKKGIIRPVHSSFNHPILGVLKPNGQFRLCSDLRQLNK
RVYMSGWPVPYIDQCLAQIQGSKIFTALDCAQGYWTIKIDERDQHKLAFTFGKQQYAW
TRLPFGYINAGHEFAVFMHKAMPDAAERGTLSYVDDILIKSTTFEEHIAELKYVLKLSLQ
KAQWCRSKVNFLGHEVTAEGINPQKKKVEAIKNTKSPTNLKELRSFLGMMNYSRKFID
NYAEITKPLLQLLKKGAKWEWNECHEQAVSELKTKLIQAPCLAYPEGGKPFYVETGFTD
KSVSAVLFQKQEKLNKIIAYASKSLSPVEIKFNNCEKALLATVWALQYFRSFIQGEKIIVE
TAHQSLQYLQSGRLKDGNLSNSRITAWTMSLMGWPLEIRYKQDTKNPVAQGLAELHDC
INVEHKDESPENDFLEEQSLSPYKSYEEEYCKSLPCVYVDGCSFHTEEAEKMLVAGIGIV
WNNIFPEISIGYKIGPKSSQVAELAAVYKAIQMAIEYDLKEFVLITDSDYVRNSFVEYFPG
WKRSNMLRSNKKPVKHGKLFCKIDELVTTYGLTIYWKKVRGHSKHPGADKDGNDLAD
SLAKKAAIDGEILDIDDLMGTIQVDVTTRGQAQKESQPNVAMWSQDSPSEDLIESQKGD
PIIGLFYKYIQNPQDHPITMEDCADNEELRVLMKGESQFSLQDGLLIRTSKNGITQWVVP
AAYRGLMLQHAHDAPTAGHRGEKITYELLRDYAYWPHMLQDVRTYCQGCLVCPQFQP
QGPNHRAPLMKRGISMPWSDIQIDFIGPVTTSSKGNRYMLTVTCLFTKWVECLPCKTCSS
SVCASLLINHIFSRFGLPQRIESDRGSHFTSEVMTKMWEILGVKRKLHIAYRPASSGGVER
YNQSIVNILKKFVKESGKDWDVKLPLVLMAIRATPSAATKLSPFEMITGRKMVLPQHLL
YRTTDHNLMNATTAHQYVENLRKHLQHAFAFAQRNIEKAAVSTKTYYDLKTTQREYQ
VDDKVYLYNFARDQVKERKFLPSWKGPYIITDKLSPVVYKVKIPKGDDFVEKWVHINQ
LRVCHPSSQLRKIEQVES*
>Loki-Xla consensus sequence *env* gene
ATGCCATCCAAGATGATCGCTATCTTATATGCAGTCCTGATCTTGGGGAAAGAGTTC
CATGCTGAGCCGATTGTGATGCCAGGTCCATCATCCGGGATCATGCTACAGGATACG
GCAGGATTCATCATGACCAACAAGAGGATCCTATCCCAGAAGGTTTATATGAGCCTA
GATCCACGTTGTGTCATCGAAAGACAAGTCAACGTATCCAATATTCGATCACCGGAG
ATTCAAACCTGGTACCAGCTACATATCGAATATTCACAAGAGCGGGTTACTCAGATC
CTGGAGCAAACCCGGAAAACTTTGACCAGAGAGCCCTTTTCTTCAAACCAAAGACC
AAAGCGGTTCATTTCTGCCATTGTTGCTGCTATAATCTTTGCGGTAGTAGGAACTGTG
TTTGCAACCGGTGTAACAATCGCTAATTCTGTATCCCTTAAAACCATGGAATCTGAG
ATTAGTTCATTAAAGACTAATTTACAACATATTCATGCAGAGATGGAAAGACAAAA
GAAAGATTTAACAGACTTGTATGCCGTTGTGGAAGACACTGTCGTTACCACGAACTT
ACACTTTAAATTGTTAGACCGCTCAATAGATCTACACCAAGGTCATGAGCAGTTTAA
GCAAGAACTTATTTTGTTAAAAGAACGTATCGACTTTCAACTTCAGGTTTTCCATGA
CGAGGAGCAAAGTGGGATGCAAGACCTAAGGGAAGATGTTACAAAGTGTCGATGG
AGAAAAGTAGAGGATATCCAAATGAACTTAGCATTTAATATGGGAGTTGCCATACC
ATTATACATCGATCCCATTAATATGGAAATTTGCTTTTTACTATCAATCCCAATGGTT
AATAGGAACAATATATTTCAAATGAAGAAAATCCACAATGTAGGTACGTGGAGAGA
CAATATTCACATTAAAATTCAAACTCCTGACATTGTAGCATATCAATCTTGGAAACC
GGACTGGCTAGCTGTTCCTATTCTGGATAATTGTGCAACATTCCGGGATAATAATTTT
CAATGTGAAGGCGATCCGTTCAATTATGATGCAACTAACGCTTTGTGCGGACTTGAA
TTTGAGAAACATGGATCAGAGAAATGCCAGGAGATCTTGTCTAAAATTGAGGAGAA
AGACTGGGTAAATGCAATTTTAATAAAAGACAAATGGCTCGTTAGTACTCGTCTGCC
AGACCTGACTGTTTTTTCTAACGACCACATTATGGCGAAAGTTATTGCTTTACCTTCC
AGTGTATCGATCATCAAGGTACCCCCTAATTCCAGAGTTGTGATTGGAAATCGCACC

CTATATCCTGTAAATATTGAAGACATTGAAGGAGAGGTGGAAAGCATTGATGCATTC
CAGAATCAGGAAATCACTATAAGTCCGGATTTAAAATTTTTGTTNGACAATAAGCCC
AGACATACCGTTCGTATGGTTATTAAAAAGCATAATGTTACCATGGATACCCCACAA
TACTCATCAACCGACGACGAGATATTGTTTCCCTCTGTTTGGTCATGGGGGGAAAAT
TACATATGGAGCATTTTGTTCGCAGTTATAGTAGGGTGGTTAACTGTAATCACCGTC
TTGGCAATCCGTTGGATGAAGTTCATCAAAAGAACTCTAGATAAGCTATGGCTTAAA
TCTCGTCGATTGAGTCTACATAATGAGTATGATGAAGGAAGAGTTTGA
>Loki-Xla consensus sequence Env protein
MPSKMIAILYAVLILGKEFHAEPIVMPGPSSGIMLQDTAGFIMTNKRILSQKVYMSLDPR
CVIERQVNVSNIRSPEIQTWYQLHIEYSQERVTQILEQTRKTLTREPFSSNQRPKRFISAIV
AAIIFAVVGTVFATGVTIANSVSLKTMESEISSLKTNLQHIHAEMERQKKDLTDLYAVVE
DTVVTTNLHFKLLDRSIDLHQGHEQFKQELILLKERIDFQLQVFHDEEQSGMQDLREDV
TKCRWRKVEDIQMNLAFNMGVAIPLYIDPINMEICFLLSIPMVNRNNIFQMKKIHNVGT
WRDNIHIKIQTPDIVAYQSWKPDWLAVPILDNCATFRDNNFQCEGDPFNYDATNALCGL
EFEKHGSEKCQEILSKIEEKDWVNAILIKDKWLVSTRLPDLTVFSNDHIMAKVIALPSSVS
IIKVPPNSRVVIGNRTLYPVNIEDIEGEVESIDAFQNQEITISPDLKFL?DNKPRHTVRMVIK
KHNVTMDTPQYSSTDDEILFPSVWSWGENYIWSILFAVIVGWLTVITVLAIRWMKFIKRT
LDKLWLKSRRLSLHNEYDEGRV*
>Loki-Xtr consensus sequence
TTGTCGGATTTTCAGATAAGATAATTATTAATTACAAGAATATTCTGATTTGTTAATG
ATAACGTGATTATGTTACTTAATAGAATATATATTCATGAATGATATTATGCTTTTTG
ATGAATCATATAGGAATAAATAATGAATGTTATTTTGTTTAGTTGATCAGAAACGTA
TATGTTGAAAAGGAGGTTTGTATCGGATTACTAGCCCAGATGTATAACCTTTGTGTT
AAAGCAAATATCAAAGGTTATATAATCCTTGGTGTCTAGGTGAACAAGATGAGAAC
ACCTTGAAGGATCTTACAAACTCTCCACAATATAGAGCAtACCTGTTGAGAGGGTAA
TAATTAATCAAGGACAGACCAATGACATATATTGTATGTGTACTAAGGTATTAAGTA
GTTACACAGCATTGTTCTAGTGTGGAGCAATGTCAACAATAGTGTGTAACATCGAAT
TAAGGAAGGTCAAAAGTTATTTTATATTTTCTGAATATTTTCAGACCACAATATTGA
GGAAACAGTGACATCCACAGATGGTGGAGGACATTACGGGTCCAAAGGTCACTGTA
ATAAATATTTAAATCAACTGGTGATAGTTTAACTTAAAGACTGCCCTTTGATGCCAG
TTGGCTAGTTGATAAATCCCaGAGGTGGAGTTTTAGATAGAGATGAGGGATAAGTTA
GGAGACAGGAGGACTaAGAAAAAAAAAAAGAAAAGAGGGATCTAAAAGAAAGAAC
CTGAAAAAGAaAATCTAAGGGGATCTAAAAACTGGAAGAGGAAGGAGAAACTACTC
TCAACTATAAGAAAACTCTACTCTGGTAAATATAAATTTTATTATACTTTTTCTCGAT
TAAGTTGTATGCTTGGATATTTATTATTCCTATGTATATTATAGTCTTACCCTAAACA
TTGACTTTCCCAGAGTTGCAAACTGGGAATAATCTGTTTAATTTGGGTCCATTTTATT
TTTCATTAGTATAGATATATGATATCCAAGTTAGCGATTTAATGGAGAAATAAGAAA
ATTTAAGGATAGAATTCCTTAAGACCTCAAAATCAATTATTTACCCTTCCTTTGATTC
AACATATTTAATTACCTTGTTATTAAGTATTGAAAGTGAACGTACATATATAATCTTG
TGGATGCTTGCCATTTATTGAAGAATGTCTATTGGTCAATGTTCATTTTGATTAATTA
TTAGATAGATATCAGGGAGTGTTAAGTAATAATTTCATTTTTGTAATATTCATATTTA
AATAAATTAATTATTTTATTAAATAATTGTCATGATTTCTCACCTTTGTACCAATTGA
TTGGGAAAGCTCGATATCGAATTCAGAGTTGGGCATATATTATTTAAGACTCCTCCT
GTTACTAATAAGGAGCCTTGAATTCTTTTAGATATACGCAGAATTAGCCCGACATAT
ATTTGGAGGCACTGCTGAGATCGAGATTATTTTCTCATATTGGTACAAATAGTGAAA
AACCATTCCAATACTTCTGTCAAGCTAATTATTTTGTTTATTAATTAAGAGACCATGA

GTGTTGCAAGCAGTGAGAATGGGAATACCCCAGAGCAAATGATCCACAGAATTGTT
GAATTCATATATACACGGGTAAAATTTGATATAGGGATGGATGTTTGGATTAATGAT
TTGAAGAGAAATGCTAGAGCAGACCTTAATGATATCTGGGCAACAGACGGCCTGGT
TGAACGCTATTTAACATGTATGAACTTAGCCTCTAACAGTTCAGACAAAGAAAAAA
AATTAATATGTGCCCTACCTCCTGTTTTATATGCATTAATGGAGGTAGATACCCTATC
AAAAGCTAGAAAAGAAAATGTAGAAAAATTGAGGTCTCAGTTACATGAGATAGAAT
CAGATATCCAACATATGAAATCTGATAAAAACCAACTAGAGATAGATATCTCAAAC
AAGGAAGCTTTTATTAGGGATTTAGAAACTGATGCTGAACGTAGTCGCATAGCATAC
TGTAAGCTCAGAAAGGATTTCGAAGAATTACAGCAACAGCACATTAAAAGTCAACA
GTCAGGTGTATATGGTGTGGAGAGATCTCCTGATCCTCCAATAATCACTCCCCAATA
TAAAAATTGGGCACCATATATTTCTAGAGCCAAACCAATTGATAATGTTGATTCAGA
GGCACAAGATGTTTTGGGCCGGTTGGCTGAAGCTTGCCAAGCTGTGACCACCTTAAT
AGGTAATCACAGAACAAGGGTTCTTAGAGATTCTTCTTCAGACACGTCCCCAGAGAG
GAATGTGAGGAATGTGCGAAATCAAACACCTATGTCAAATGTATCCCATAGAAGTA
GTGTTTCCAATACTGATGCAAGTGAAGTTGAGAGTGATGAGAGAGCATGGGGATCC
GGTCCTTTCCCAAACAGGGGGAAGAATACCCATAAGAGGGGTTCAAAAGGGAAAAA
TCCCCAATCTCCAAAAGAAAAACAAAATTCTGCCAGCATAATTAAATTTTTAAATAC
TGCTATACCAAAATTTTCTAAGAAAGGTTCTTCCCAAATTGCAAACCACTTAGAACT
ATATGAATCTACTATGGATTCATTAAAATTATCTGAGGCAGACAAAATCAGGTTTCT
TCCATGGGCCTTTGATGACAGATACCGTCATTACTTTTCCTCCTTTAAAGAGAGAGG
AATCACCAAATGGCAGTCAGTCCTACATGAAATTAAATCAGAATTTGGGCCCTATCG
AACTACCACTGCCGCAAAGAGAGAAATTTATAAACTTACATGTAGATCTAATCAAA
GCCCTCGAGAATTTCTCTCTGTACTTAAAAATGCCTATGGTTTAGCTTACAGAAACC
CCGATTGGGAATCTGTGGAATTTAAACAAGTATTCTATGAGGCCTTACCAACCCAGA
TCAAATTAAGTCTAGCTCATGATCTGGATTTTGAAAACCCTCTAGAAAGACTGGTAA
CATCAGCAACTTTGCTGTTTAACATTAGTGAGGGCCAAAACTTAAATGATAGGAAAT
TTAAAAAATTCCCAGAGCACAATGTTGATGAGTCCGGGGTAAAACCTCATTTAAATT
TTGAGTCCCAGCCAAAGAAAATACCTTCAGCCACTGGGCCTAATCAACAAAACCAA
AAACAGCTAAATCCCAGACAAACTAAACCTCAAATTGCCAAGGGGTCAGAAGGAAA
TAGTTCAGGTTCTGGAAGCCAAGGTTACCGTCCTCGTTTCAACCAAGGGTACCAaGG
ATACCGAGGTAACCAAAGATATGGGGGTTACCAAGGATACCGAGGTAACCAAAGAT
ATGGGGGTTACCAAGGATACCGGGATCCCCAGGGATACAGACGTTGGAACAATAGG
CCCAGGCAGCAATGGAGAAAAGGGCAGGAATCACCAGATTCACCCAGGGGTAAAT
CACCCACAGGGAAcCAAAATGGTCCACGGAATCAAAACCCAGGTAGACCTAGCAGA
TTTGATCAGCTTGCAGATCAGGTTTCCAAGTTAACTGACATTGTAAGTAAATTATCTC
AAGTTTCTGCAGGAAAACAACCTGAAGTTTTTTTtAGAGGAAAAAGCGGGGCCAAGC
TCCGCCAAATAAAGGAAGCAGAACCAAATTGCTATATGGAAATGGACAGGAAAGA
ATTAAATACAGTCAATGATCATGTTTTGCATTGCCCCGACCCTGAATTCAGGGCCGG
GAAGGGAAAATCTAATATTTCATCTGTTTTACAGCCTAATAAGGACAATTCTTATAA
TCCGGTGTTATTTGAACCAGTGTGTGAAGTACCTTCGGTTCAGTCTGTTCCATTAATA
GAaTCCATAGGAACACAAATGCCCTGCAGACAGAACAAGGAGGAGCAATCTCCAAA
TATCCTGACGCTGCAGAGAGATCACTTTTTAAATCAACAGATGATAAAACACTCCAA
TTTCCTCTGTAATCTCTTACAGTTAGATGGCCGGTATTTTATAGATATTTGTCTCCAG
GATGGATGTGTCGGACCAATCCAAGGTTTATTGGATACCGGATCTCAGGTCACTATT
TTGTCTTTCAGCTATTTCCAACAGATCTTAGAGTCCTCAAAGAATAAACCGAGATTG
ACGGTATTTGATGGATCTCTGGTAAGTGTAGGTGGTGATAGTCTACAAGTGAAAGGC

```
AAGTCCTGGCTGACATATAAGATTGGTAAAAAGACCATTAGACATCCAACATTGATT
GTAGATCTTCCATATGATCGGCTGATCATTGGAATTGATCTGCTCAAAAGGTTAAGA
TCCATAATTGACTTTATTAATGAAGCAATATGGACTCAGGTAAAAACCCCAATCTCT
TTTGAACGTTCCCACAATGCAAAACCGCAAAATAGTTGCCATATGATTGAAGAAAG
ACCTAATTCTGTCGAAATCCATTTCCGGAACAGTCAGGTACCAGATGTCTCAATCCT
AAAAAACGGTAAACCTGAAGAAAGCATTAATACTTCCTTCCACATCATTAATATCCA
AAAGGACAAAATTGAAAAGATAATCCTAGAGGATGATGTGTTAACCATTTCTCTTAA
AGGGGGAAATCGGGAAGTAGCAGGCATAACCAGACATATCCAATCATTGGTTAAAA
TCGAGAAGACCAATGATACTATGCTGGTTCCCGTACAGGTAAATAATATTGCCCGGG
TGAAATATGCCAAGCTGGATCTCAAATCCGAAGTCAGCTATATAAGTTTAGGACTTC
TAAAGCAGGTTACAAATCCTCAGATGATTAAATTTTCTTCCTTACAGGAATGGGTTT
ATGACCTGGACAAAGACGACCAGAGTCAAAATGTGATTGCAAAATGCATTTTATCT
ATCTCCCTAGGAAACAAATCTGCGGAACACGTTTTCAATGTGTTAAAGGAACCGCAA
TATCAAATGTATTTAGGAAATGACATTATACATCGGTTTGCCATACAAGTTGATTTG
ATCAATAATGTTTTGTGGTCCAGACTATCAGGGAATCCAGAGGAATTCCAGGATAAA
GAACAAGCCTTGAGATGGGGACAACGAATACCCTATGCAGTGGATATCTTAGTCGC
CGGAAAGGTAAAAGTCCCTGGGGGATGTAAATCATTTCTTTTACCCATTCGAGTAAA
GAAAGGGCAAAAGCTGAGGAATCCAGATGCATTAATTTGCTTATCCAATCGAATGC
AACAATTTGGATTGAAGGTAAAACCGAATCCTATAATAAATATTCATCAGGATCCAT
TATATATGGTTATCGATAACATGAGTCCTTATAACATTATGTTACAAGAGGACACAA
TGATTGGCCTAGCCATCGATTCAGAATATTATACTTTTGGGTTCCAAAATGATATAA
TTGGATTAATACCTGACGAATATCTAACAGAAGAACAAATGGTGGATCAACATTGCT
TTTCAACACCTGAGGGATTGTTCAGTATTCACTCTGTTTATCCTTTCAATCCTGAGGA
TGGTACATGTTACGTTGAGGAGACATCATTGATCTTTAACCATGAAATCGAGAAAGG
TGAGTATGAGTCATTCCAACAGGGGAAAGATCAGATATATCACACCCCAGATGACC
TGACCAGCAAGCTTGAAGAAGCTTATGAGGTAAATCAACCTGAAATTTTCCCAGAAT
TTCAAGAGCGGGTGGAAGAGCAACTCTCTATAGCTGATGCATGTTCTGATGAATCAG
AACGACAACAGCTAAAAGAACTATTTTGGGAATTCCAGGATATGTTTGCCAAGGATT
CTTATGACTGCGGGGAAACTAACCTACATGTTACAAGAATCCAGACCGATCCAGAT
GCACCCCCTGTATTTATTAAACAATACCGACTTCCGTTAGCGGCTTATGAGTCACTCT
CGGAAATCGTTAAAAATCTGGAGAAGAAGGGTATCATCCGCCCAGTACATAGTTCTT
TTAATCATCCAATACTTGGAGTCCTTAAACCCAATGGTCAGTTTCGTCTTTGCTCAGA
CTTGAGACAGTTGAACAAACAAGTCTACATGTCTGGTTGGCCTGTACCATACATTGA
CCAATGCTTGGCACAAATCCAAGGGTCTAAAATTTTTACCGCCCTTGATTGCGCACA
AGGATATTGGACAATTAAAATAGATGAAAGGGATCAATATAAACTAGCCTTTACAT
TTGGAAAGCAACAATATGCATGGACCCGATTACCTTTCGGATACATTAATGCAGGCC
ATGAGTTTGCTGTATTCATGCACAAAGCCATGCCTGATGCTGCTGAACGAGGTACAT
TATCATATGTGGATGACATATTGATAAAAAGTACAACTTTTGgAGAACATATTGCAG
AATTGAGATATGTATTAAATCAACTAAAGAAGGCAGGTGTTAAACTTTCTTTACAGA
AGGCTCAATGGTGCCGATCTAAGGTAAATTTCCTTGGACATGAAATTACTGCCGAAG
GAATCAATCCACAGgAAGATCGTTCCTGGGTATGnaTGAATTATTCACGTAAGTTCAT
AGACAACTATGCAGAAATTACAAAACCTCTTTTGCAGTTGCTGAAGAAAGGAGCAA
AATGGAAATGGGATGAATGTCACGATCAAGCAATATCTGAACTTAAAACTAAGCTT
ATCCAGGCTCCATGCCTTGCATATCCAGAAGGTGGAAAACCCTTCTATATCGAAACA
GGTTTTACCAACAAAAGTGTAAGTGCGGTTCTTTTCCAAAAACAAGAAAAACTGAAT
AAAATAATCGCTTATGCCAGCAAGTCGTTATCATCAGTTGAAATAAAATTTAATAAC
```

TGTGAAAAAGCATTATTAGCAACTGTATGGGCTTTGCAATATTTCAGGAGTTTTATT
CAAGGTGAAAAGATAATTGTGGAAACTGCACATCAATCACTACAATATTTACAGAG
TGGTCGATTAAAGGATGGAAACCTGTCAAATAGTAGGATAACTGCATGGACAATGT
CCCTTATGGGATGGCCATTGGAGATCAGATATAAACAGGATACTAAAAATCCAGTT
GCTCAAGGATTAGCTGAACTCCATGATTGTATTAATGTTGAACACAAAGATGAGTCA
CCAGAAAATGACTTTCTGGAGGAACAATCACTATCTCCATATAAATCATATGAAGAA
GAATACTGCAAATCATTACCATGTGTATATGTTGATGGCTGTTCTTTCCATACAGAA
GAAGCTGAAAAGATGCTAGTTGCAGGTATTGGGATAGTATGGAATAATATATTCCC
AGAGATATCCACAGGATACAAGATTGGTTCCAAAAGTAGCCAAGTCGCAGAGCTTG
CtGCTGTGTACAAAGCCGTACAAATGGCTATTGAATATGATATTAAAGAATTTGTTTT
AATTACAGATTCCGATTATGTTCGGAACAGCTTTGTGGAATATTTTCCCGGATGGAA
AAGATCCAATATGATGAGAAGTAATAAGAAACCAGTCAAACATGGTAAGTTGTTTT
GTAAAATCGATGAATTAGTTACTACCCACGGCCTTAACATATATTGGAAGAAGGTAA
GGGGTCATTCGAAACATCCAGGTGCAGATAAAGATGGAAATGACTTAGCAGATTCC
CTTGCCAAGAAAGCAGCCATGGAGGGAGAGATCTTGGATATTGATGACCTCATGGG
AACTATCCAAGTGAATGCCACAACCAGGAGTCAGACCCAAAAGGAATCTCAACCTA
ATGTGGCAATGTGGAGCCAAGATTCCCCTAGCGAAGATCTCATTGAGAGTCAGAAA
GGGGATCCTATTATTGGTTTATTTTATAAATATATTCAAAATCCAAAGGATCATCCC
ATAACCATGGAAGAATGTGCTGATAAAGAGGAGTTACGTATTTTGGTGAAAGGCGA
GTCCCAATTTTCATTACAAGATGGATTGCTGATACGAACCTCGAAAAATGGTATTAC
ACAATGGGTAGTACCCGCAGCATATCGGGGTTTGATGTTACAACATGCTCATGATGC
CCCGACAGCTGGTCATCGAGGAGAGAAAATAACATATGAATTGTTAAGGGATTATG
CTTATTGGCCACACATGTTACAGGATGTCAGGACTTATTGTCAAGGATGTTTGATAT
GTCCTCAATTCCAGCCACAAGGACCCAATCACCGGGCACCGCTGATGAAAAGAGGA
ATATCCATGCCATGGTCAGACCTCCAAATTGATTTCATTGGACCAGTAACAACATCC
TCAAAAGGAAACCGGTATATGCTAACTGTAACTTGTTTATTTACAAAATGGGTAGAA
TGTTTGCCATGCAAGACGTGTAGTTCTAGTGTATGCGCATCACTGCTCATTAACCAT
ATATTTTCCAGATTTGGTCTTCCCCAACGTATTGAATCTGACCGCGGAAGTCACTTCA
CCAGCGAAGTAATGACTAAGATGTGGGAAATTCTAGGAGTAAAAAGGAAGCTACAC
ATAGCTTACAGACCAGCTTCTAGTGGTGGAGTAGAGCGTTACAATCAATCAATCATA
AACATTCTGAAAAAGTTTGTCAAAGAATCTGGTAAAGACTGGGATGTCAAATTACC
ACTAGTTCTCATGGCTATTAGAGCCACCCCGAGTGCGGCAACCAAAATTTCTCCATT
TGAGATGATTACGGGTAGAAAAATGGTATTACCTCAACATCTACTTTATAGAACCAC
AGATCACAATTTGATGAATGCAATGACTGCACATCAATATGTAGAAAATTTACGAA
AGCATCTACAGCATGCTTTTGCGTTTGCTCAAAGAAATATTGAGAAAGCTGCAGTAA
GTACGAAAACGTACTATGATCTAAAAACAACTCAAAAGGAATATCAGGTTGATGAT
AAAGTTTATCTTTATAACTTTGCCAGAGACCAAGTAAAGGAAAGGAAATTCTTACCT
TCCTGGAAAGGTCCATATATAATCACTGATAAATTATCTCCTGTGGTTTATAAAGTT
AAAATTCCTAAGGGGGATGATTTTGTTGAAAAATGGGTGCACATTAACCAATTACGT
GTTTGTCATCCTAGCCAACAGCTTCGGAGAATAGAGCAAGTCGAGAGCTGAGTAGT
AAGAAATCAGGAGATTTTAGTTAAAGATGGAAGTAGAGATTAATATTGTATATAGA
TGTTGTATCATGATATGCCTATGTCTCTCTACCTTCTAAATATCCTTTATCTTCTCAGT
GAACTGAAACTTGCTATGTATATATGACGGCTGAACCTAATGATTTCTTTTGATTATC
TATAGGCGCAAAGATGCCATCCAAGATGATCGCTATTTTATATGCAGTCCTGATCTT
GGGGAAAGAGTTACACGCTGAGCGGATTGTGATGCCAGGTCCATCATCCGGGATCA
TGCTACAGGATACAGCGGGATTCATCATGACCAACAAGAGAATCTTGTCCCAGAAG

ATTTATATGAGCCTAGATCCACGTTGTGTCATCGAAAGACAAGTCAACGTATCCAAT
ATTCGATCACCGGAGATTCAAACCTGGTACCAGCTGCATATTGAATATTCACAAAAG
CGGATTACTCAAATCCTGGAGCAAACCCGGAAAACTTTAATCAGAGAGCCTTTTTCG
TCAAACCAAAGACCAAAACGGTTCATTTCTGCTATTGTTGCTGCTATAATCTTTGCA
GTAGTAGGAAACGTGTTTGCAACCGGAGTAACAATCGCTAATTCCATATCCCTTAAA
ACCATGGAATCTGAGATTAATTCATTAAAAACTAATTTACAACATATTCATGTGGAG
ATGGAAAGACAAAAGAAAGATTTAACAGACTTGTATACCGTTGTGGAAGACACTGT
CATTGCTACGAACTTGCACTCTAAATTATTAAACCACTCGATAAATCTTCACCAAAG
TCATGAACTGTTTAAACAAGAACTGATTTTGCTAAAGGAGCGTATCGACTTTCAACT
TCAAGTTTTCCATGACGAGGTGCAAAGTGGAATGCAAGACCTGAGAGAAGGGCGAA
TTCCTCTCTATTTTGTATCATACGAGATTATTCGCAAGATGTTACAAAATGCAGATGG
AGATAAAGTAGAGGATATCCAAATGAACTTAGCGTTTAATATGGGAGTTGCCATAC
CATTATACATTGATCCTGTTAAAATGGAAATCTGCTTTTTACTATCAATTCCAGTGGT
CAATAGGAACAATATATTTCAAATGAAAAAAATCCACAATGTAGGTACATGGAGAG
ATAATATTCATGTTAAGATCCAAACCCCTGATATCGTGGCATATCAATCTTGGAGAC
CGGACTGGCTATCTGTTCCTATTCTAAATAATTGTGCAATTTTTAGAGATAATAACTT
TCAATGTGAAGGTGACCCATTTAATTATGATACAACTAATGCTTTGTGCGGACTTGA
CTTTGAGAAACATGGATCAGAAAAATGCCAGGTAACTTTGTCTAAAATCGAGGAGA
AAGATTGGGTGAATGCAGTTTTAATAAAAGACAAATGGCTCGTTAGTACTCGTTTAC
CTGACCTAACTGTTTTTTCTAAGGACCACATTATGGCAAAAATTATTGCTATACCAA
CTCATGTATCGATTATCAAAGTACCCCCTAATTCCAGAGTTGTAATTGGAAATCGCA
CaTTATATCCTGTAAACATTGAGGACACTGAAGGAGAGGTGGAAAGCATTGATGCAT
TCCAAAATCAGGAGATTACCATAAGTCCGGATTTAAAATTTTTGTTGGACAATAAGC
CCAGCCATACCATCCGTATGGTTATAAAAAAACATAATATCACCATGGATACTCCAC
AATTTTCACCAACCGATGACGAAATACTGTTTCCTTCTGTTTGGTCATGGGGGGAAA
ACTATATATGGAGTATATTGATCACAATTATAGTAGGATGGTTAACTGTAATCACCA
TCTTAGCAATTCGTTGGATGAAGATCATCAAAAAGACTCTAGATAAACTATGGCTTA
AATCTCGCCGATCAAGTTTACATAATGAGTATGATGAAGGAAGAGTTTGAAAACCCT
AAGTAATTTTCCTAAAATGTTCATATTATAGGAAACATTTTAGAAAAACAAGGGGGA
TTTGTCGGATTTTCAGATAAGATAATTATTAATTACAAGAATATTCTGATTTGTTAAT
GATAACGTGATTATGTTACTTAATAGAATATATATTCATGAATGATATTATGCTTTTT
GATGAATCATATAGGAATAAATAATGAATGTTATTTTGTTTAGTTGATCAGAAACGT
ATATGTTGAAAAGGAGGTTTGTATCGGATTACTAGCCCAGATGTATAACCTTTGTGT
TAAAGCAAATATCAAAGGTTATATAATCCTTGGTGTCTAGGTGAACAAGATGAGAA
CACCTTGAAGGATCTTACAAACTCTCCACAATATAGAGCACACCTGTTGAGAGGGTA
ATAATTAATCAAGGACAGACCAATGACATATATTGTATGTGTACTAAGGTATTAAGT
AGTTACACAGCATTGTTCTAGTGTGGAGCAATGTCAACAATAGTGTGTAACATCGAA
TTAAGGAAGGTCAAAAGTTATTTTATATTTTCTGAATATTTTCAGACCACAATATTGA
GGAAACAGTGACATCCACAGATGGTGGAGGACATTACGGGTCCAAAGGTCACTGTA
ATAAATATTTAAATCAACTGGTGATAGTTTAACTTAAAGACTGCCCTTTGATGCCAG
TTGGCTAGTTGATAAATCCCTGAGGTGGAGTTTTAGATAGAGATGAGGGATAAGTTA
GGAGACAAGAGGACTGAGAAAAAAAAAAAGAAAAGAGGGATCTAAAAGAAAGAA
CCTGAAAAAGAAAATCTAAGGGGATCTAAAAACTGGAAGAGGAAGGAGAAACTAC
TCTCAACTATAAGAAAACTCTACTCTGGTAAATATAAATTTTGTTATACTTTTTCTCG
ATTAAGTTGTATGCTTGGATATTTATTATTCCTATGTATATTATAGTCTTACCCTAAA
CATTGACTTTCCCAGAGTTGCAAACTGGGAATAATCTGTTTAATTTGGGTCCATTTTA

TTTTTCATTAGTATAGATATATGATATCCAAGTTAgCAATCTAATGGGGAAATAAGA
ACATTTAAAGATAGAATTCCTTAAGACCTCAAAATCAATCATTTACCCTTCCTTTGAT
TCAACATATTTAATTACCTTGTTATTAAGTATTGAAAGTGAACGTACATATATTATCT
TGTGGATGCTTGCCATTTATTGAATAATGTCTATTGGTCAATGTTCATTTTGATTAAT
TATTAGATAGATATCAGGGAGTGTTAAGTAATAATTTCATTTTTGTAATATTCATATT
TAAATAAATTAATTATTTTATTAAATAATTGTCATGATTTCTCACCTTTGTACCAATT
GATTGGGAAAGCTCGATATCGAATTCAGAGTTGGGCATATATTATTTAAGACTCCTC
CTGTTACTAATAAGGAGCCTTAAATTCTTTTAGA
>Loki-Xtr consensus sequence 5’LTR
TTGTCGGATTTTCAGATAAGATAATTATTAATTACAAGAATATTCTGATTTGTTAATG
ATAACGTGATTATGTTACTTAATAGAATATATATTCATGAATGATATTATGCTTTTTG
ATGAATCATATAGGAATAAATAATGAATGTTATTTTGTTTAGTTGATCAGAAACGTA
TATGTTGAAAAGGAGGTTTGTATCGGATTACTAGCCCAGATGTATAACCTTTGTGTT
AAAGCAAATATCAAAGGTTATATAATCCTTGGTGTCTAGGTGAACAAGATGAGAAC
ACCTTGAAGGATCTTACAAACTCTCCACAATATAGAGCAtACCTGTTGAGAGGGTAA
TAATTAATCAAGGACAGACCAATGACATATATTGTATGTGTACTAAGGTATTAAGTA
GTTACACAGCATTGTTCTAGTGTGGAGCAATGTCAACAATAGTGTGTAACATCGAAT
TAAGGAAGGTCAAAAGTTATTTTATATTTTCTGAATATTTTCAGACCACAATATTGA
GGAAACAGTGACATCCACAGATGGTGGAGGACATTACGGGTCCAAAGGTCACTGTA
ATAAATATTTAAATCAACTGGTGATAGTTTAACTTAAAGACTGCCCTTTGATGCCAG
TTGGCTAGTTGATAAATCCCaGAGGTGGAGTTTTAGATAGAGATGAGGGATAAGTTA
GGAGACAGGAGGACTaAGAAAAAAAAAAAGAAAAGAGGGATCTAAAAGAAAGAAC
CTGAAAAAGAaAATCTAAGGGGATCTAAAAACTGGAAGAGGAAGGAGAAACTACTC
TCAACTATAAGAAAACTCTACTCTGGTAAATATAAATTTTATTATACTTTTTCTCGAT
TAAGTTGTATGCTTGGATATTTATTATTCCTATGTATATTATAGTCTTACCCTAAACA
TTGACTTTCCCAGAGTTGCAAACTGGGAATAATCTGTTTAATTTGGGTCCATTTTATT
TTTCATTAGTATAGATATATGATATCCAAGTTAGCGATTTAATGGAGAAATAAGAAA
ATTTAAGGATAGAATTCCTTAAGACCTCAAAATCAATTATTTACCCTTCCTTTGATTC
AACATATTTAATTACCTTGTTATTAAGTATTGAAAGTGAACGTACATATATAATCTTG
TGGATGCTTGCCATTTATTGAAGAATGTCTATTGGTCAATGTTCATTTTGATTAATTA
TTAGATAGATATCAGGGAGTGTTAAGTAATAATTTCATTTTTGTAATATTCATATTTA
AATAAATTAATTATTTTATTAAATAATTGTCATGATTTCTCACCTTTGTACCAATTGA
TTGGGAAAGCTCGATATCGAATTCAGAGTTGGGCATATATTATTTAAGACTCCTCCT
GTTACTAATAAGGAGCCTTGAATTCTTTTAGA
>Loki-Xtr consensus sequence 3’LTR
TTGTCGGATTTTCAGATAAGATAATTATTAATTACAAGAATATTCTGATTTGTTAATG
ATAACGTGATTATGTTACTTAATAGAATATATATTCATGAATGATATTATGCTTTTTG
ATGAATCATATAGGAATAAATAATGAATGTTATTTTGTTTAGTTGATCAGAAACGTA
TATGTTGAAAAGGAGGTTTGTATCGGATTACTAGCCCAGATGTATAACCTTTGTGTT
AAAGCAAATATCAAAGGTTATATAATCCTTGGTGTCTAGGTGAACAAGATGAGAAC
ACCTTGAAGGATCTTACAAACTCTCCACAATATAGAGCACACCTGTTGAGAGGGTAA
TAATTAATCAAGGACAGACCAATGACATATATTGTATGTGTACTAAGGTATTAAGTA
GTTACACAGCATTGTTCTAGTGTGGAGCAATGTCAACAATAGTGTGTAACATCGAAT
TAAGGAAGGTCAAAAGTTATTTTATATTTTCTGAATATTTTCAGACCACAATATTGA
GGAAACAGTGACATCCACAGATGGTGGAGGACATTACGGGTCCAAAGGTCACTGTA
ATAAATATTTAAATCAACTGGTGATAGTTTAACTTAAAGACTGCCCTTTGATGCCAG

TTGGCTAGTTGATAAATCCCTGAGGTGGAGTTTAGATAGAGATGAGGGATAAGTTA
GGAGACAAGAGGACTGAGAAAAAAAAAAAGAAAAGAGGGATCTAAAAGAAAGAA
CCTGAAAAAGAAAATCTAAGGGGATCTAAAAACTGGAAGAGGAAGGAGAAACTAC
TCTCAACTATAAGAAAACTCTACTCTGGTAAATATAAATTTTGTTATACTTTTTCTCG
ATTAAGTTGTATGCTTGGATATTTATTATTCCTATGTATATTATAGTCTTACCCTAAA
CATTGACTTTCCCAGAGTTGCAAACTGGGAATAATCTGTTTAATTTGGGTCCATTTTA
TTTTTCATTAGTATAGATATATGATATCCAAGTTAgCAATCTAATGGGGAAATAAGA
ACATTTAAAGATAGAATTCCTTAAGACCTCAAAATCAATCATTTACCCTTCCTTTGAT
TCAACATATTTAATTACCTTGTTATTAAGTATTGAAAGTGAACGTACATATATTATCT
TGTGGATGCTTGCCATTTATTGAATAATGTCTATTGGTCAATGTTCATTTTGATTAAT
TATTAGATAGATATCAGGGAGTGTTAAGTAATAATTTCATTTTTGTAATATTCATATT
TAAATAAATTAATTATTTTATTAAATAATTGTCATGATTTCTCACCTTTGTACCAATT
GATTGGGAAAGCTCGATATCGAATTCAGAGTTGGGCATATATTATTTAAGACTCCTC
CTGTTACTAATAAGGAGCCTTAAATTCTTTTAGA
>Loki-Xtr consensus sequence *spc7-gag-pol* gene
ATGAGTGTTGCAAGCAGTGAGAATGGGAATACCCCAGAGCAAATGATCCACAGAAT
TGTTGAATTCATATATACACGGGTAAAATTTGATATAGGGATGGATGTTTGGATTAA
TGATTTGAAGAGAAATGCTAGAGCAGACCTTAATGATATCTGGGCAACAGACGGCC
TGGTTGAACGCTATTTAACATGTATGAACTTAGCCTCTAACAGTTCAGACAAAGAAA
AAAAATTAATATGTGCCCTACCTCCTGTTTTATATGCATTAATGGAGGTAGATACCC
TATCAAAAGCTAGAAAAGAAAATGTAGAAAAATTGAGGTCTCAGTTACATGAGATA
GAATCAGATATCCAACATATGAAATCTGATAAAAACCAACTAGAGATAGATATCTC
AAACAAGGAAGCTTTTATTAGGGATTTAGAAACTGATGCTGAACGTAGTCGCATAG
CATACTGTAAGCTCAGAAAGGATTTCGAAGAATTACAGCAACAGCACATTAAAAGT
CAACAGTCAGGTGTATATGGTGTGGAGAGATCTCCTGATCCTCCAATAATCACTCCC
CAATATAAAAATTGGGCACCATATATTTCTAGAGCCAAACCAATTGATAATGTTGAT
TCAGAGGCACAAGATGTTTTGGGCCGGTTGGCTGAAGCTTGCCAAGCTGTGACCACC
TTAATAGGTAATCACAGAACAAGGGTTCTTAGAGATTCTTCTTCAGACACGTCCCCA
GAGAGGAATGTGAGGAATGTGCGAAATCAAACACCTATGTCAAATGTATCCCATAG
AAGTAGTGTTTCCAATACTGATGCAAGTGAAGTTGAGAGTGATGAGAGAGCATGGG
GATCCGGTCCTTTCCCAAACAGGGGGAAGAATACCCATAAGAGGGGTTCAAAAGGG
AAAAATCCCCAATCTCCAAAAGAAAAACAAAATTCTGCCAGCATAATTAAATTTTTA
AATACTGCTATACCAAAATTTTCTAAGAAAGGTTCTTCCCAAATTGCAAACCACTTA
GAACTATATGAATCTACTATGGATTCATTAAAATTATCTGAGGCAGACAAAATCAGG
TTTCTTCCATGGGCCTTTGATGACAGATACCGTCATTACTTTTCCTCCTTTAAAGAGA
GAGGAATCACCAAATGGCAGTCAGTCCTACATGAAATTAAATCAGAATTTGGGCCC
TATCGAACTACCACTGCCGCAAAGAGAGAAATTTATAAACTTACATGTAGATCTAAT
CAAAGCCCTCGAGAATTTCTCTCTGTACTTAAAAATGCCTATGGTTTAGCTTACAGA
AACCCCGATTGGGAATCTGTGGAATTTAAACAAGTATTCTATGAGGCCTTACCAACC
CAGATCAAATTAAGTCTAGCTCATGATCTGGATTTTGAAAACCCTCTAGAAAGACTG
GTAACATCAGCAACTTTGCTGTTTAACATTAGTGAGGGCCAAAACTTAAATGATAGG
AAATTTAAAAAATTCCCAGAGCACAATGTTGATGAGTCCGGGGTAAAACCTCATTTA
AATTTTGAGTCCCAGCCAAAGAAAATACCTTCAGCCACTGGGCCTAATCAACAAAA
CCAAAAACAGCTAAATCCCAGACAAACTAAACCTCAAATTGCCAAGGGGTCAGAAG
GAAATAGTTCAGGTTCTGGAAGCCAAGGTTACCGTCCTCGTTTCAACCAAGGGTACC
AAGGATACCGAGGTAACCAAAGATATGGGGGTTACCAAGGATACCGAGGTAACCAA

AGATATGGGGGTTACCAAGGATACCGGGATCCCCAGGGATACAGACGTTGGAACAA
TAGGCCCAGGCAGCAATGGAGAAAAGGGCAGGAATCACCAGATTCACCCAGGGGT
AAATCACCCACAGGGAACCAAAATGGTCCACGGAATCAAAACCCAGGTAGACCTAG
CAGATTTGATCAGCTTGCAGATCAGGTTTCCAAGTTAACTGACATTGTAAGTAAATT
ATCTCAAGTTTCTGCAGGAAAACAACCTGAAGTTTTTTTAGAGGAAAAAGCGGGGC
CAAGCTCCGCCAAATAAAGGAAGCAGAACCAAATTGCTATATGGAAATGGACAGGA
AAGAATTAAATACAGTCAATGATCATGTTTTGCATTGCCCCGACCCTGAATTCAGGG
CCGGGAAGGGAAAATCTAATATTTCATCTGTTTTACAGCCTAATAAGGACAATTCTT
ATAATCCGGTGTTATTTGAACCAGTGTGTGAAGTACCTTCGGTTCAGTCTGTTCCATT
AATAGAATCCATAGGAACACAAATGCCCTGCAGACAGAACAAGGAGGAGCAATCTC
CAAATATCCTGACGCTGCAGAGAGATCACTTTTTAAATCAACAGATGATAAAACACT
CCAATTTCCTCTGTAATCTCTTACAGTTAGATGGCCGGTATTTTATAGATATTTGTCT
CCAGGATGGATGTGTCGGACCAATCCAAGGTTTATTGGATACCGGATCTCAGGTCAC
TATTTTGTCTTTCAGCTATTTCCAACAGATCTTAGAGTCCTCAAAGAATAAACCGAG
ATTGACGGTATTTGATGGATCTCTGGTAAGTGTAGGTGGTGATAGTCTACAAGTGAA
AGGCAAGTCCTGGCTGACATATAAGATTGGTAAAAAGACCATTAGACATCCAACAT
TGATTGTAGATCTTCCATATGATCGGCTGATCATTGGAATTGATCTGCTCAAAAGGT
TAAGATCCATAATTGACTTTATTAATGAAGCAATATGGACTCAGGTAAAAACCCCAA
TCTCTTTTGAACGTTCCCACAATGCAAAACCGCAAAATAGTTGCCATATGATTGAAG
AAAGACCTAATTCTGTCGAAATCCATTTCCGGAACAGTCAGGTACCAGATGTCTCAA
TCCTAAAAAACGGTAAACCTGAAGAAAGCATTAATACTTCCTTCCACATCATTAATA
TCCAAAAGGACAAAATTGAAAAGATAATCCTAGAGGATGATGTGTTAACCATTTCTC
TTAAAGGGGGAAATCGGGAAGTAGCAGGCATAACCAGACATATCCAATCATTGGTT
AAAATCGAGAAGACCAATGATACTATGCTGGTTCCCGTACAGGTAAATAATATTGCC
CGGGTGAAATATGCCAAGCTGGATCTCAAATCCGAAGTCAGCTATATAAGTTTAGG
ACTTCTAAAGCAGGTTACAAATCCTCAGATGATTAAATTTTCTTCCTTACAGGAATG
GGTTTATGACCTGGACAAAGACGACCAGAGTCAAAATGTGATTGCAAAATGCATTTT
ATCTATCTCCCTAGGAAACAAATCTGCGGAACACGTTTTCAATGTGTTAAAGGAACC
GCAATATCAAATGTATTTAGGAAATGACATTATACATCGGTTTGCCATACAAGTTGA
TTTGATCAATAATGTTTTGTGGTCCAGACTATCAGGGAATCCAGAGGAATTCCAGGA
TAAAGAACAAGCCTTGAGATGGGGACAACGAATACCCTATGCAGTGGATATCTTAG
TCGCCGGAAAGGTAAAAGTCCCTGGGGGATGTAAATCATTTCTTTTACCCATTCGAG
TAAAGAAAGGGCAAAAGCTGAGGAATCCAGATGCATTAATTTGCTTATCCAATCGA
ATGCAACAATTTGGATTGAAGGTAAAACCGAATCCTATAATAAATATTCATCAGGAT
CCATTATATATGGTTATCGATAACATGAGTCCTTATAACATTATGTTACAAGAGGAC
ACAATGATTGGCCTAGCCATCGATTCAGAATATTATACTTTTGGGTTCCAAAATGAT
ATAATTGGATTAATACCTGACGAATATCTAACAGAAGAACAAATGGTGGATCAACA
TTGCTTTTCAACACCTGAGGGATTGTTCAGTATTCACTCTGTTTATCCTTTCAATCCT
GAGGATGGTACATGTTACGTTGAGGAGACATCATTGATCTTTAACCATGAAATCGAG
AAAGGTGAGTATGAGTCATTCCAACAGGGGAAAGATCAGATATATCACACCCCAGA
TGACCTGACCAGCAAGCTTGAAGAAGCTTATGAGGTAAATCAACCTGAAATTTTCCC
AGAATTTCAAGAGCGGGTGGAAGAGCAACTCTCTATAGCTGATGCATGTTCTGATGA
ATCAGAACGACAACAGCTAAAAGAACTATTTTGGGAATTCCAGGATATGTTTGCCA
AGGATTCTTATGACTGCGGGGAAACTAACCTACATGTTACAAGAATCCAGACCGATC
CAGATGCACCCCCTGTATTTATTAAACAATACCGACTTCCGTTAGCGGCTTATGAGT
CACTCTCGGAAATCGTTAAAAATCTGGAGAAGAAGGGTATCATCCGCCCAGTACAT

AGTTCTTTTAATCATCCAATACTTGGAGTCCTTAAACCCAATGGTCAGTTTCGTCTTT
GCTCAGACTTGAGACAGTTGAACAAACAAGTCTACATGTCTGGTTGGCCTGTACCAT
ACATTGACCAATGCTTGGCACAAATCCAAGGGTCTAAAATTTTACCGCCCTTGATT
GCGCACAAGGATATTGGACAATTAAAATAGATGAAAGGGATCAATATAAACTAGCC
TTTACATTTGGAAAGCAACAATATGCATGGACCCGATTACCTTTCGGATACATTAAT
GCAGGCCATGAGTTTGCTGTATTCATGCACAAAGCCATGCCTGATGCTGCTGAACGA
GGTACATTATCATATGTGGATGACATATTGATAAAAAGTACAACTTTTGGAGAACAT
ATTGCAGAATTGAGATATGTATTAAATCAACTAAAGAAGGCAGGTGTTAAACTTTCT
TTACAGAAGGCTCAATGGTGCCGATCTAAGGTAAATTTCCTTGGACATGAAATTACT
GCCGAAGGAATCAATCCACAGGAAGATCGTTCCTGGGTATGNATGAATTATTCACGT
AAGTTCATAGACAACTATGCAGAAATTACAAAACCTCTTTTGCAGTTGCTGAAGAAA
GGAGCAAAATGGAAATGGGATGAATGTCACGATCAAGCAATATCTGAACTTAAAAC
TAAGCTTATCCAGGCTCCATGCCTTGCATATCCAGAAGGTGGAAAACCCTTCTATAT
CGAAACAGGTTTTACCAACAAAAGTGTAAGTGCGGTTCTTTTCCAAAAACAAGAAA
AACTGAATAAAATAATCGCTTATGCCAGCAAGTCGTTATCATCAGTTGAAATAAAAT
TTAATAACTGTGAAAAAGCATTATTAGCAACTGTATGGGCTTTGCAATATTTCAGGA
GTTTTATTCAAGGTGAAAAGATAATTGTGGAAACTGCACATCAATCACTACAATATT
TACAGAGTGGTCGATTAAAGGATGGAAACCTGTCAAATAGTAGGATAACTGCATGG
ACAATGTCCCTTATGGGATGGCCATTGGAGATCAGATATAAACAGGATACTAAAAA
TCCAGTTGCTCAAGGATTAGCTGAACTCCATGATTGTATTAATGTTGAACACAAAGA
TGAGTCACCAGAAAATGACTTTCTGGAGGAACAATCACTATCTCCATATAAATCATA
TGAAGAAGAATACTGCAAATCATTACCATGTGTATATGTTGATGGCTGTTCTTTCCA
TACAGAAGAAGCTGAAAAGATGCTAGTTGCAGGTATTGGGATAGTATGGAATAATA
TATTCCCAGAGATATCCACAGGATACAAGATTGGTTCCAAAAGTAGCCAAGTCGCA
GAGCTTGCTGCTGTGTACAAAGCCGTACAAATGGCTATTGAATATGATATTAAAGAA
TTTGTTTTAATTACAGATTCCGATTATGTTCGGAACAGCTTTGTGGAATATTTTCCCG
GATGGAAAAGATCCAATATGATGAGAAGTAATAAGAAACCAGTCAAACATGGTAAG
TTGTTTTGTAAAATCGATGAATTAGTTACTACCCACGGCCTTAACATATATTGGAAG
AAGGTAAGGGGTCATTCGAAACATCCAGGTGCAGATAAAGATGGAAATGACTTAGC
AGATTCCCTTGCCAAGAAAGCAGCCATGGAGGGAGAGATCTTGGATATTGATGACC
TCATGGGAACTATCCAAGTGAATGCCACAACCAGGAGTCAGACCCAAAAGGAATCT
CAACCTAATGTGGCAATGTGGAGCCAAGATTCCCCTAGCGAAGATCTCATTGAGAGT
CAGAAAGGGGATCCTATTATTGGTTTATTTTATAAATATATTCAAAATCCAAAGGAT
CATCCCATAACCATGGAAGAATGTGCTGATAAAGAGGAGTTACGTATTTTGGTGAA
AGGCGAGTCCCAATTTTCATTACAAGATGGATTGCTGATACGAACCTCGAAAAATGG
TATTACACAATGGGTAGTACCCGCAGCATATCGGGGTTTGATGTTACAACATGCTCA
TGATGCCCCGACAGCTGGTCATCGAGGAGAGAAAATAACATATGAATTGTTAAGGG
ATTATGCTTATTGGCCACACATGTTACAGGATGTCAGGACTTATTGTCAAGGATGTT
TGATATGTCCTCAATTCCAGCCACAAGGACCCAATCACCGGGCACCGCTGATGAAA
AGAGGAATATCCATGCCATGGTCAGACCTCCAAATTGATTTCATTGGACCAGTAACA
ACATCCTCAAAAGGAAACCGGTATATGCTAACTGTAACTTGTTTATTTACAAAATGG
GTAGAATGTTTGCCATGCAAGACGTGTAGTTCTAGTGTATGCGCATCACTGCTCATT
AACCATATATTTTCCAGATTTGGTCTTCCCCAACGTATTGAATCTGACCGCGGAAGT
CACTTCACCAGCGAAGTAATGACTAAGATGTGGGAAATTCTAGGAGTAAAAAGGAA
GCTACACATAGCTTACAGACCAGCTTCTAGTGGTGGAGTAGAGCGTTACAATCAATC
AATCATAAACATTCTGAAAAAGTTTGTCAAAGAATCTGGTAAAGACTGGGATGTCA

AATTACCACTAGTTCTCATGGCTATTAGAGCCACCCCGAGTGCGGCAACCAAAATTT
CTCCATTTGAGATGATTACGGGTAGAAAAATGGTATTACCTCAACATCTACTTTATA
GAACCACAGATCACAATTTGATGAATGCAATGACTGCACATCAATATGTAGAAAAT
TTACGAAAGCATCTACAGCATGCTTTTGCGTTTGCTCAAAGAAATATTGAGAAAGCT
GCAGTAAGTACGAAAACGTACTATGATCTAAAAACAACTCAAAAGGAATATCAGGT
TGATGATAAAGTTTATCTTTATAACTTTGCCAGAGACCAAGTAAAGGAAAGGAAATT
CTTACCTTCCTGGAAAGGTCCATATATAATCACTGATAAATTATCTCCTGTGGTTTAT
AAAGTTAAAATTCCTAAGGGGGATGATTTTGTTGAAAAATGGGTGCACATTAACCA
ATTACGTGTTTGTCATCCTAGCCAACAGCTTCGGAGAATAGAGCAAGTCGAGAGCTG
A
>Loki-Xtr consensus sequence Spc7-Gag-Pol polyprotein
MSVASSENGNTPEQMIHRIVEFIYTRVKFDIGMDVWINDLKRNARADLNDIWATDGLVE
RYLTCMNLASNSSDKEKKLICALPPVLYALMEVDTLSKARKENVEKLRSQLHEIESDIQH
MKSDKNQLEIDISNKEAFIRDLETDAERSRIAYCKLRKDFEELQQQHIKSQQSGVYGVER
SPDPPIITPQYKNWAPYISRAKPIDNVDSEAQDVLGRLAEACQAVTTLIGNHRTRVLRDS
SSDTSPERNVRNVRNQTPMSNVSHRSSVSNTDASEVESDERAWGSGPFPNRGKNTHKR
GSKGKNPQSPKEKQNSASIIKFLNTAIPKFSKKGSSQIANHLELYESTMDSLKLSEADKIR
FLPWAFDDRYRHYFSSFKERGITKWQSVLHEIKSEFGPYRTTTAAKREIYKLTCRSNQSP
REFLSVLKNAYGLAYRNPDWESVEFKQVFYEALPTQIKLSLAHDLDFENPLERLVTSAT
LLFNISEGQNLNDRKFKKFPEHNVDESGVKPHLNFESQPKKIPSATGPNQQNQKQLNPR
QTKPQIAKGSEGNSSGSGSQGYRPRFNQGYQGYRGNQRYGGYQGYRGNQRYGGYQGY
RDPQGYRRWNNRPRQQWRKGQESPDSPRGKSPTGNQNGPRNQNPGRPSRFDQLADQV
SKLTDIVSKLSQVSAGKQPEVFFRGKSGAKLRQIKEAEPNCYMEMDRKELNTVNDHVL
HCPDPEFRAGKGKSNISSVLQPNKDNSYNPVLFEPVCEVPSVQSVPLIESIGTQMPCRQN
KEEQSPNILTLQRDHFLNQQMIKHSNFLCNLLQLDGRYFIDICLQDGCVGPIQGLLDTGS
QVTILSFSYFQQILESSKNKPRLTVFDGSLVSVGGDSLQVKGKSWLTYKIGKKTIRHPTLI
VDLPYDRLIIGIDLLKRLRSIIDFINEAIWTQVKTPISFERSHNAKPQNSCHMIEERPNSVEI
HFRNSQVPDVSILKNGKPEESINTSFHIINIQKDKIEKIILEDDVLTISLKGGNREVAGITRHI
QSLVKIEKTNDTMLVPVQVNNIARVKYAKLDLKSEVSYISLGLLKQVTNPQMIKFSSLQE
WVYDLDKDDQSQNVIAKCILSISLGNKSAEHVFNVLKEPQYQMYLGNDIIHRFAIQVDLI
NNVLWSRLSGNPEEFQDKEQALRWGQRIPYAVDILVAGKVKVPGGCKSFLLPIRVKKG
QKLRNPDALICLSNRMQQFGLKVKPNPIINIHQDPLYMVIDNMSPYNIMLQEDTMIGLAI
DSEYYTFGFQNDIIGLIPDEYLTEEQMVDQHCFSTPEGLFSIHSVYPFNPEDGTCYVEETS
LIFNHEIEKGEYESFQQGKDQIYHTPDDLTSKLEEAYEVNQPEIFPEFQERVEEQLSIADA
CSDESERQQLKELFWEFQDMFAKDSYDCGETNLHVTRIQTDPDAPPVFIKQYRLPLAAY
ESLSEIVKNLEKKGIIRPVHSSFNHPILGVLKPNGQFRLCSDLRQLNKQVYMSGWPVPYI
DQCLAQIQGSKIFTALDCAQGYWTIKIDERDQYKLAFTFGKQQYAWTRLPFGYINAGHE
FAVFMHKAMPDAAERGTLSYVDDILIKSTTFGEHIAELRYVLNQLKKAGVKLSLQKAQ
WCRSKVNFLGHEITAEGINPQEDRSWV?MNYSRKFIDNYAEITKPLLQLLKKGAKWKW
DECHDQAISELKTKLIQAPCLAYPEGGKPFYIETGFTNKSVSAVLFQKQEKLNKIIAYASK
SLSSVEIKFNNCEKALLATVWALQYFRSFIQGEKIIVETAHQSLQYLQSGRLKDGNLSNS
RITAWTMSLMGWPLEIRYKQDTKNPVAQGLAELHDCINVEHKDESPENDFLEEQSLSPY
KSYEEEYCKSLPCVYVDGCSFHTEEAEKMLVAGIGIVWNNIFPEISTGYKIGSKSSQVAE
LAAVYKAVQMAIEYDIKEFVLITDSDYVRNSFVEYFPGWKRSNMMRSNKKPVKHGKLF
CKIDELVTTHGLNIYWKKVRGHSKHPGADKDGNDLADSLAKKAAMEGEILDIDDLMGT
IQVNATTRSQTQKESQPNVAMWSQDSPSEDLIESQKGDPIIGLFYKYIQNPKDHPITMEEC

ADKEELRILVKGESQFSLQDGLLIRTSKNGITQWVVPAAYRGLMLQHAHDAPTAGHRG
EKITYELLRDYAYWPHMLQDVRTYCQGCLICPQFQPQGPNHRAPLMKRGISMPWSDLQ
IDFIGPVTTSSKGNRYMLTVTCLFTKWVECLPCKTCSSSVCASLLINHIFSRFGLPQRIESD
RGSHFTSEVMTKMWEILGVKRKLHIAYRPASSGGVERYNQSIINILKKFVKESGKDWDV
KLPLVLMAIRATPSAATKISPFEMITGRKMVLPQHLLYRTTDHNLMNAMTAHQYVENL
RKHLQHAFAFAQRNIEKAAVSTKTYYDLKTTQKEYQVDDKVYLYNFARDQVKERKFL
PSWKGPYIITDKLSPVVYKVKIPKGDDFVEKWVHINQLRVCHPSQQLRRIEQVES*
>Loki-Xtr consensus sequence *env* gene
ATGCCATCCAAGATGATCGCTATTTTATATGCAGTCCTGATCTTGGGGAAAGAGTTA
CACGCTGAGCGGATTGTGATGCCAGGTCCATCATCCGGGATCATGCTACAGGATACA
GCGGGATTCATCATGACCAACAAGAGAATCTTGTCCCAGAAGATTTATATGAGCCTA
GATCCACGTTGTGTCATCGAAAGACAAGTCAACGTATCCAATATTCGATCACCGGAG
ATTCAAACCTGGTACCAGCTGCATATTGAATATTCACAAAAGCGGATTACTCAAATC
CTGGAGCAAACCCGGAAAACTTTAATCAGAGAGCCTTTTTCGTCAAACCAAAGACC
AAAACGGTTCATTTCTGCTATTGTTGCTGCTATAATCTTTGCAGTAGTAGGAAACGT
GTTTGCAACCGGAGTAACAATCGCTAATTCCATATCCCTTAAAACCATGGAATCTGA
GATTAATTCATTAAAAACTAATTTACAACATATTCATGTGGAGATGGAAAGACAAA
AGAAAGATTTAACAGACTTGTATACCGTTGTGGAAGACACTGTCATTGCTACGAACT
TGCACTCTAAATTATTAAACCACTCGATAAATCTTCACCAAAGTCATGAACTGTTTA
AACAAGAACTGATTTTGCTAAAGGAGCGTATCGACTTTCAACTTCAAGTTTTCCATG
ACGAGGTGCAAAGTGGAATGCAAGACCTGAGAGAAGGGCGAATTCCTCTCTATTTT
GTATCATACGAGATTATTCGCAAGATGTTACAAAATGCAGATGGAGATAAAGTAGA
GGATATCCAAATGAACTTAGCGTTTAATATGGGAGTTGCCATACCATTATACATTGA
TCCTGTTAAAATGGAAATCTGCTTTTTACTATCAATTCCAGTGGTCAATAGGAACAA
TATATTTCAAATGAAAAAAATCCACAATGTAGGTACATGGAGAGATAATATTCATGT
TAAGATCCAAACCCCTGATATCGTGGCATATCAATCTTGGAGACCGGACTGGCTATC
TGTTCCTATTCTAAATAATTGTGCAATTTTTAGAGATAATAACTTTCAATGTGAAGGT
GACCCATTTAATTATGATACAACTAATGCTTTGTGCGGACTTGACTTTGAGAAACAT
GGATCAGAAAAATGCCAGGTAACTTTGTCTAAAATCGAGGAGAAAGATTGGGTGAA
TGCAGTTTTAATAAAAGACAAATGGCTCGTTAGTACTCGTTTACCTGACCTAACTGT
TTTTTCTAAGGACCACATTATGGCAAAAATTATTGCTATACCAACTCATGTATCGATT
ATCAAAGTACCCCCTAATTCCAGAGTTGTAATTGGAAATCGCACATTATATCCTGTA
AACATTGAGGACACTGAAGGAGAGGTGGAAAGCATTGATGCATTCCAAAATCAGGA
GATTACCATAAGTCCGGATTTAAAATTTTTGTTGGACAATAAGCCCAGCCATACCAT
CCGTATGGTTATAAAAAAACATAATATCACCATGGATACTCCACAATTTTCACCAAC
CGATGACGAAATACTGTTTCCTTCTGTTTGGTCATGGGGGGAAAACTATATATGGAG
TATATTGATCACAATTATAGTAGGATGGTTAACTGTAATCACCATCTTAGCAATTCG
TTGGATGAAGATCATCAAAAAGACTCTAGATAAACTATGGCTTAAATCTCGCCGATC
AAGTTTACATAATGAGTATGATGAAGGAAGAGTTTGA
>Loki-Xtr consensus sequence Env protein
MPSKMIAILYAVLILGKELHAERIVMPGPSSGIMLQDTAGFIMTNKRILSQKIYMSLDPRC
VIERQVNVSNIRSPEIQTWYQLHIEYSQKRITQILEQTRKTLIREPFSSNQRPKRFISAIVAA
IIFAVVGNVFATGVTIANSISLKTMESEINSLKTNLQHIHVEMERQKKDLTDLYTVVEDT
VIATNLHSKLLNHSINLHQSHELFKQELILLKERIDFQLQVFHDEVQSGMQDLREGRIPLY
FVSYEIIRKMLQNADGDKVEDIQMNLAFNMGVAIPLYIDPVKMEICFLLSIPVVNRNNIF
QMKKIHNVGTWRDNIHVKIQTPDIVAYQSWRPDWLSVPILNNCAIFRDNNFQCEGDPFN

YDTTNALCGLDFEKHGSEKCQVTLSKIEEKDWVNAVLIKDKWLVSTRLPDLTVFSKDHI
MAKIIAIPTHVSIIKVPPNSRVVIGNRTLYPVNIEDTEGEVESIDAFQNQEITISPDLKFLLD
NKPSHTIRMVIKKHNITMDTPQFSPTDDEILFPSVWSWGENYIWSILITIIVGWLTVITILAI
RWMKIIKKTLDKLWLKSRRSSLHNEYDEGRV*

Dataset S2. RT protein alignment generated by Mafft 7.402 and used to reconstruct the phylogenetic tree in Fig. S3.

```
>Loki-Xla
QYRLPLAAYESLSEIVKNLEKKGIIRPVHSSFNHPILGVLKPNGQFRLCSDLR------
QLNKRVYMSGWPVPYIDQCLAQIQGSK-IFTALDCAQGYWTIKIDERDQHKLAFTF-------
GKQQYAWTRLPFGYINAGHEFAVFMHKAMPD----------------AAERGTLSYVDDILIKSTTFEEHIAELKYVL--------
KLSLQKAQWCRSKVNFLGHEVTAEGINPQKKKVEAIKNTKSPTNLKELRSFLGMMNYSRKFIDNYAEITKPLLQLLKKG
A----KWEW
>Loki-Xtr
QYRLPLAAYESLSEIVKNLEKKGIIRPVHSSFNHPILGVLKPNGQFRLCSDLR------
QLNKQVYMSGWPVPYIDQCLAQIQGSK-IFTALDCAQGYWTIKIDERDQYKLAFTF-------
GKQQYAWTRLPFGYINAGHEFAVFMHKAMPD----------------
AAERGTLSYVDDILIKSTTFGEHIAELRYVLNQLKKAGVKLSLQKAQWCRSKVNFLGHEITAEGINPQ-----------------------
-VMNYSRKFIDNYAEITKPLLQLLKKGA----KWKW
>Loki-Npa
QYRLPLACYDSLAEIIRNLKERGVIRSVHSSYNNPILGILKPNGQWRLCADLR------
QLNKRVYMSGWPVPYIDQCLVQMQGAR-IFTTLDCAQGYWTIKVSPEDQYKLAFTF-------
GKEQYTWTRMPFGYINSGHEFAVFLHKAMPD----------------
ATERGNLTYVDDILLKSTSFEKHVQEIRHVLNQLREAGVKISLQKAQWCRTKVNFLGHEVTSEGLNPQKKKVEAVMKS
KTPTNVSELRSFLGMMNYSRKFIDNYAEISKPLLTLLKKDV----PWMW
>Loki-Rca
QYRLPLASYESLAEIIRKLHERGIIRPIHSSYNHPILGILKPNGQWRLCADLR------QLNKRVYMSGWPVPYIDQCLAQLQ-
SQ-VFSALDCAQGYWTIKVSEEDQYKLAFSF-------GNQQYCWVRMPFGYINSGHEFAVFMHKAMPD----------------
ALERGTLAYVDDILIKSTDMTKHLTELRHILGQLR-
AGVKLSIQKAQWCRTRVNFLGHEVTSEGLNPQKKKVEAVFNCKVPTNLKELRSFLGMVNYSRKFIDNYAEISKPLLHLL
KKDT----EWSW
>Loki-Ape
QYKIPLAAYESIQEILDKLLEKRIVRECNSTYNSPIWGALKPTGKWRLTIDYR------
PLNKQVPLSRWPMIHLDQELAKVKEAR-FFSTVDVANGFWTMTVDPVDQYKLAFSF-------
CNRQYTWNRCPFGYSNSPAEFNIFLHKAMSD----------------
ASARGNLIYVDDILMRSRSFDDHLAEIRHVLGQLAAAGAKLAIAKGQWCRTKVEYVGLTVGANGIEPKAERIRAIQDIM
APVNVSKLQSFLGVCNYSRQFIADYAEIARPLTELLGKDK----PFLW
>Loki-Pfl
QYKIPLAAYESIQEILDKLLEKQIVRECNSTYNSPIWPVLKPTGKWRLTIDYR------
PLNKQVPLSRWPMIHLDQELAKVKEAR-FFSTVDVANGFWTMRVDPVDQYKLAFSF-------
CNRQYTWNRCPFGYSNSPAEFNIFLHKAMSD----------------
ASARGNLIYVDDILMRSRSFDEHLAEIRHILGQLAAAGAKLAITKGQWCRTKVEYVGLTVGANGIEPKAERMRAIQDIM
APVNVSKLRSFLGVCNYSRQFIADYAEIARPLTELLGKDK----PFLW
>Loki-Pyo
QYKIPLAAYESIQDILDTLLEKRIIRECNSTYNSPIWPVLKPTGKWRLTIDYR------
PLNKQVPLSRWPMIHLDQELAKVKGAT-FFSTVDVANGFWTMKVDQADQYKLAFSF-------
GGSQYTWNRCPFGYSNSPSEFNIFLHKAMSD----------------
AATRGNLIYVDDILMRSQTFEDHITEIRHVLTQLSNAGAKLALTKGQWCRTKVKYVGLTVGADGIEPQAGRIRAIQDIK
APTRLTELRSFLGVCNYSRQFIEDYAEIARPLNELLHKDK----PFEW
>Loki-Sau
QYRIPLAAYESIQDILDKLLQKQIIRECNSTYNSPIWPVLKPTGKWRLTIDYR------
PLNKQVPLSRWPMIHLDQELAKVKGAC-FFSTVDIANGFWTMKVDPADQYKLAFSF-------
CNRQYTWNRCPFGYSNSPAEFNIFLHKAMSD----------------
AAARGNLIYVDDILMRSCTFETHLAEIRHVLNQLTRAGAKLAIAKGQWCRTKVEYVGLTVGGSGIEPQTGRIRAIQDIKA
PTNVSELRSFLGVCNYSRQFIEDYAEIARPLTELLCKDK----LFQW
>Loki-Ofa
QYRIPLAAYESIQEILDNLLQKQIIRECNSTYNSPIWPVLKPTGKWRLTIDYR------
PLNKQVPLSRWPMIHLDQELAKVKGAC-FFSTVDVANGFWTMRVNPADQYKLAFSF-------
GNRQFTWNRCPFGYSNSPAEFNIFLHKAMSD----------------AAARGNLIYVDDILMRSRTLE-
HLAEIRHVLNQLSAAGAKLAIAKGQWCRTKVEYVGLTVGADGIEPQAGRIRAIQDIKAPTGVSELRSFLGVCNYSRQFIE
DYAEIARPLTELLRKDK----SFEW
>Loki-Bsp
QYRIPLAAYESIQEILDTLLKKEIIRECNSTYNSPIWPVLKPTGKWRLTIDYR------
PLNKNVPLSRWPMIHLDQELAKIKDAH-FFSTVDVANGFWTMRVDPADQYKLAFSF-------
GNRQYTWNRCPFGYSNSPAEFNIFLHKAMSD----------------
```

```
AAARGNLIYVDDILMRSQTFEGHLAEIRHVLSQLLAAGAKLALTKGQWCRTKVKYVGLTVGQHGIEPQADRIQAIRTIK
APNTLSGLRSFLGVCNYSRQFVEDYAEIARPLTELLRNDT----PYIW
>Loki-Cne
QYKIPLAAYGSIQEILNKLLEKNIIRECNSTYNSPIWPVLKPTGKWRLTIDYR------
ALNKQVPLSRWPMIHLDQELAKVKDAR-FFSTVDVANGFWTMKVDPADRYKLAFSF-------GNRQY--
NRCPFGYSNSPAEFNIFLHKAMSD----------------ATERGNLIYVDDILMRSRTFKEHLEEI-
HVLGQLAAAGAKLALSKGQWCRKKVDYVGLTVGPNGIEPQADRIRAIQDIAAPTNLSELRSFLGVCNYSRQFIEDYAD
MARPLTELLRKNT----PFDL
>Loki-Lch
QYRIPLSSYESVQEILHTLMEREIIRECNSTYNSPIWPVLKPNGKWRLTIDYR------
QLNKQVPLSHWPMTYLDQNLTKIAGAK-YFTTLDIANGFWTVKVDPKDQYKLAFSF-------
GKRQYTWNRCPFGYLNSPADFNIFLHKAMPD----------------
AEERGNLIYVDDILIKSNSASDHLDEIRYVLTQLKNAGAKLSLTKGQWCRKRVNYLGYDVGPDGLRPQQNRIKAIQAIR
SPTNITELRSFLGICNYSRQFIENYAEISQPLTKLLRKNE----PWVW
>Loki-Str
QYKIPLAAYAAVQEIIDSLLAKRIIRECNSTYSAPVWPVLKPTGKWRLTIDYR------
QLNKLVPLSRWPMTQLDQELPKVANAK-YFSTVDVANGFWTMTVDPRDQHKLAFSF-------
SNKLFTFNRCPFGYANSPSEFNIFLHKAMPD----------------
AASRGTIIYVDDVLMRSETWSHHLNEMDHVLTQLGTAGAKLAIMKGQWCRTKVNYVGLLVGAEGILPQSNRIQAVRNI
KTPTNLHEVRSFLGVCNYSRQFIENYADLSKPLTHLLQKDT----PFVW
>Loki-Tfu
QYKIPLESYEPVQEIINSMLEKGVIRLCNSTCSAPIWPVLTPNGKWRPTIDYR------
KLNQQVSLSRWPMTQLEQEIPRVKGAT-IFSTLDVASGFWTIPVHPDDQHKLAFMF-------CNR-
YTFTRCPFGYANSPAEFNIFLNKACPD----------------
ARVRGNLIYVDDVLMKSTTMTDHLKEIDHVLQQLTKAGAKIALHKGQWCKTKVNYVGPLIAPQGIERQTNRIQAIQNI
KPPANISELRSFLGVCSYSRQFIKDYSDIARPLMALLKKDS----PFTW
>Loki-Lro
QYKIPLESYEPVQDIITSMLEKGVIRPCNSTYSAPIWPVLKPNGKWRPTIDYR------
KLNQQVPLSRWPMTQLEQEIPRIRGAT-IFSTLDVASGFWTIPVHPDDQHKLAFRF-------
DNRQYTFTRCPFGYANSPAEFNIFLNKACPD----------------
ARVRGNLVFVDDILLKHTTVADHLKEIDHILNQLTTAGAKIALHKGQWCKIKVNYVGLLIGPQGIEPQTNRIQAIQNIKPP
ANVSELRSFLGVCNYSRQFIENYSDIARPLTALLKKDC----PFIW
>Loki-Dcl
QYKIPLESYEPVQEIIDSMLEKGVIRPCNSTYSAPIWPVLKPNGKWRPTIDYR------
KLNQQVPLSRWPMTQLEQEIPRIRGAT-IFSTLDVASGFWTIPVHPDDQHKLAFTF-------
SNRQYTFTRCPFGYANSPAEFNIFLNKACPD----------------
ARVRGNLIYVDDVLMKSMTVADHLKEIDHVLHQLTAAGAKIALHKGQWGKSKVNYVGLLIGPQGIEPQSSRIQAIQTIK
PPVNVSELRSFLGVCNYSRQFIENYSDIARPLTTLLKKDC----PFAW
>Loki-Sri
QYKIPLASYEPVQEIIDNLLEKGIIRPCNSTYSAPLWPVLKPNGKWRLTIDYR------
KLNQQVPLSRWPMTQLEQELPKIRDAK-YFSTLDIASGFWTIPVHVADQHKLAFTF-------
ANRQYTFTRCPFGYANSPAEFNIFLNKACPD----------------
ANDRGTLIYVDDILMRSSTLDTHLEEIDHVLEQLTAAGAKISLSKCQWCKTKVNYVGLLVGPRGVEPQINRVQGVTNIK
APTNVSELRSFLGVCNYSRQFIENYADLAKPLTELLRKDA----LFVW
>Loki-Sme
QYKIPLASYEPVQEIIDDLLHRGIIRPCNSTYSAPVWPVLKPNGKWRLTIDYR------
KLNQQVPLSRWPMTQLNQELPKVKDAK-YFSSMDIASGFWTIPVHIADQHKLAFTF-------
AGRQYTFTRCPFGYANSPAEFNIFLNKSCPD----------------
ASERGTLIYVDDILIRTPTMKGHIDEIEHVLDQLTTAGAKISLSKCQWCKTKVNYVGLLVGSNGVEPQTNRVQGVMNIK
TPTNISELRSFLGVCNYSRQFIPDFADIAKPLTDLLKKDS----LFTF
>Loki-Oja
QYKIPLASYEPVQEIIDDLTNKGIIRPCNSTYSAPIWPVRKPNGKWRLTIDYR------
RLNQQVPLSRWPMTQLEQELPRVKEAK-YFSTLDVASGFWTIPVHPDDQHKLAFTF-------
ARQQYTFTRCPFGYANSPAEFNIFLNKACPD----------------
ATQRGTLIYVDDILIRNATLEDHIEEIEHVLKQLTSAGAKISLSKCQWCRTKVNYVGLLVGPTGVQPQVGRVQGLKDIAA
PTNLKELRSFLGVCNYSRQFIEDYADIAKPLTELLKKDV----PFRW
>Loki-Ame
QYKIPLASYGPVQEIIDDLLDKGIIRPCNSTYSAPLWPVLKPNGKWRLTIDYR------
RLNDQVPLSRWPMTQLEQELPRVRDAK-YFSTLDVASGFWTIPVHVEDQHKLAFTF-------
AGRQFTFTRCPFGYSNSPAEFNIFLNKACPD----------------
ARERGTLIYVDDVLIRNNSLDAHLEEIDHVLDQLTKAGAKISLAKCQWCKTKVNYVGLLVGPDGVLPQPCRAQGIVDIA
EPKTIHALRSFLGVCNYSRQFIENFAELAKPLYQLLKQDV----PFIW
```

```
>Loki-Bpe
QYKIPLASYEPIQEIIDDLLTKGVIRPCNSTYSAPIWPVLKPNGKWRLTVDYR------
QLNTRVPLSRWPMTQLGQDLPKVRDAK-YFSTLDIASGFWTIPVHPDDQHKLAFSF-------
ANRQYTFTRCPFGLANSPAEFNIFLNKACPN-----------------
AASRGTLIYVDDILIRTTTLDDHLKEIDHVLDQLTSAGAKISLSKCQWCKTKVNYVGLL--------
RSSRIQGITNLKVPANVSELRSFLGVCNYSRQFVEHYAELAKPLTDLLKKDE----TFSW
>Loki-Lca
QYKIPLAAYESIQETLENLLKKNIIRECNSTYNSPIWPVLKPTGKWRLTIDYR------
QLNKTVPLSRWPMAEIDHGLNQIKGAK-VLTTMDLANGFWTMPVKETDQYKLAFTF-------
DRIQYTWNRCPFGYSNSPADFNIFLHKAMGD-----------------
AKERGTIVYVDDILVKDSSWQEHLKSLQHTLEQLKEAGAKISIQKGQWARKRVDYLGFQIGTEGLLPQTNRIEALLALK
SPNTVPTLRSFLGICNYLRQFVDDYAGITRPLVKLLQKDE----PWEW
>Loki-Pma
QYKIPLAAYESIQETLENLLKKNIIRKCNSTYNSPIWPVLKPTGKWRLTIDYR------
QLNKTVPLSRWPMAEIDQNLNQIKGAK-ILTTMDLSN-FWTMPVKEADQYKL--TF-------NGIQYTWNRCPFGYSNSPAD-
NIFLHKAMGD-----------------AKERGTIVYVDDILVKDSSWEEHLESLQHTLEQLE--
AAKISIQKGQWARKRVDYLGFQIGT-
GMLPQTKRLEALLALKPPTTVTTLRSFLGICNYLRQFVDDYAGIARPLVKLLQKDE----PWEW
>Loki-Aro
KYRIPIAAHDPVQDTITEILHAGIIRKCNSTYNSPIWPVLKPSGKQSLTVDYR------KLNKQVPLSRCLWPTWMKPLQK-
SRMP-HTSQADISNGFCTIPVREQDQHKLAFSF-------NNVQ-TWTVTPFGYANSPAEFNIFFHKAIPD-----------------
VNSKNIAIYVDDILMWSDSWGEHQQLLSYVLTRLSQAGAKISLDKGQWCKRAVNYLGFQVGGDGIRP---
RTEAPKAIPAPTTLKQLCSFLGICNFSQHFVADYAKISQPLTRLLKKDV----PWTW
>KoRV
QYPMSKEAREGIRPHIQRFLDLGILVPCQSPWNTPLLPVKKPTNDYRPVQDLR------
EVNKRVQDIHPTVPNPYNLLSSLPSHT-WYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKG-
NTGQLTWTRLPQGFKNSPTLFDEALHRDLASFRA-------------
LNPQVVMLQYVDDLLVAAPTYRDCKEGTRRLLQELSKLGYRVSAKKAQLCREEVTYLGYLLKGGKRWLTPARKATV
MKIPTPTTPRQVREFLGTAGFCRLWIPGFASLAAPLYPLTREKV----PFTW
>MuLV
QYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPTNDYRPVQDLR------
EVNKRVEDIHPTVPNPYNLLSGLPSHQ-WYTVLDLKDAFFCLRLHPTSQSLFAFEWRDPEMG-
ISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRI-------------
QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM
GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGT----LFEW
>FLV
QYPMPHEAYQGIKPHIRRMLDQGILKPCQSPWNTPLLPVKKPTEDYRPVQDLR------
EVNKRVEDIHPTVPNPYNLLSTLPSHP-WYTVLDLKDAFFCLRLHSESQLLFAFEWRDPEIG-
LSGQLTWTRLPQGFKNSPTLFDEALHSDLADFRV-------------
RYPALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGYRASAKKAQICLQEVTYLGYSLKDGQRWLTKARKEAILS
IPVPKNSRQVREFLGTAGYCRLWIPGFAELAAPLYPLTRPGT----LFQW
>REV
QYPITLEAKRSLRETIHKFRAAGILRPVHSPWNTPLLPVRKSTSEYRMVQDLR------
EVNKRVETIHPTVPNPYTLLSLLPDRI-WYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEG-
ESGQLTWTRLPQGFKNSPTLFDVALNRDLQGFRL-------------
DHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIP
VPKTKRQVREFLGTIGYCRLWIPGFAELAQPLYAATRGGND---PLVW
>WDSV
QYPLPKDKTEGLRPLISSLENQGILIKCHSPCNTPIFPIKKARDEYRMIHDLR------
AINNIVAPLTAVVASPTTVLSNLPSLH-WFTVIDLSNAFFSVPIHKDSQYLFAFTF-------
EGHQYTWTVLPQGFIHSPTLFSQALYQSLHK-----------------
IKSSEICIYMDDVLIASKDRDTNLKDTAVMLQHLASEGHKVSKKKLQLCQQEVVYLGQLLTPEGRKILPDRKVTVSQFQ
QPTTIRQIRAFLGLVGYCRHWIPEFSIHSKFLEKQLKKDTA--EPFQL
>Xen-1
QYPLKPAQSESINKQLQTLLENGAIKRQSSPCNTPLFPVKEKKDKYRLVQDLR------
AVNEATVMETPLVSNPHTILSGIPSAT-HFRAVDLTNAFYSIPLREDCQYLFAFTH-------
ERQQYVWTVLPQGAQNSPTHFSLALTSILDS-----------------
HPEITLLQYVDDLLLCAPDLPTCEASSTDLLSFLADQGCKASKEKLQWCQTTVVFLGQCISQGTRHITEGRIKTLQDIPLP
KGHKPLHAFLGLISYCRSWIPEASLLMQPLYDVLKSD-----PFTM
>SnRV
QYPVPDASHASIKETVEALLEQGVLRKCNSTVNSAIWPVGKPDGSWRLTIDYR------
PLNSAVSCPYPTVASTPELFAKLKKYQ-VYSSLDISNGFWSIRLEEECQYLFAFTF-------
```

```
DTQQYTWTRLPQGFHASPGIFHQALYNGLAS-----------------
SQGCKLLQYVDDILLMSEDRDHHLRSLAILLQGLKDLGVKINPKKSHFCKDQVQYLGVNVGADTRSLIDARSQLIRTLDI
PLTVQGLRSALGLFNFCRAWIPEFSRKTQSLYDMLKGDCK---KLKW
>ZFERV-2
QYPINKSALPEIKITIEELKRKGIIKEVNAPTNSPIQAVAKPDGGWRLVTNYK------
ALNKQTVPDTRYLINARDATGAAPEGR-LLSKIDLANGFWSVPLKPESRARTAFTF-------
ENKQYVVYNVLPQGFKNSPNAFQAIMMEILKG-------------------
LPIVVYIDDILIVTNDEEEHLKILDETLSRLTKEGFRLNYKKMEIAKDDVEFLGFQITGNERFLSQTTENKLKELRGETTLK
QLQSMLGTLNFVRDLIPGYARHAKPLYNATKGGV-----LTW
>AciFLERV
QYPLNPGAVKEMDLIVRELLTLGVIRQELNPTNSPIQAVKKPEGGWRPVINFK------
ALNRRTVANRASLINPQGTLKTLKLKP-FKSCIDLANGFFSLRLTRGSQGKTAFTH-------
KGRSYVWERLPQGYKNSPNVFQAAVMDVLDG-------------------
LGATIYIDDVFIADDTEEQHLDRLQKIVERLTAAGLKLNLKKCQFGRKRVDYLGFQVSDD-
LGLSEVYRKKIEQIEPPSSLNELQKVLGLCNYVRDYVPGYQRYAKPLYAKLKKPDE--AAWNW
>AliFLERV
QYPLHPEAVAEMDKIVKELHALGIIREEPNPTNSPIQAVKKPEGGWRPVINFK------
ALNRRTVANRASLINPQGALKTLRVKK-YKSCIDLANGFFSLRLARESQGKTAFTH-------
KGKSYVWQRLPQGYRNSPNVFQSAVLEILED-------------------
VGASVYIDDVFIADDTEEEHLKRLEEVVKRISAAGLKLNLAKCQFGQFKVNYLGFQVTTD-
LGLSDGYREKIQQIQPPRCENDLQKILGLCNYVRDHVPNYQKYAKPLYACLKKEAT--KAWSW
>FFV
QYHINPKAKPDIQIVINDLLKQGVLIQKESTMNTPVYPVPKPNGRWRMVLDYR------
AVNKVTPLIAVQNQHSYGILGSLFKGR-YKTTIDLSNGFWAHPIVPEDYWITAFTW-------
QGKQYCWTVLPQGFLNSPGLFTGDVVDLLQG----------------IPN--
VEVYVDDVYISHDSEKEHLEYLDILFNRLKEAGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENITAPTTLKQL
QSILGLLNFARNFIPDFTELIAPLYALIPKSTK--NYVPW
>BFV
QYHINPRAKADIQIVIDDLLRQGVLRQQNSEMNTPVYPVPKADGRWRMVLDYR------
EVNKVTPLVATQNCHSASILNTLYRGP-YKSTLDLANGFWAHPIKPEDYWITAFTW-------
GGKTYCWTVLPQGFLNSPALFTADVVDILKD----------------IPN--
VQVYVDDVYVSSATEQEHLDILETIFNRLSTAGYIVSLKKSKLAKETVEFLGFSISQNGRGLTDSYKQKLMDLQPPTTLR
QLQSILGLINFARNFLPNFAELVAPLYQLIPKAKG--QCIPW
>EFV
QYRINPKAKADIQIVIDDLLKQGVLKQQTSPMNTPVYPVPKPDGRWRMVLDYR------
AVNKVTPAIATQNCHSASLLNTLYRGQ-YKTTLDLANGFWAHPIQESDQWITSFTW-------
NGKSYVWTTLPQGFLNSPALFTADVVDLLKD----------------IPN--
VEVYVDDVYFSNDTEEEHLKTMDLLFQKLQTAGYIVSLKKSKLGQHTVDFLGFQITQTGRGLTDSYKSKLLDITPPNTL
KQLQSILGLLNFARNFIPNYSELITPLYQLIPLAKG--IYIPW
>SFV
QYHINPKAKPSIQIVINDLLKQGVLRQSTSPMNTPVYPVPKPDGKWRMVLDYR------
AVNKTIPLIAAQNQHSLGILTNLIRHK-YKSTIDLSNGFWAHPITEDSQWITAFTW-------
EGKQHVWTRLPQGFLNSPALFTADVVDILKE----------------VPG--
VSVYVDDIYISSPTMEEHFQVLDSIFRKLLETGYIVSLKKSALARYEVNFLGFVISETGRGLTSEFRERLQEITPPTTLKQL
QSILGFLNFARNFVPNFSELVQPLYQLISTASG--NFIQW
>CoEFV
QYPINTKAIPSIQVVINELLEQGVLVKQTSPTNMAVLPVPKPDGTWRLVLDYR------
ALNKHSEPVRAQNQHSSGILANIERKA-YKSSVDLANGFWSHPIREEDWPKMAFTW-------
CGFQYLWTRLPQGFLNSPALFSADVISLVGQ----------------LPG--
VYCYVDDIYLTHDTEEEHLKILDQVLEILIKAGYVINIKKSKLCRKVVDFLGFSLSDEGRGLSDQYREKLAAIKPPQTLRQ
LQSVMGLLNFSRLFVKGFSELAKPLYDL-------------
>Amn-san
-----------MKEYISENLQRGFIRPSTSPAGAGFFFVEKKDGGLRPCIDYR------GLNKITVKNRYPLPLISELFDQLKGAK-
IFSKLDLRGAYNLIRIREGDEWKTAFNT-------RDGHYEYLVMPFGLCNAPAVFQEFVNDIFRD----------------
LLGKSVVVVYLDDILIFSQDLETHRSQVKEALSRLRENFLFAKLEKCTFEVPKISFLGYIISSRGFEMDPAKVSAIQKWPLPQ
STKAIQRFIGFANYYRQFIKDFS---------------------
>Sushi-ichi
------------------LASGIIRPSSSPLAAGFFFVAKEDGGLRPCIDFR------KLNNITVKNKYPLPLMSSTFEPLTHAR-
VFTKLDLRNAYHLVQIRKGDEWKTAFNT-------HLGHFEYLVMPFGLSNAPAVFQELVNDVLRD----------------
MINVFVVVVYLDDILIFSRTMEEHHQHVRLVLQRLLENRLFIKAEKCIFHSASVGYLGYIVEEGRVRADPAKIQAVVEWPR
PTDRTQLRRFLGFAGFIRRFIKGFA---------------------
>CRM
```

```
-------------------LDKGYVRESLSPCAVPVILVPKKDGTWRMCVDCR------AINNITIRYRHPIPRLDDMLDELSGAI-
VFSKVDLRSGYHQIRMKLGDEWKTAFKT-------KFGLYEWLVMPFGLTNAPSTFMRLMNEVLRA----------------
FIGKFVVVYFDDILIYSKSMDEHVDHMRAVFNALRDARLFGNLEKCTFCTDRVSFLGYVVTPQGIEVDQAKVEAIHGW
PMPKTITQVRSFLGLAGFYRRFVKDFS----------------------
>Reina
-------------------LEQGVIQHSSSPFASPVLLVKKKDGEWRLCVDYR------RLNAHTVKNRYPMPVFDEIVDELCGTK-
IFTKLDHRSGYHQIRIKEGDEFKTAFQT-------HNGHYEYRVMPFGLTGAPATFQDFMNKILTP-----------------
FLRKCVVVFLDDVLIYSRDMEEHVLQVKQVFQKLKDHQLKLKLSKCRFAQTTLEFLGHIISAEGIATDPEKVQVIRDWPI
PNNVKEVRSFLGMAGYYRRFVAHYA----------------------
>Del
-------------------LNKGFIRGSTSPWGAHVLFDPKKDDSKRMCIDYK-------LNSVTVKNKYPLPRIDDLFDQLNGA--
YFSKIDLRFRYHQLRIR-ADIPKTAFRT-------RYGHYEFLVMPFGLTNVPTAFMNLMNRVFRE-----------------
YLDKFIVVFVDYVLIYSRTQKDHEHHLRISLQLLRNNQLYAKLSKCEFWMEKVKFLGHVVSREGIVVDPVKVKAVMN
WELPKNIFEIRSFLGLAGYYRRFIKGFA----------------------
>Galadriel
-------------------LDAGLIQPSKAPYGAPVLFQKKQDGTMRMCVDYR------ALNKATIKNKYSVPLVQDLMDRLSKAC-
WFTKLDLRAGYWQVRIAEGDEPKTTCVT-------RYGSYEFLVMPFGLTNAPATFCNLMNNVLFD----------------
YLDDFVVVYLDDIVIYSRTLEEHVNHLSLVLSQLRKYTLYVKMEKCEFAQQEIKFLGHLVSKNQVRMDPKKVQAIVDW
QAPRHVKDLRSFLGLANYYRKFIAGYS----------------------
>Maggy
-------------------LKKGFIRPSSSSVASPVLFVKKQGGGLRFCVDYR------ALNNITVKDRYPLPLVRETLNNLAGMK-
FFSKIDIVSAFNNIRIKKGEEYLTAFRT-------RFGLYESLVMPFGLTGAPATFQRYINDSLRE-----------------
YLDVFCTAYLDDILIYSRTRTEHEEHLKLVLEALRKAGLYANAAKCEFFVTETKFLGLLVGVEGVKMDPEKITAVLDW
QTPKKLTDVQAFLGFGNFYRRFIRDFA----------------------
>Skippy
-------------------IRKGYIRPSKSSAGFPVMFVPKPNNKLRLVVDYR------QLNEITEKDRTSLPLITELKDRLFGKK-
WFTALDLKSAYNLIRIKEADEWKTAFRT-------KYGLFEYLVMPFGLTNAPAVFQRMITNVLRE-----------------
YLDIFVVCYLDDILIFSDTEEEHTEHVHKVLKALQDANMLVEPTKSHFHQSQVTYLGHEISHNEIRMDRRKIAAVAEWK
VPTSVKETQSFLGFANYYRRFIKDFA----------------------
>Pyggy
-------------------LDKGFIRVSSLLAFLLVLFAKKLEGGLRLCIDYW------ALNAITRKNRYLLPLIRETLNNISKAK-
WFTKLDVIAAFYKIRVAEGDEWKTAFRT-------RFSLYEWLVTPFGMANSLSTFQRYINWTLRE-----------------
YLDEFCSAYLDDVLIYTDSLEQHQDHVRKVLRKLQESGLNVDIKKCEFGVKSTKYLGLIIDAEGIRMDPEKVKAIMEWE
-PLRLRAYVHFWDSQTSTGDLSRDFS----------------------
>Gypsy
-------------------LKDGIIRPSRSPYNSPTWVVDKKNPNKRLVIDFR------KLNEKTIPDRYPMPSIPMILANLGKAK-
FFTTLDLKSGYHQIYLAEHDREKTSFSV-------NGGKYEFCRLPFGLRNASSIFQRALDDVLRE-----------------
QIGKICYVYVDDVIIFSENESDHVRHIDTVLKCLIDANMRVSQEKTRFFKESVEYLGFIVSKDGTKSDPEKVKAIQEYPEP
DCVYKVRSFLGLASYYRVFIKDFA----------------------
>Tom
-------------------LEQGLIRESNSPYNSPTWVVPKKKAKYRVVIDYR------KLNEITIPDRFPIPNMDEILGKLGKCQ-
YFTTIDLARGFHQIEMDSESIQKTAFST-------KRGHYEYVRMPFGLRNAPATFQRCMNNILRP----------------
LINKHCLVYLDDMIIFSTSLDEHLNSLQLVFEKLSESNLKLQLDKCEFLKKEATFLGHIVTPDGIKPNPLKVEAIASYPIPT
KVKEIRAFLGMTGYYRKFIPSYA----------------------
>Zam
-------------------LDQQIIRSSHSPWSAPVWVVPKKNRKWRLVIDYR------KLNDKTISDRYPIPNINDILDSIGKAK-
YFSTLDLTSGFHQIEMNPKDIAKTAFTV-------EGGHYEFTRMPFGLKNAPATFQRVMDSVLGD-----------------
LNGTICLFYLDDIIIFSPSLQKHLLDIKMVFEKLRAANFKLQPSKSEFLRKEIEFLGHIVTQDGVKPNPNKISAIKKFPCPTN
RRAIKSFLGLLGYYRKFIRDFA----------------------
>Cer3
-------------------IQSGRIVESNTPWTSPIVLIKKKNGSLRVCLDFR------KLNEVTVPDNYPLPRIDSIIEQIGGSK-
YFTSLDMANGYLQLRLDAESSDKCGFIT-------ENKVYAYTHLPFGLRSAASYFQRDLKQVLG-------------------
LEKEVTVYIDDVLIFSKTFEEHLNSLRKVLERFRKFNLKVSPNKCEFFKQSITFLGHEINADNYKPNRINVDAIVNLPTPRN
VGDVRRFIGMSGFFRKFLPNFS----------------------
>Mdg3
-------------------LENGFIRPSDSEYASPIVLVKKKTGDLRMCVDFR------KLNKMTMKDNYPLPLIDDLLDRMNEKT-
VFTKLDLKNGFFHVHVKKESIKYTSFVT-------PLGQYEWLRMPFGLKNAPSVFQRFVNKIFAD----------------
MIENKVVVYMDDILLATENINEHLETLKEIFKRLVENKLELRIDKCEFMQSSIKYLGFIINKDGIMPNDKGIEAIKNFPIPN
NVHTVQSFLGLCSYFRRFIKDFS----------------------
>Mdg1
-------------------IKDGIVEQSISEYNSPLLLVPKKEKRWRLVVDYR------QINKKLLADKFPLPRIEDILDQLGRAK-
YFSCLDLMSGFHQIELDERSRNITSFST-------STGAYRYTRLPFGLKIAPNSFQRMMTLAFSG----------------
```

```
LTPSQAFLYMDDLVVIGCSEKHMLKNLTDVFKLCRQHNLKLHPEKCTFFMKEVTYLGHKCTDKGILPDDSKYEVIKNY
PKPVNADEARRFVAFCNYYRRFIKNFA---------------------
>TF1
-------------------LKSGIIRESKAINACPVMFVPKKEGTLRMVVDYK------PLNKYVKPNIYPLPLIEQLLAKIQGST-
IFTKLDLKSAYHLIRVRKGDEHKLAFRC-------PRGVFEYLVMPYGISTAPAHFQYFINTILGE----------------
AKESHVVCYMDDILIHSKSESEHVKHVKDVLQKLKNANLIINQAKCEFHQSQVKFIGYHISEKGFTPCQENIDKVLQWK
QPKNRKELRQFLGSVNYLRKFIPKTS---------------------
>Gmr1
-------------------LKLGVIEESRSAWASPIVLVPKPDGTHRFCNDFR------RLNEVSDFDSYPMPRVDELIERLGPAR-
YLSTLDLTKGYWQVPLAPSSREKTAFAT-------PGGLFQYTVLPFGVHGAPATFQRMMDQVLRP----------------
HSSYAAAYIDDIIIHSASWDEHVKHVRAVLNGLRAAGLTANPAKCRLGREETAYLGYRVGRGNVRPQEDKVAAIREWP
QPQTKKQVRSFLGLVGYYQRFIPGYA---------------------
>Woot
-------------------LDLGVIKREASPYASPMTVGKKKDGTVRICLDAR------MINSKMIADCESPPAADELLRRFHEIR-
YMSTIDLRSSYWQIPLSPESRQYTAFLY-------NGRSYTYQVLPFGLKTAVGSFSRAMDVVLGT----------------
EVREFVVNYIDDLLVASETLNEHLEHLRQVFEKLKQARMTINLEKSNFIQKEVKFLGHILTINGIKADPEKISAIRNFPVPQ
KTKHVRAFLGLCNFYRKFCARYA---------------------
>Cer4
-------------------LDLEVLEPITSDWAAPIVVIRKKDGKVRVCADFK------CLNNSLIEEIHPLPTSDDLFGTLQGC--
IFSKIDLKDAYLQIALDSESQKLAVINT-------HKGLFKYRRMTFGLKPAPAKFQKIIDKMIAG----------------LPG--
VAAYLDDVIVSANSLEEHEKVLHELLKRIKDYGFRISPEKCNFAQSEITFLGFIIDKRGRRPDPKKTSVIRSMKAPTDQKQ
LMSFLGAICFYGRFVPKMS---------------------
>Mag
-------------------LAAGVIKPVDSDWATPLVVVRKADGGLRICADYK------VLNKVLAIDRFPVPKMEDLFSNLSGNK-
FFTKLDLSQAYNQIVLSERSSEYTVINT-------HRGLFKYSRLVYGLASSPGIFQKLMVNMFKN----------------VPN--
VVVFYDDILIRNQDLDSHLKSIKEVLDILERYGLKIKRSKCEFMVTEVRYLGFIIDQNGVRVDPEKVKSIATMPHPNNVTE
LKSFIGMVNFYSKFIQDLA---------------------
>SPM
---------------LDRLGKAGVIKKVKSRWAAPIVTVPKSDGSVRLCGDYK------QVNKVLEEDVYPLPTCEDLFANLAGGK-
VFSKIDLSNAYLQLELTDKSKELLTIHT-------HKGLYQYERLPFGVSTAPAVFQSVMDRILGG----------------MKG--
VCCYLDDILISSENEEDHIKTLNEVLGRLQEYGVKAKKEKCSFMVSRVIYLGHEIDGKGVHPTKEKVEGICDAKRPENK
DELRTFIGIVVYYAKFVPNLA---------------------
>SURL
-------------------ERLDVITPVDTDWVSSLVVVMKKNGQLRVCLDPR------DLNRAIKREHYQLPSRAEITAHFAGAK-
YFSKLDASSGFWQIQLDDESSKLCTFIT-------PYGRYKFLRLPFGICSAPEVYHKIVHQMFAH----------------IPG--
VNTMMDDVIVWGTTQQEHDNRLREVLSIARRMNLKLNKDKCEFSVKKLTFIGDLISDQGVQPDPKKVSAILNMERPKC
RKDVQRFLGMINYQGKFIPDLA---------------------
>Kabuki
-------------------QELGICRPSKIAWASPLHIVPKKNGDLRPCGDYR------QLNAITKPDRYPIPRRHDFTYILDDKK-
IFSKLDINRAFHCIDVAPEDIEKTAIIT-------PFGLFEFPKMTFGLRNAPQTFQRFMNNTVLH----------------GLE-
FLFSYLDDVIIASESPTQHNEHLRIVFERFNSYGITINLDKCCFGQPKVEFLGHEVSVNGITPLKNKVEAIINFPKPETVSDL
RRFLGMVNFYRPHLPNSA---------------------
>Tor4a
---------------LDQFLENKLIEPCSSPWNSPSLLVRKKDGRFRLVIDYR------RLNDATLQMHHPLPNLEDSISYLEKSR-
VYSMCDMIKGFHQIDLDEDSKQKTAFSN-------EFGQFQYTRMPMGCKNAPSFFMRIMDKALLG----------------
VKKTEIIAYMDDLLCHSQTEVEHIKILEKLFQILAVNNLRINSKKAAFFTDSVNFCGYDISNGKICPSQDKIEAIQKLKIPRT
KDAAQSLFGALNYHRRFIRNFA---------------------
>Tor1
-------------------EDAGIIEKASSSFNAPLQLVRKSSGGYRICVDMR------SLNNRLAESKWPLPSLAETLESLAGTA-
FFSCVDIRQAFFHMALTDESKHLTAFSA-------LNCQYQFRRLPMGLKISPSVYQMAMKETLGN----------------
DLGNKAVVYLDDVLVTGRTEDEHLEALDVVLDRLRKAGFLLNPDKCILGVKKTTFLGHEVTTEGYYPKTDNLAAIREF
PKPTNKKALRRFIGMTAFYSTLVPKLQ---------------------
>Athila4-1
-----------------------IYPISDSTWVSPVHCVPKKITGHRMCIDYR------KLNAASRKDHFPLPFIDQMLERLANHP-
YYCFLDGYSGFFQIPIHPNDQEKTTFTC-------PYGTFAYKRMPFGLCNAPATFQRCMTSIFSD----------------
LIEEMVEVFMDDFSVYGPSFSSCLLNLGRVLTRCEETNLVLNWEKCHFMVKEGIVLDHKISEKGIEVDKGKVEVMMQL
QPPKTVKDIRSFLGHAGFYRRFIKDFA---------------------
>Diaspora
-----------------------IYPISDSQWVSPVQVVPKKQNNWRVCIHYR------RLNQVTKKDHFPLPFIDQILECLAGKS-
HYCFLDGFSGYMQITIALEDQEKTTFTC-------LFGTFAYRRMSFGLCNAPGTFQRCMISIFSD----------------
FLENCIEEFMDDFTVYGSSFDGCLDSLEKVLNRRIETNLVLNFEKCHFMVEQGIVLGHIISNKGIEVDPAKISVISQLPYPS
CV-EVRSFLGHAGFYRCFIRDFS---------------------
```

```
>RetroSor1
-------------------LDAKVIREVIPEWLANVVLVPKKNGKMRMCIDFT------DLNKACVKDSFPLPRIDTSVDKAAGCQ-
RFSLLDCFSGYHQIWLKKEDEGKASFTT-------PFGTYCYTRMPEGLKNAGATFSRMMGKVLGS-----------------
QLQRNIIAYVDDVVVMSKRKEDHIKDLQETFVNLRSAGLKLNPEKCVFGVSKGKMLGYIISSEGIRANPDKTKAIMSMA
EPSNKKEVQRLTGRIAALNRFISRSA---------------------
>Tat4-1
-------------------LGAGSIVEVKPEWLANPVVVKKKNGKLRVCIDFT------DLNKACPKDSFPLPHIDRMVEATTGNE-
LLSFMDAFSGYNQIPMHKDDQEKTSFIT-------DRGTYCYKVMPFGLKNAGARYQRLVNQMFAP-----------------
QLGKTMEVYIDDMLVKSKKSADHIEHLTACFETLNKYNMKLNPAKCSFGVTSGEFLGYIVTKRGIEANPKQIQAILDLK
SPRNKKEVQRLTGRIAALNIFIARST---------------------
>Cigr-1
-------------------LEEEIIEPSNSPWRAQVLVAREGSKKPRMVIDYS------QVNVFTELDAYPLPSIESIVNKVTQDK-
VYSSLDLRSAYHQIPLKDSDRPFTAFEA-------EGQLYQYKRLPFGVTNGVSAFQRIINRFIQD-----------------
NHLNKVYAYLDDITVTGETLEEHDRNLKKLLVAAKRHNLTINEEKSKLRMTSIDFLGYRISYGDARPDPERLRPLLNLPP
PDTPRELKRIQGMFAYYAKWISRFS---------------------
>CaMV
--KYSPMDREEFDKQIKELLDLKVIKPSKSPHMAPAFLVNNEAGKKRMVVNYK------
AMNKATVGDAYNLPNKDELLTLIRGKK-IFSSFDCKSGFWQVLLDQESRPLTAFTC-------
PQGHYEWNVVPFGLKQAPSIFQRHMDEAFRV-----------------
FRKFCCVYVDDILVFSNNEEDHLLHVAMILQKCNQHGIILSKKKAQLFKKKINFLGLEIDEGTHKPQGHILEHINKFPDTE
DKKQLQRFLGILTYASDYIPKLAQIRKPLQAKLKENV----PWRW
>SbCMV
--PYTIRDVQEFKEECEDLLKKGLIRESQSPHSAPAFYVENHNGKRRMVINYK------
KMNEATIGDSYKLPRKDFILEKIKGSL-WFSSLDAKSGYYQLRLHENTKPLTAFSCP------
PQKHYEWNVLSFGLKQAPSIYQRFMDQSLKG-----------------
LEHICLAYIDDILIFTKSKEQHVNDVRIVLQRIKEKGIIISKKKSKLIQQEIEYLGLKIQGNEIDLSPHTQEKILQFPDEEDRK
QIQRFLGCINYIANFFKNLALERKHLQKKISVKN----PWKW
>ComYMV
--HVTPGDEEAMTRQINLLLQMKVIRPSESKHRSTAFIVRSGTGKERMVFNYK------
LLNENTESDQYSLPGINTIISKVGRSK-IYSKFDLKSGFWQVAMEEESVPWTAFLA-------
GNKLYEWLVMPFGLKNAPAIFQRKMDNVFKG-----------------
TEKFIAVYIDDILVFSETAEQHSQHLYTMLQLCKENGLILSPTKMKIGTPEIDFLGASLGCTKIKLQPHIISKICDFSDKATP
EGMRSWLGILSYARNYIQDIGKLVQPLRQKMAPTG----DKRM
>RTBV
--PYTPADKEVFEKQIKELLDNKLIKKADCRHRTAAFIVRNHSQKPRIVYNYK------
RLNDNMHTDPFNIPHKISMINLIQKAN-IFSKFDLKAGFHHMKLKDDFKDWTTFTC-------
SEGLYTWNVCPFGIANAPCAFQRFMQESFGD-------------------
KFALLYIDDILIASNNEKEHIEHLKIFFNRVKEVGCVLSKKKSKMFLKEVEYLGVEIKEGKISLQPHIVDKIKKFDKKNTL
KGLQAYLGLLNYARGYIKDLSKLVGPLYKKTGKNG----QRIF
>CSVMV
--LYQETDLPEFKMHIEEMIKEGFIEEEDKKYSSPAFIVNKHSGKTRMVIDYK------
DLNKKAKVVKYPIPNKDTLIHRSIQAR-YYSKFDCKSGFYHIKLEEDSKKYTAFTV-------
PQGYYQWKVLPFGYHNSPSIFQQFMDRIFRP-----------------
YYDFIIVYIDDILVFSKTIEEHKIHIAKFRDITLANGLIISKKKTELCKEKIDFLGVQIEQGGIELQPHIINKILEKHTKKNKTE
LQSILGLLNQIRHFIPHLAQILLPIQKKLKIKD----EEIW
>PVCV
--GMNPEHLQLAIKECDELQQFDLIEPSDSQWACEAFYVNKRSGKLRLVINYQ------
PLNHFLQDDKFPIPNKLTLFSHLSKAK-LFSKFDLKSGFWQLGIHPNERPKTGFCI-------
PDRHFQWKVMPFGLKTAPSLFQKAMIKIFQP-----------------
ILFSALVYIDDILLFSETLEDHIKLLNQFISLVKKFGVMLSAKKMILAQNKIQFLGMDFADGTFSPAGHISLELQKFPDTLS
VKQIQQFLGIVNYIRDFIPEVTEHISPLSDMLKKKP----PAWG
>HERV-L
QYCI-PGGIAEISATIKDLKDAGVVIPTTSPFNSPIWPVQKTDGSWRMTVDYH------
KLNQVVTPIAVPVPDVVSLLEQISPGT-WYAAIDLANAFFSIPVHKAHQKQFAFSW-------
QGQQYTFTVLPQGYINSPALCHNLIQRDLDD-----------------
PQDITLVHYIDDIMLIGSSEQEVANTLDLLVRHLHARGWEINLTKIQGPSTSVKFLGVLVQWCACDIVPSKVDKLLHLAP
PTTKKEAQCLVGLFGFWRQHIPHLGVLLWPIYQVTQKAA----SFEW
>MuERV-L
QYCI-PGEIAEITATIKDLKDAGVVVPTTSPFNSPIWPVQKTDGSWRMTVDYR------
KLNQVVTPIAAAVPDVVSLLEQISPGT-WYAAIDLANAFFSVPVHKDHQKQIAFSW-------
QGQQYTFTVLPQVYINSPALCHNLVRRDLDR-----------------
```

```
PQSITLVHYIDDIMLVGPSEQEVATTLDSLVTHMRIRGWEINPTKIQGPSTSVKFLGVQWCG-
ACRDIPSKVDKLLHLAPPTTKKEAQRLVGLFGFWRQHIPHLGVLLRPIYQVTRKAA----SFVW
>Tor2
-----------VKAQIDNMAKDGIIEPCDKGFHSPLVIVRKKTGGLRICSDFKS-----SLNQCLDETTWPLPQMDHLFANIHGHR-
IFTSLDVSKAYWNLLIDPRDRHKTNFTF-------DNKCWMYVRLPFGLKFSGDAFCRSISSMLDR----------------
VSKSNFCNYVDDVLAYSSDCETHVKVLQQIFSACRHFGARLGAKKCTFGQASTAFMGRIISAEGISIPPENMETILALKPP
TTRKQVQSLIGNFCWLKSWVSANL---------------------
>SRV-1
QWPLTSEKLAAAQQLVQEQLEAGHITESNSPWNTPIFVIKKKSGKWRLLQDLR------AVNATMVLMGALQPGLPSP-
VAIPQGY-LKIIIDLKDCFFSIPLHPSDQKRFAFSLPSTNFKEPMQRFQWKVLPQRMANSPTLCQKYVATAIHKVRH-------
-----AWKQMYIIHYMDDILIAGKDGQQVLQCFDQLKQELTIAGLHIAPEKIQL-QDPYTYLGFELNGPKITNQK----
AVIRKDKLQTLNDFQKLLGDINWLRPYLKLTTADLKPLFDTLKGDSNPNS----
>MMTV
QWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLR------AVNATMHDMGALQPGLPSP-
VAVPKGW-EIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRD-------
-----KYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQK-YDNLKYLGTHIQGDSVSYQK----
LQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNGDSNPIS----
>JSRV
----------------------------------------------------------------MMHMGALQPGLPTP-SAIPDKS-
YIIVIDLKDCFYTIPLAPQDCKRFAFSLPSVNFKEPMQRYQWRVLPQGMTNSPTLCQKFVATAIAPVRQ------------
RFPQLYLVHYMDDILLAHTDEHLLYQAFSILKQHLSLNGLVIADEKIQT-HFPYNYLGFSLYPRVYNTQL----
VKLQTDHLKTLNDFQKLLGDINWIRPYLKLPTYTLQPLFDILKGDSDPAS----
>LDV
QWPLTAQKLDAVQNIIQDLLKDGRIIPSRSQWNSPIFVIQKKDSKFRMLHDLR------AVNALIKDWGALQPGTPWP-
GAIPSEW-PVIAMDISDCFFSIPLAERDSERFAFTIPSPNLREPAKRYQWTVLPQGMKNSPYICQQVVAEVIRPIRE----------
--RFRDAVIIHYMDDILIAAAEERQTEVIFEAVKTTCHEKGLKINEAKTQR-APEVSYLGWRVMETGIAP------
LTVQLPETTNLVQLQRLLGIIQWLKPILSLRPEQLQVFYDLLKGKTRPEE----
>RSV
QWPLPEGKLVALTQLVEKELQLGHIEPSLSCWNTPVFVIRKASGSYRLLHDLR------AVNAKLVPFGAVQQGAPVL-
SALPRGW-PLMVLDLKDCFFSIPLAEQDREAFAFTLPSVNNQAPARRFQWKVLPQGMTCSPTICQLVVGQVLEPLRL-----
--------KHPSLCMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQR-EPGVQYLGYKLGSTYVAPVG-----
LVAEPRIATLWDVQKLVGSLQWLRPALGIPPRLMGPFYEQLRGSDPNEA----
>HIV-1
QWPLTEEKIKALVEICTEMEKEGKISKGENPYNTPVFAIKKKSTKWRKLVDFR------ELNKRTQDFWEVQLGIPHP-
AGLKKKK-SVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRK-------
-----QNPDIVIYQYMDDLYVGSDEIGQHRTKIEELRQHLLRWGLTTPDKKHQK-EPPFLWMGYELHPDKWTVQP----
IVLPEKDSWTVNDIQKLVGKLNWASQIYPGIK--VRQLCKLLRGTKALTE----
>FIV
QWPLTNEKIEALTEIVERLEREGKVKRDNNPWNTPVFAIKKKSGKWRMLIDFR------ELNKLTEKGAEVQLGLPHP-
AGLQIKK-QVTVLDIGDAYFTIPLDPDYAPYTAFTLPRKNNAGPGRRFVWCSLPQGWILSPLIYQSTLDNIIQPFIR----------
--QNPQLDIYQYMDDIYIGSNSKKEHKEKVEELRKLLLWWGFETPEDKLQE-EPPYTWMGYELHPLTWTIQQ----
KQLDIPEQPTLNELQKLAGKINWASQAIPDLS--IKALTNMMRGNQNLNSTRQW
>EIAV
QWPLTKEKLEGAKEIVQRLLSEGKISESNNPYNSPIFVIKKRSGKWRLLQDLR------ELNKTVQVGTEISRGLPHP-
GGLIKCK-HMTVLDIGDAYFTIPLDPEFRPYTAFTIPSINHQEPDKRYVWKCLPQGFVLSPYIYQKTLQEILQPFRE----------
--RYPEVQLYQYMDDLFMGSNSKKQHKELIIELRAILLEKGFETPDDKLQE-VPPYSWLGYQLCPENWKVQK----
MQLDMVKNPTLNDVQKLMGNITWMSSGIPGLT--VKHIAATTKGCLELNQKVIW
>HTLV-2
PISFKPERLQALNDLVSKALEAGHIEPYSGPGNNPVFPVKKPNGKWRFIHDLR------
ATNAITTTLTSPSPGPPDLTSLPTALP-
HLQTIDLTDAFFQIPLPKQYQPYFAFTIPQPCNYGPGTRYAWTVLPQGFKNSPTLFEQQLAAVLNPMRK------------
MFPTSTIVQYMDDILLASPTNEELQQLSQLTLQALTTHGLPISQEKTQQTPGQIRFLGQVISPNHITYES---
TPTIPIKSQWTLTELQVILGEIQWVSKGTPILRKHLQSLYSALHGYRDPRACIT-
>BLV
QFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKPNGTWRFVHDLR------
ATNALTKPIPALSPGPPDLTAIPTHLP-
HIICLDLKDAFFQIPVEDRFRSYFAFTLPTPGGLQPHRRFAWRVLPQGFINSPALFERALQEPLRQVSA------------
AFSQSLLVSYMDDILIASPTEEQRSQCYQALAARLRDLGFQVASEKTRQTPSPVPFLGQMVHNQIVTYQS---
LPTLQISSPISLHQLQAVLGDLQWVSRGTPTTRRPLQLLYSSLKGIDDPRAIIQ-
>PreDIRS
------EHAAFVTDEDEKLCSTGAAELSDVKVISALSVSVNADAKCRLVMDLT------
TVNPYITANKIKLENVAIAKSLIPKSG-FMLTFDMKSGYHQARMADSELIYLAFRW-------
```

```
EGKTFWMKALPFGLSSAPEYFTKLFRHPLATLR---------------
GDGVNCLLYLDDLLVWSETYEGACEASAKVRALFGKLGVVLNNEKSSVTQREVKWLGVVFNLTTLKISKNRIARLLNR
KRPSAKDRLK-FTGALNSMHDVLGPMAAITKSLFCFIASVT--------
>CbrDIRS
-RPSVKRHEEFLIAEIKRLEEEGVVVR-EEPCVSPLHVV-EQGKKKRMILDLS------
ELNESLVPPRFKLENLKTAWPFLENAK-FAATFDFKSGYHHIKIHRDSRDLLGFSLSNP----
PAAYFLFKGLPFGLSIAPWLFTKIFKVLVQKWR---------------
AEGIKIFLYLDDGLVVGETEREVARAVRKVREDLHNAGVCVAEEKSFWVDAKFTWLGYECDLVEVRGTEKRMDELRR
SLAPTVLDRMK-FLGCQASFELVTGDIGVAARWMMQVVGENQKKKS----
>Bel
---------------------RGHMEKSDTCFYLPHHAVIKLDTKCRVVFDGSGKDSGVSLNDRLHIGPPIQRDLFGVCLRFRQHQ-
YVLCADVEKMFRGIKVFKPHTNFQRIVWRTTENE-PLLHFRLLTVTYGLAPSPFLAVRVLKQLADDHGEY----------
PAAAHALLHDAYVDDIPTGANTFEELMILKDELIALLDKGKFKLRKWSSNSWDSPVKVLGIQWNPGDVLYLNLKGCDA
TISTKRELLSQLSRIYDPLGLVAPVTVLL-----------------------
>Kobel
---------------------NGYAEKKGKEWYIPHHGVYHPRNKIRVVFDCAAKCDGTSLNDVLLPGPNLMNDLKGVLMRFREFP-
VAFASDIECMFYQVKVPESQRDLLRFLWLPKDLDQKPKVYRMAVHLFGAVSAPSCANYSLRRTGIDNIAR----------
EETVNTLHNNFYMDDGLKSVENVEQAIQLVKELKALCKEGGFNLTKWSSNERLPTERTLGVLWNAEDEFQFKCQIKKT
-EVTRRSMLSEVSSIYDPLGFIAPLLIPA-----------------------
>Tamy
---------------------SGYAEETGRTFYLPHFAVVHPMKKIRIVFDAASRYEGKSLNDALLPGPDLLQSLFGVLLRFRQGP-
VAVVADIKEMFLQIKIREQDRDSLRFLWRGEDRASKPREYRMTSVIFGAASSPATAIFVKNRNAEEHQSH----------
PEAVKAIVRNHYMDDYLQSFATIEEAIDTAATVDSIHKEAGFELRQWASNEQSNEEKTLGLRWLIKDALAFNGGLRNAP
KVTKREVTSAVMSTFDPLGLIAPLL--------------------------
>Purbel
---------------------MGFAKKYGPVHYISHHAVIRPDTPLRIVFNSSAIYKGHCLNEYWKKGPDLLNNLFGVLLRFREHP-
VAICADISKMYHRILIPERDQHVHRFLWRDMNQEKEPDVYMMKVVTFGDKPAAAMAQIGVRLTAEEGEKY----------
PEAAAILKRNIYMDDICDSVETEDTAKKRIAEVDELLEAGGFKVKGWQSNKPIAEEKVLGVVWERHDTFSYKVKLQET
DKKTKRKILGKVARIYDPIGFAAPIIIKA-----------------------
>Zebel
---------------------AGYVAKKNESWYIPHH-VVHHNNKDRIVFNCSFQHQGQSLNSQLLPGPTLGPSLLGVLLRFRQHA-
VAVSGDIKGMFHQVRLLPKDKQVLRFLWRDLCRDREPDIYEWQVLPFGTTCSPCCAIFALQHHAQGHKDM----------
PKLVNIVENSFYVDNCLYSTPTAKEAKDVVDGLRQLLAEGGFDLRQWACNVPDLQEPTLGLRWNCLDSLGYNLRSAEP
LEPTMRNMYKTLASQYDPLGFIIPFTT-------------------------
>Cer10-1
---------------------AGIVERADVKYYTPHRAVIKETTKLRIVLDASSPPPEPSLNDCLYAGDNLVTPLYGILLRGRIYP-
YVVVADIEKAFHQVRLPTEFRNASLFLWL-KDITKPPTTYRFTRIPFGVASSPWLLAASILHFLDRN-------------
PPLNQRVRENLYVDNCLIGTFDKSEITEIIKNSKDIFMKMKMNLREYVTNSEHRVIKLLGCEWDSHDILSVALAVLDIDH
PTKRQVASKMAETFDPLGLITPLIVSF-----------------------
>Ty1B
-----------YHKEVNQLLKMKTWDTDPKRVINSMFIFRKRDHKARFVARGDIQHP--DTSNTVHHY-----
ALMTSLSLALDNNYHITQLDISSAYLYADIKEELYIRPPPHLGM-----
NDKLIRLKKSLYGLKQSGANWYETIKSYLKQCGEEVRGWSCVF----
ENSQVTICLFVDDMVLFSKNLNSNKRIIDKLKMQYDTKIINLGESDEIQ----YDILGLEIKYQRGKYMKLGMEN--------------
---------------------------------------
>CoDi5.6
-----------FGKEIGSLAQKGTDTIDRKNIPKDRVIT-----YARIVVDYR------
PQNRTAGGNTADLTTAKLMNSVISTDNARYACYDVKNFYLGTPLDRYEYMRIPLHLIPQHI--
TNKYVEIRKGIYGLPQAGMLANKLLKERLAPHGLEVRDTPGLFI--HKTRPIMFTLVVDDFGVKYVGKEH----
ADHLVSVLK-RYYTLSEDWTGS-----LYCGVSLKWNY--------------------KERWVDISMPGYIKNALQRYN----------------------
>Zeco1
-----------KQAEIFNWYKNNVFE-EDQKCVSTRWVCLKETPKARLVARGFEEL-NIHELQKDSPT-
CASDSLRLLLAVICQNKWQVHSMDIKSAFLQGQLSREIYVRPPPEVGK-----
ENVLWKLNKCVYGLADASLYWYIKVKIMLSTGSKLSKVDPAVFYLDQCKVTGVLACHVDDFLWAGSQNFS-----
TNVIPILK-SALNVGREEHEH----FCYVGMDFVT----------------------INGVVHVHQHRYIEN----------------------------
>Ty4
-----------YHKELQNLKDMKVFDVDN--IVPTNTIFKKRNYKARIVCRGDTQSP-----DTYSVTSLNHNHIKIFLMMQTTEI-
CLWTLDINHAFLYAKLEEEIYIPHP----------GDVYVKLNKALYGLKQSPKEWNDHLRQYLNGIGKDNSYTPGLY--
QTEDKNLMIAVYVDDCVIAASNEQR----LDEFINKLK-SNFELKITGTLIDVLDTDILGMDLVYNK--------------------
RLGTIDLTLKSFINRMDKKYK----------------------
>1731
```

```
-----------MGLEYKALLANETWKLRNRRCVACKWVYLKRDFKARLVAKGCSQKFGVDYFETFSPV-
CRLESVRLILALAAEMQLYLHHMDVCTAYLNSELKDTVYMKQPQGFTDAAN--
PDQVLLLRKAIYGLKQSGREWNSKLDGVLKDLGKACNHEPCLYQQSGQGNLMLILVYVDDLILACQSRED----
MEDLKAKIS-ESFECTDKGPLH-----LFLGMEVQRDG--------------------DLGEITLGHSQYIKELLRDYG----------------------
>pCretro6
-----------MQREYDQLTKLGCWDLAGRKAIGCKWVYIKRNYKARLVAQGFSQVPGVDYDETYAPV-
MRPESLHILAAIAVILNLEWDIENAVGAYLNSQLKLTIYMRQPEGFDDG----
SGRVCKLNLALYGLKQSGREWNLLLDEFLRGIGRASSVDPCVYLRIDEGSPTFLAVHVDDFSLFAKTREI----
MDKLKGELS-SRFEMTDLGPVR-----QILGYEVIRER--------------------DQRTLMLRQAAYIRKVLDRFN----------------------
>Tnt-1
-----------MQEEMESLQKNGTYKLKGKRPLKCKWVFLKKDYKARLVVKGFEQKKGIDFDEIFSPV-
VKMTSIRTILSLAASLDLEVEQLDVKTAFLHGDLEEEIYMEQPEGFEVAGK--
KHMVCKLNKSLYGLKQAPRQWYMKFDSFMKSQTLKTYSDPCVYFKRSENNFIILLLYVDDMLIVGKDKGL----
IAKLKGDLS-KSFDMKDLGPAQ-----QILGMKIVRER--------------------TSRKLWLSQEKYIERVLERFN----------------------
>Hydra1-2
-----------MEDEISSLQENNTFTLEGKHAVGGRWVYIKSNYKARYVAKGYSQVLGVDYIETFSPT-
ANITSIRALMQMAAQYDLELHQMDVKTAYLHAPIDCKIYMEQPKGFEVKSN--
DKLVCKLNKSLYGLKQLGRNWNKILHVYLTENDIQNLADYCVYSKRSGKDRVIIIIWVDDLIIAASDNSL----
LKDVKEMLT-SKFKMKDLGKLK-----HFLGIDFDQ----------------------SKGTVRVNQKRYILRILERFN----------------------
>Copia
------------NTELNAHKINNTWTIENKNIVDSRWVFVKYNYKARLVARGFTQKYQIDYEETFAPV-
ARISSFRFILSLAIQYNLKVHQMDVKTAFLNGSLKEEIYMRLPQGISCN----
SDNVCKLNKAIYGLKQAARCWFEVFEQALKECEVNSSVDRCIYIDKNINENIYVLLYVDDVVIATGDMTR----
MNNFKRYLM-EKFRMTDLNEIK-----HFIGIRIEM----------------------HEDKIYLSQSAYVKKILSKFN----------------------
>Koala
-----------MDEEYQALVTNKTWHLPNKNIIDCKWVYVKRKYKARLVAKGFRQRYGIDYEDTFSPV-
IKMTTIRIILAIAVSKGWFLRQLDVKNAFLHGILEEEVYMYQPPGYEDKQH--
PNYVCKLDKALYGLKQAPRAWFARLSHKLNQLGQESKADTSLFFYNREGLTVFLLIYVDDIIVVSSKSEA----
IPILLQNLQ-QDFALKDLGNLH-----YFLGIEVNQ----------------------SPNGIVLTQAKYANDLLRRSG----------------------
>SIRE1-4
-----------MQEELEQFKRNEVWELEGTNVIGTKWIFNKTNNKARLVAQGYTQIEGVDFDETFAPV-
ARLESIRLLLGVACILKFKLYQMDVKSAFLNGYLNEEVYVEQPKGFADPTH--
PDHVYRLKKALYGLKQAPRAWYERLTEFLTQQGRKGGIDKTLFVKQDAENLMIAQIYVDDIVFGGMSNEM----
LRHFVQQMQ-SEFEMSLVGELT-----YFLGLQVKQ----------------------MEDSIFLSQSRYAKNIVKKFG----------------------
>Vitico1-1
------------KEEIAAIEKNETWELEDKNVIGVKWVFTKYLHKAQLVAKGYAQQHGVDYDDTFSPI-
ALFETVRTLLALAAHMHWCVYQFDVKSAFLNGELVEEVYVSQPEGFIVPGK--
EEHVYRLKKTLFGLKQAPRAWYSKIDSYFVENGERSKSDPNLYLKRQ-----------DDMIYMGSSSFL----INEFKACMK-
KKFEMSDLGLLH-----FFLGLEA-------------------------EDDTERADARRFRSLVG-------------------------
>CoDi6.3
-----------IRKEFRDMVKRKVWRRSNRRCVKSKWVFIKRNFRARLVACGYSQIPGVDFSESYSPV-
ANDITIRLLLVAMILFGLSAKIVDVETAFLYGELEEEVYMENPEGLEDSN---
DDEALLLLTTIYGLVQAARQYYKKARGILRKIGTGGDVDPCLFVKKSSLGIVFIALYVDDNLLVGHPKA-----
IECAIEQMKRHGLILKVEDDLK-----DYLSCEIQFSK--------------------DKTKAWLGQPHLISNLMSKFG----------------------
>CoDi6.7
-----------MQEEVQAHTENKLWELQGTPIIPAVWSMRKRRWKARLAFDGSKQIHGVNFWETYAPV-
ASWPTIRYILTLALINRWHMQQIDFVLAYTQAEAECEMFMKIPKGFTVEHD--
EEYVLRIKKNYYGLKQAGRVWNQHLVSKLQECGKQSEHDQCLFY----RGRSVYVLYTDDSILAGPDLEE----
LEQIKIDMANSGLKLTSEPGVS-----DFLGVKIDR----------------------KGDEIHLTQPHLINSILDDLR----------------------
```

Dataset S3. RT alignment generated by Mafft 7.402 and used to reconstruct the phylogenetic tree in Fig. 2.

```
>ZFERV-2
QYPINKSALPEIKITIEELKRKGIIKEV-DNAPTNSPIQAVAKP---DGGWRLVTNYK-ALNKQTVPDTRYLINARDATGAA--
PEGRLLSKIDLANGFWSVPLKPESRARTAFTF-------ENKQYVYNVLPQGFKNSPNAFQAIMMEILKGLPI-------
VVYIDDILIVTNDEEEHLKILDETLSRLTKEGFRLNYKKMEIAKDDVEFLGFQITGNERFLSQTTENKLKELKEITTLKQL
QSMLGTLNFVRDLIPGYARHAKPLYNATKGGV-----LTW
>AciFLERV
QYPLNPGAVKEMDLIVRELLTLGVIRQE-LNPITNSPIQAVKKPESSGGGWRPVINFK-
ALNRRTVANRASLINPQGTLKTL--KLKPFKSCIDLANGFFSLRLTRGSQGKTAFTH-------
KGRSYVWERLPQGYKNSPNVFQAAVMDVLDGLGA-------
TIYIDDVFIADDTEEQHLDRLQKIVERLTAAGLKLNLKKCQFGRKRVDYLGFQVSDD-
LGLSEVYRKKIEQIEPPSSLNELQKVLGLCNYVRDYVPGYQRYAKPLYAKLKKPDETAA-WNW
>AliFLERV
QYPLHPEAVAEMDKIVKELHALGIIREE-PNPLTNSPIQAVKKPESAGGGWRPVINFK-
ALNRRTVANRASLINPQGALKTL--RVKKYKSCIDLANGFFSLRLARESQGKTAFTH-------
KGKSYVWQRLPQGYRNSPNVFQSAVLEILEDVGA-------
SVYIDDVFIADDTEEEHLKRLEEVVKRISAAGLKLNLAKCQFGQFKVNYLGFQVTTD-
LGLSDGYREKIQQIQPPRCENDLQKILGLCNYVRDHVPNYQKYAKPLYACLKKEATTKA-WSW
>FFV
QYHINPKAKPDIQIVINDLLKQGVLIQ--KESTMNTPVYPVPKP---NGRWRMVLDYR-AVNKVTPLIAVQNQHSYGILGSL-
-FKGRYKTTIDLSNGFWAHPIVPEDYWITAFTW-------QGKQYCWTVLPQGFLNSPGLFTGDVVDLLQGIPN--V----
EVYVDDVYISHDSEKEHLEYLDILFNRLKEAGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENITAPTTLKQLQ
SILGLLNFARNFIPDFTELIAPLYALIPKSTKN-Y-VPW
>BFV
QYHINPRAKADIQIVIDDLLRQGVLRQ--QNSEMNTPVYPVPKA---DGRWRMVLDYR-
EVNKVTPLVATQNCHSASILNTL--YRGPYKSTLDLANGFWAHPIKPEDYWITAFTW-------
GGKTYCWTVLPQGFLNSPALFTADVVDILKDIPN--V----
QVYVDDVYVSSATEQEHLDILETIFNRLSTAGYIVSLKKSKLAKETVEFLGFSISQNGRGLTDSYKQKLMDLQPPTTLRQ
LQSILGLINFARNFLPNFAELVAPLYQLIPKAKGQ-C-IPW
>EFV
QYRINPKAKADIQIVIDDLLKQGVLKQ--QTSPMNTPVYPVPKP---DGRWRMVLDYR-
AVNKVTPAIATQNCHSASLLNTL--YRGQYKTTLDLANGFWAHPIQESDQWITSFTW-------
NGKSYVWTTLPQGFLNSPALFTADVVDLLKDIPN--V----
EVYVDDVYFSNDTEEEHLKTMDLLFQKLQTAGYIVSLKKSKLGQHTVDFLGFQITQTGRGLTDSYKSKLLDITPPNTLK
QLQSILGLLNFARNFIPNYSELITPLYQLIPLAKGI-Y-IPW
>SFV
QYHINPKAKPSIQIVINDLLKQGVLRQ--STSPMNTPVYPVPKP---DGKWRMVLDYR-AVNKTIPLIAAQNQHSLGILTNL--
IRHKYKSTIDLSNGFWAHPITEDSQWITAFTW-------EGKQHVWTRLPQGFLNSPALFTADVVDILKEVPG--V----
SVYVDDIYISSPTMEEHFQVLDSIFRKLLETGYIVSLKKSALARYEVNFLGFVISETGRGLTSEFRERLQEITPPTTLKQLQS
ILGFLNFARNFVPNFSELVQPLYQLISTASGN-F-IQW
>CoEFV
QYPINTKAIPSIQVVINELLEQGVLVK--QTSPTNMAVLPVPKP---DGTWRLVLDYR-ALNKHSEPVRAQNQHSSGILANI--
ERKAYKSSVDLANGFWSHPIREEDWPKMAFTW-------CGFQYLWTRLPQGFLNSPALFSADVISLVGQLPG--V----
YCYVDDIYLTHDTEEEHLKILDQVLEILIKAGYVINIKKSKLCRKVVDFLGFSLSDEGRGLSDQYREKLAAIKPPQTLRQL
QSVMGLLNFSRLFVKGFSELAKPLYDL------------
>Loki-Xla
QYRLPLAAYESLSEIVKNLEKKGIIRP--VHSSFNHPILGVLKP---NGQFRLCSDLR-QLNKRVYMSGWPVPYIDQCLAQI--
QGSKIFTALDCAQGYWTIKIDERDQHKLAFTF-------GKQQYAWTRLPFGYINAGHEFAVFMHKAMPDAAERGT----
LSYVDDILIKSTTFEEHIAELKYVL--------
KLSLQKAQWCRSKVNFLGHEVTAEGINPQKKKVEAIKNTKSPTNLKELRSFLGMMNYSRKFIDNYAEITKPLLQLLKKG
A---K-WEW
>Loki-Xtr
QYRLPLAAYESLSEIVKNLEKKGIIRP--VHSSFNHPILGVLKP---NGQFRLCSDLR-QLNKQVYMSGWPVPYIDQCLAQI--
QGSKIFTALDCAQGYWTIKIDERDQYKLAFTF-------GKQQYAWTRLPFGYINAGHEFAVFMHKAMPDAAERGT----
LSYVDDILIKSTTFGEHIAELRYVLNQLKKAGVKLSLQKAQWCRSKVNFLGHEITAEGINPQ----------------EDRSW--
VMNYSRKFIDNYAEITKPLLQLLKKGA---K-WKW
>Loki-Npa
QYRLPLACYDSLAEIIRNLKERGVIRS--VHSSYNNPILGILKP---NGQWRLCADLR-
QLNKRVYMSGWPVPYIDQCLVQM--QGARIFTTLDCAQGYWTIKVSPEDQYKLAFTF-------
```

```
GKEQYTWTRMPFGYINSGHEFAVFLHKAMPDATERGN----
LTYVDDILLKSTSFEKHVQEIRHVLNQLREAGVKISLQKAQWCRTKVNFLGHEVTSEGLNPQKKKVEAVMKSKTPTNV
SELRSFLGMMNYSRKFIDNYAEISKPLLTLLKKDV---P-WMW
>Loki-Rca
QYRLPLASYESLAEIIRKLHERGIIRP--IHSSYNHPILGILKP---NGQWRLCADLR-QLNKRVYMSGWPVPYIDQCLAQL--
Q-SQVFSALDCAQGYWTIKVSEEDQYKLAFSF-------GNQQYCWVRMPFGYINSGHEFAVFMHKAMPDALERGT----
LAYVDDILIKSTDMTKHLTELRHILGQLR-
AGVKLSIQKAQWCRTRVNFLGHEVTSEGLNPQKKKVEAVFNCKVPTNLKELRSFLGMVNYSRKFIDNYAEISKPLLHLL
KKDT---E-WSW
>Loki-Ape
QYKIPLAAYESIQEILDKLLEKRIVRE--CNSTYNSPIWGALKP---TGKWRLTIDYR-PLNKQVPLSRWPMIHLDQELAKV--
KEARFFSTVDVANGFWTMTVDPVDQYKLAFSF-------CNRQYTWNRCPFGYSNSPAEFNIFLHKAMSDASARGN----
LIYVDDILMRSRSFDDHLAEIRHVLGQLAAAGAKLAIAKGQWCRTKVEYVGLTVGANGIEPKAERIRAIQDIMAPVNVS
KLQSFLGVCNYSRQFIADYAEIARPLTELLGKDK---P-FLW
>Loki-Pfl
QYKIPLAAYESIQEILDKLLEKQIVRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-PLNKQVPLSRWPMIHLDQELAKV--
KEARFFSTVDVANGFWTMRVDPVDQYKLAFSF-------CNRQYTWNRCPFGYSNSPAEFNIFLHKAMSDASARGN----
LIYVDDILMRSRSFDEHLAEIRHILGQLAAAGAKLAITKGQWCRTKVEYVGLTVGANGIEPKAERMRAIQDIMAPVNVS
KLRSFLGVCNYSRQFIADYAEIARPLTELLGKDK---P-FLW
>Loki-Pyo
QYKIPLAAYESIQDILDTLLEKRIIRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-PLNKQVPLSRWPMIHLDQELAKV--
KGATFFSTVDVANGFWTMKVDQADQYKLAFSF-------GGSQYTWNRCPFGYSNSPSEFNIFLHKAMSDAATRGN----
LIYVDDILMRSQTFEDHITEIRHVLTQLSNAGAKLALTKGQWCRTKVKYVGLTVGADGIEPQAGRIRAIQDIKAPTRLTE
LRSFLGVCNYSRQFIEDYAEIARPLNELLHKDK---P-FEW
>Loki-Sau
QYRIPLAAYESIQDILDKLLQKQIIRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-PLNKQVPLSRWPMIHLDQELAKV--
KGACFFSTVDIANGFWTMKVDPADQYKLAFSF-------CNRQYTWNRCPFGYSNSPAEFNIFLHKAMSDAAARGN----
LIYVDDILMRSCTFETHLAEIRHVLNQLTRAGAKLAIAKGQWCRTKVEYVGLTVGGSGIEPQTGRIRAIQDIKAPTNVSE
LRSFLGVCNYSRQFIEDYAEIARPLTELLCKDK---L-FQW
>Loki-Ofa
QYRIPLAAYESIQEILDNLLQKQIIRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-PLNKQVPLSRWPMIHLDQELAKV--
KGACFFSTVDVANGFWTMRVNPADQYKLAFSF-------GNRQFTWNRCPFGYSNSPAEFNIFLHKAMSDAAARGN----
LIYVDDILMRSRTLE-
HLAEIRHVLNQLSAAGAKLAIAKGQWCRTKVEYVGLTVGADGIEPQAGRIRAIQDIKAPTGVSELRSFLGVCNYSRQFIE
DYAEIARPLTELLRKDK---S-FEW
>Loki-Bsp
QYRIPLAAYESIQEILDTLLKKEIIRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-PLNKNVPLSRWPMIHLDQELAKI--
KDAHFFSTVDVANGFWTMRVDPADQYKLAFSF-------GNRQYTWNRCPFGYSNSPAEFNIFLHKAMSDAAARGN----
LIYVDDILMRSQTFEGHLAEIRHVLSQLLAAGAKLALTKGQWCRTKVKYVGLTVGQHGIEPQADRIQAIRTIKAPNTLSG
LRSFLGVCNYSRQFVEDYAEIARPLTELLRNDT---P-YIW
>Loki-Cne
QYKIPLAAYGSIQEILNKLLEKNIIRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-ALNKQVPLSRWPMIHLDQELAKV--
KDARFFSTVDVANGFWTMKVDPADRYKLAFSF-------GNRQY--NRCPFGYSNSPAEFNIFLHKAMSDATERGN----
LIYVDDILMRSRTFKEHLEEI-
HVLGQLAAAGAKLALSKGQWCRKKVDYVGLTVGPNGIEPQADRIRAIQDIAAPTNLSELRSFLGVCNYSRQFIEDYAD
MARPLTELLRKNT---P-FDL
>Loki-Lch
QYRIPLSSYESVQEILHTLMEREIIRE--CNSTYNSPIWPVLKP---NGKWRLTIDYR-QLNKQVPLSHWPMTYLDQNLTKI--
AGAKYFTTLDIANGFWTVKVDPKDQYKLAFSF-------GKRQYTWNRCPFGYLNSPADFNIFLHKAMPDAEERGN----
LIYVDDILIKSNSASDHLDEIRYVLTQLKNAGAKLSLTKGQWCRKRVNYLGYDVGPDGLRPQQNRIKAIQAIRSPTNITE
LRSFLGICNYSRQFIENYAEISQPLTKLLRKNE---P-WVW
>Loki-Str
QYKIPLAAYAAVQEIIDSLLAKRIIRE--CNSTYSAPVWPVLKP---TGKWRLTIDYR-QLNKLVPLSRWPMTQLDQELPKV--
ANAKYFSTVDVANGFWTMTVDPRDQHKLAFSF-------SNKLFTFNRCPFGYANSPSEFNIFLHKAMPDAASRGT----
IIYVDDVLMRSETWSHHLNEMDHVLTQLGTAGAKLAIMKGQWCRTKVNYVGLLVGAEGILPQSNRIQAVRNIKTPTNL
HEVRSFLGVCNYSRQFIENYADLSKPLTHLLQKDT---P-FVW
>Loki-Tfu
QYKIPLESYEPVQEIINSMLEKGVIRL--CNSTCSAPIWPVLTP---NGKWRPTIDYR-KLNQQVSLSRWPMTQLEQEIPRV--
KGATIFSTLDVASGFWTIPVHPDDQHKLAFMF-------CNR-YTFTRCPFGYANSPAEFNIFLNKACPDARVRGN----
LIYVDDVLMKSTTMTDHLKEIDHVLQQLTKAGAKIALHKGQWCKTKVNYVGPLIAPQGIERQTNRIQAIQNIKPPANISE
LRSFLGVCSYSRQFIKDYSDIARPLMALLKKDS---P-FTW
>Loki-Lro
```

```
QYKIPLESYEPVQDIITSMLEKGVIRP--CNSTYSAPIWPVLKP---NGKWRPTIDYR-KLNQQVPLSRWPMTQLEQEIPRI--
RGATIFSTLDVASGFWTIPVHPDDQHKLAFRF-------DNRQYTFTRCPFGYANSPAEFNIFLNKACPDARVRGN----
LVFVDDILLKHTTVADHLKEIDHILNQLTTAGAKIALHKGQWCKIKVNYVGLLIGPQGIEPQTNRIQAIQNIKPPANVSEL
RSFLGVCNYSRQFIENYSDIARPLTALLKKDC---P-FIW
>Loki-Dcl
QYKIPLESYEPVQEIIDSMLEKGVIRP--CNSTYSAPIWPVLKP---NGKWRPTIDYR-KLNQQVPLSRWPMTQLEQEIPRI--
RGATIFSTLDVASGFWTIPVHPDDQHKLAFTF-------SNRQYTFTRCPFGYANSPAEFNIFLNKACPDARVRGN----
LIYVDDVLMKSMTVADHLKEIDHVLHQLTAAGAKIALHKGQWGKSKVNYVGLLIGPQGIEPQSSRIQAIQTIKPPVNVS
ELRSFLGVCNYSRQFIENYSDIARPLTTLLKKDC---P-FAW
>Loki-Sri
QYKIPLASYEPVQEIIDNLLEKGIIRP--CNSTYSAPLWPVLKP---NGKWRLTIDYR-KLNQQVPLSRWPMTQLEQELPKI--
RDAKYFSTLDIASGFWTIPVHVADQHKLAFTF-------ANRQYTFTRCPFGYANSPAEFNIFLNKACPDANDRGT----
LIYVDDILMRSSTLDTHLEEIDHVLEQLTAAGAKISLSKCQWCKTKVNYVGLLVGPRGVEPQINRVQGVTNIKAPTNVSE
LRSFLGVCNYSRQFIENYADLAKPLTELLRKDA---L-FVW
>Loki-Sme
QYKIPLASYEPVQEIIDDLLHRGIIRP--CNSTYSAPVWPVLKP---NGKWRLTIDYR-KLNQQVPLSRWPMTQLNQELPKV--
KDAKYFSSMDIASGFWTIPVHIADQHKLAFTF-------AGRQYTFTRCPFGYANSPAEFNIFLNKSCPDASERGT----
LIYVDDILIRTPTMKGHIDEIEHVLDQLTTAGAKISLSKCQWCKTKVNYVGLLVGSNGVEPQTNRVQGVMNIKTPTNISE
LRSFLGVCNYSRQFIPDFADIAKPLTDLLKKDS---L-FTF
>Loki-Oja
QYKIPLASYEPVQEIIDDLTNKGIIRP--CNSTYSAPIWPVRKP---NGKWRLTIDYR-RLNQQVPLSRWPMTQLEQELPRV--
KEAKYFSTLDVASGFWTIPVHPDDQHKLAFTF-------ARQQYTFTRCPFGYANSPAEFNIFLNKACPDATQRGT----
LIYVDDILIRNATLEDHIEEIEHVLKQLTSAGAKISLSKCQWCRTKVNYVGLLVGPTGVQPQVGRVQGLKDIAAPTNLKE
LRSFLGVCNYSRQFIEDYADIAKPLTELLKKDV---P-FRW
>Loki-Ame
QYKIPLASYGPVQEIIDDLLDKGIIRP--CNSTYSAPLWPVLKP---NGKWRLTIDYR-RLNDQVPLSRWPMTQLEQELPRV--
RDAKYFSTLDVASGFWTIPVHVEDQHKLAFTF-------AGRQFTFTRCPFGYSNSPAEFNIFLNKACPDARERGT----
LIYVDDVLIRNNSLDAHLEEIDHVLDQLTKAGAKISLAKCQWCKTKVNYVGLLVGPDGVLPQPCRAQGIVDIAEPKTIH
ALRSFLGVCNYSRQFIENFAELAKPLYQLLKQDV---P-FIW
>Loki-Bpe
QYKIPLASYEPIQEIIDDLLTKGVIRP--CNSTYSAPIWPVLKP---NGKWRLTVDYR-QLNTRVPLSRWPMTQLGQDLPKV--
RDAKYFSTLDIASGFWTIPVHPDDQHKLAFSF-------ANRQYTFTRCPFGLANSPAEFNIFLNKACPNAASRGT----
LIYVDDILIRTTTLDDHLKEIDHVLDQLTSAGAKISLSKCQWCKTKVNYVGLL--------
RSSRIQGITNLKVPANVSELRSFLGVCNYSRQFVEHYAELAKPLTDLLKKDE---T-FSW
>Loki-Lca
QYKIPLAAYESIQETLENLLKKNIIRE--CNSTYNSPIWPVLKP---TGKWRLTIDYR-QLNKTVPLSRWPMAEIDHGLNQI--
KGAKVLTTMDLANGFWTMPVKETDQYKLAFTF-------DRIQYTWNRCPFGYSNSPADFNIFLHKAMGDAKERGT----
IVYVDDILVKDSSWQEHLKSLQHTLEQLKEAGAKISIQKGQWARKRVDYLGFQIGTEGLLPQTNRIEALLALKSPNTVPT
LRSFLGICNYLRQFVDDYAGITRPLVKLLQKDE---P-WEW
>Loki-Pma
QYKIPLAAYESIQETLENLLKKNIIRK--CNSTYNSPIWPVLKP---TGKWRLTIDYR-QLNKTVPLSRWPMAEIDQNLNQI--
KGAKILTTMDLSN-FWTMPVKEADQYKL--TF-------NGIQYTWNRCPFGYSNSPAD-NIFLHKAMGDAKERGT----
IVYVDDILVKDSSWEEHLESLQHTLEQLE--AAKISIQKGQWARKRVDYLGFQIGT-
GMLPQTKRLEALLALKPPTTVTTLRSFLGICNYLRQFVDDYAGIARPLVKLLQKDE---P-WEW
>Loki-Aro
KYRIPIAAHDPVQDTITEILHAGIIRK--CNSTYNSPIWPVLKP---SGKQSLTVDYR-KLNKQVPLSRCLWPTWMKPLQK---
SRMPHTSQADISNGFCTIPVREQDQHKLAFSF-------NNVQ-TWTVTPFGYANSPAEFNIFFHKAIPDVNSKNI----
AIYVDDILMWSDSWGEHQQLLSYVLTRLSQAGAKISLDKGQWCKRAVNYLGFQVGGDGIRP---
RTEAPKAIPAPTTLKQLCSFLGICNFSQHFVADYAKISQPLTRLLKKDV---P-WTW
>SnRV
QYPVPDASHASIKETVEALLEQGVLRK--CNSTVNSAIWPVGKP---DGSWRLTIDYR-
PLNSAVSCPYPTVASTPELFAKLE-KKYQVYSSLDISNGFWSIRLEEECQYLFAFTF-------
DTQQYTWTRLPQGFHASPGIFHQALYNGLASCKTESQGCKLLQYVDDILLMSEDRDHHLRSLAILLQGLKDLGVKINPK
KSHFCKDQVQYLGVNVGADTRSLIDARSQLIRTLDIPLTVQGLRSALGLFNFCRAWIPEFSRKTQSLYDMLKGDC---K-
LKW
>KoRV
QYPMSKEAREGIRPHIQRFLDLGILVP--CQSPWNTPLLPVKKP--GTNDYRPVQDLR-
EVNKRVQDIHPTVPNPYNLLSSLP-PSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPE-
KGNTGQLTWTRLPQGFKNSPTLFDEALHRDLASFRALNPQVVMLQYVDDLLVAAPTYRDCKEGTRRLLQELSKLGYR
VSAKKAQLCREEVTYLGYLLKGGKRWLTPARKATVMKIPTPTTPRQVREFLGTAGFCRLWIPGFASLAAPLYPLTREKV
---P-FTW
>MuLV
```

```
QYPMSQEARLGIKPHIQRLLDQGILVP--CQSPWNTPLLPVKKP--GTNDYRPVQDLR-
EVNKRVEDIHPTVPNPYNLLSGLP-PSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWRDPE-
MGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGYRAS
AKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGT--
-L-FEW
>FLV
QYPMPHEAYQGIKPHIRRMLDQGILKP--CQSPWNTPLLPVKKP--GTEDYRPVQDLR-
EVNKRVEDIHPTVPNPYNLLSTLP-PSHPWYTVLDLKDAFFCLRLHSESQLLFAFEWRDPE-
IGLSGQLTWTRLPQGFKNSPTLFDEALHSDLADFRVRYPALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGYRAS
AKKAQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPKNSRQVREFLGTAGYCRLWIPGFAELAAPLYPLTRPGT---
L-FQW
>REV
QYPITLEAKRSLRETIHKFRAAGILRP--VHSPWNTPLLPVRKS--GTSEYRMVQDLR-EVNKRVETIHPTVPNPYTLLSLLP-
PDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAE-
EGESGQLTWTRLPQGFKNSPTLFDVALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRV
SGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGTIGYCRLWIPGFAELAQPLYAATRGGN---
P-LVW
>WDSV
QYPLPKDKTEGLRPLISSLENQGILIK--CHSPCNTPIFPIKKA--GRDEYRMIHDLR-AINNIVAPLTAVVASPTTVLSNLA-
PSLHWFTVIDLSNAFFSVPIHKDSQYLFAFTF-------EGHQYTWTVLPQGFIHSPTLFSQALYQSLHKIKFKISS-
EICIYMDDVLIASKDRDTNLKDTAVMLQHLASEGHKVSKKKLQLCQQEVVYLGQLLTPEGRKILPDRKVTVSQFQQPTT
IRQIRAFLGLVGYCRHWIPEFSIHSKFLEKQLKKDT---P-FQL
>Xen-1
QYPLKPAQSESINKQLQTLLENGAIKR--QSSPCNTPLFPVKEK--GEDKYRLVQDLR-AVNEATVMETPLVSNPHTILSGIP-
PSATHFRAVDLTNAFYSIPLREDCQYLFAFTH-------
ERQQYVWTVLPQGAQNSPTHFSLALTSILDSWISSHPEITLLQYVDDLLLCAPDLPTCEASSTDLLSFLADQGCKASKEK
LQWCQTTVVFLGQCISQGTRHITEGRIKTLQDIPLPKGHKPLHAFLGLISYCRSWIPEASLLMQPLYDVLKSD----P-FTM
>Amn-san
-------------------LQRGFIRP--STSPAGAGFFFVEKK---DGGLRPCIDYR-GLNKITVKNRYPLPLISELFDQL--
KGAKIFSKLDLRGAYNLIRIREGDEWKTAFNT-------RDGHYEYLVMPFGLCNAPAVFQEFVNDIFRDLLGKSV----
VVYLDDILIFSQDLETHRSQVKEALSRLRENFLFAKLEKCTFEVPKISFLGYIISSRGFEMDPAKVSAIQKWPLPQSTKAIQ
RFIGFANYYRQFIKDFS---------------------
>Sushi-ichi
-------------------LASGIIRP--SSSPLAAGFFFVAKE---DGGLRPCIDFR-KLNNITVKNKYPLPLMSSTFEPL--
THARVFTKLDLRNAYHLVQIRKGDEWKTAFNT-------HLGHFEYLVMPFGLSNAPAVFQELVNDVLRDMINVFV----
VVYLDDILIFSRTMEEHHQHVRLVLQRLLENRLFIKAEKCIFHSASVGYLGYIVEEGRVRADPAKIQAVVEWPRPTDRTQ
LRRFLGFAGFIRRFIKGFA----------------------
>CRM
-------------------LDKGYVRE--SLSPCAVPVILVPKK---DGTWRMCVDCR-AINNITIRYRHPIPRLDDMLDEL--
SGAIVFSKVDLRSGYHQIRMKLGDEWKTAFKT-------KFGLYEWLVMPFGLTNAPSTFMRLMNEVLRAFIGKFV----
VVYFDDILIYSKSMDEHVDHMRAVFNALRDARLFGNLEKCTFCTDRVSFLGYVVTPQGIEVDQAKVEAIHGWPMPKTI
TQVRSFLGLAGFYRRFVKDFS---------------------
>Reina
-------------------LEQGVIQH--SSSPFASPVLLVKKK---DGEWRLCVDYR-RLNAHTVKNRYPMPVFDEIVDEL--
CGTKIFTKLDHRSGYHQIRIKEGDEFKTAFQT-------HNGHYEYRVMPFGLTGAPATFQDFMNKILTPFLRKCV----
VVFLDDVLIYSRDMEEHVLQVKQVFQKLKDHQLKLKLSKCRFAQTTLEFLGHIISAEGIATDPEKVQVIRDWPIPNNVKE
VRSFLGMAGYYRRFVAHYA----------------------
>Del
-------------------LNKGFIRG--STSPWGAHVLFDPKK---DDSKRMCIDYK--LNSVTVKNKYPLPRIDDLFDQL--NGA-
YFSKIDLRFRYHQLRIR-ADIPKTAFRT-------RYGHYEFLVMPFGLTNVPTAFMNLMNRVFREYLDKFI----
VVFVDYVLIYSRTQKDHEHHLRISLQLLRNNQLYAKLSKCEFWMEKVKFLGHVVSREGIVVDPVKVKAVMNWELPKN
IFEIRSFLGLAGYYRRFIKGFA----------------------
>Galadriel
-------------------LDAGLIQP--SKAPYGAPVLFQKKQ---DGTMRMCVDYR-ALNKATIKNKYSVPLVQDLMDRL--
SKACWFTKLDLRAGYWQVRIAEGDEPKTTCVT-------RYGSYEFLVMPFGLTNAPATFCNLMNNVLFDYLDDFV----
VVYLDDIVIYSRTLEEHVNHLSLVLSQLRKYTLYVKMEKCEFAQQEIKFLGHLVSKNQVRMDPKKVQAIVDWQAPRHV
KDLRSFLGLANYYRKFIAGYS----------------------
>Maggy
-------------------LKKGFIRP--SSSSVASPVLFVKKQ---GGGLRFCVDYR-ALNNITVKDRYPLPLVRETLNNL--
AGMKFFSKIDIVSAFNNIRIKKGEEYLTAFRT-------RFGLYESLVMPFGLTGAPATFQRYINDSLREYLDVFC----
TAYLDDILIYSRTRTEHEEHLKLVLEALRKAGLYANAAKCEFFVTETKFLGLLVGVEGVKMDPEKITAVLDWQTPKKLT
DVQAFLGFGNFYRRFIRDFA---------------------
```

```
>Pyggy
-------------------LDKGFIRV--SSLLAFLLVLFAKKL---EGGLRLCIDYW-ALNAITRKNRYLLPLIRETLNNI--
SKAKWFTKLDVIAAFYKIRVAEGDEWKTAFRT-------RFSLYEWLVTPFGMANSLSTFQRYINWTLREYLDEFC----
SAYLDDVLIYTDSLEQHQDHVRKVLRKLQESGLNVDIKKCEFGVKSTKYLGLIIDAEGIRMDPEKVKAIMEWE-----------
-----------------------------------
>Skippy
-------------------IRKGYIRP--SKSSAGFPVMFVPKP---NSNLRLVVDYR-QLNEITEKDRTSLPLITELKDRL--
FGKKWFTALDLKSAYNLIRIKEADEWKTAFRT-------KYGLFEYLVMPFGLTNAPAVFQRMITNVLREYLDIFV----
VCYLDDILIFSDTEEEHTEHVHKVLKALQDANMLVEPTKSHFHQSQVTYLGHEISHNEIRMDRRKIAAVAEWKVPTSVK
ETQSFLGFANYYRRFIKDFA---------------------
>Gypsy
-------------------LKDGIIRP--SRSPYNSPTWVVDKK---NPNKRLVIDFR-KLNEKTIPDRYPMPSIPMILANL--
GKAKFFTTLDLKSGYHQIYLAEHDREKTSFSV-------NGGKYEFCRLPFGLRNASSIFQRALDDVLREQIGKIC----
YVYVDDVIIFSENESDHVRHIDTVLKCLIDANMRVSQEKTRFFKESVEYLGFIVSKDGTKSDPEKVKAIQEYPEPDCVYK
VRSFLGLASYYRVFIKDFA---------------------
>Tom
-------------------LEQGLIRE--SNSPYNSPTWVVPKK---KAKYRVVIDYR-KLNEITIPDRFPIPNMDEILGKL--
GKCQYFTTIDLARGFHQIEMDSESIQKTAFST-------KRGHYEYVRMPFGLRNAPATFQRCMNNILRPLINKHC----
LVYLDDMIIFSTSLDEHLNSLQLVFEKLSESNLKLQLDKCEFLKKEATFLGHIVTPDGIKPNPLKVEAIASYPIPTKVKEIR
AFLGMTGYYRKFIPSYA---------------------
>Zam
-------------------LDQQIIRS--SHSPWSAPVWVVPKK---NRKWRLVIDYR-KLNDKTISDRYPIPNINDILDSI--
GKAKYFSTLDLTSGFHQIEMNPKDIAKTAFTV-------EGGHYEFTRMPFGLKNAPATFQRVMDSVLGDLNGTIC----
LFYLDDIIIFSPSLQKHLLDIKMVFEKLRAANFKLQPSKSEFLRKEIEFLGHIVTQDGVKPNPNKISAIKKFPCPTNRRAIKS
FLGLLGYYRKFIRDFA---------------------
>Cer3
-------------------IQSGRIVE--SNTPWTSPIVLIKKK---NGSLRVCLDFR-KLNEVTVPDNYPLPRIDSIIEQI--
GGSKYFTSLDMANGYLQLRLDAESSDKCGFIT-------ENKVYAYTHLPFGLRSAASYFQRDLKQVLG--LEKEV----
TVYIDDVLIFSKTFEEHLNSLRKVLERFRKFNLKVSPNKCEFFKQSITFLGHEINADNYKPNRINVDAIVNLPTPRNVGDV
RRFIGMSGFFRKFLPNFS---------------------
>Mdg3
-------------------LENGFIRP--SDSEYASPIVLVKKK---TGDLRMCVDFR-KLNKMTMKDNYPLPLIDDLLDRM--
NEKTVFTKLDLKNGFFHVHVKKESIKYTSFVT-------PLGQYEWLRMPFGLKNAPSVFQRFVNKIFADMIENKV----
VVYMDDILLATENINEHLETLKEIFKRLVENKLELRIDKCEFMQSSIKYLGFIINKDGIMPNDKGIEAIKNFPIPNNVHTVQ
SFLGLCSYFRRFIKDFS---------------------
>TF1
-------------------LKSGIIRE--SKAINACPVMFVPKK---EGTLRMVVDYK-PLNKYVKPNIYPLPLIEQLLAKI--
QGSTIFTKLDLKSAYHLIRVRKGDEHKLAFRC-------PRGVFEYLVMPYGISTAPAHFQYFINTILGEAKESHV----
VCYMDDILIHSKSESEHVKHVKDVLQKLKNANLIINQAKCEFHQSQVKFIGYHISEKGFTPCQENIDKVLQWKQPKNRK
ELRQFLGSVNYLRKFIPKTS---------------------
>Mdg1
-------------------IKDGIVEQ--SISEYNSPLLLVPKK---EKRWRLVVDYR-QINKKLLADKFPLPRIEDILDQL--
GRAKYFSCLDLMSGFHQIELDERSRNITSFST-------STGAYRYTRLPFGLKIAPNSFQRMMTLAFSGLTPSQA----
FLYMDDLVVIGCSEKHMLKNLTDVFKLCRQHNLKLHPEKCTFFMKEVTYLGHKCTDKGILPDDSKYEVIKNYPKPVNA
DEARRFVAFCNYYRRFIKNFA---------------------
>Gmr1
-------------------LKLGVIEE--SRSAWASPIVLVPKP---DGTHRFCNDFR-RLNEVSDFDSYPMPRVDELIERL--
GPARYLSTLDLTKGYWQVPLAPSSREKTAFAT-------PGGLFQYTVLPFGVHGAPATFQRMMDQVLRPH-SSYA----
AAYIDDIIIHSASWDEHVKHVRAVLNGLRAAGLTANPAKCRLGREETAYLGYRVGRGNVRPQEDKVAAIREWPQPQTK
KQVRSFLGLVGYYQRFIPGYA---------------------
>Woot
-------------------LDLGVIKR--EASPYASPMTVGKKK---DGTVRICLDAR-MINSKMIADCESPPAADELLRRF--
HEIRYMSTIDLRSSYWQIPLSPESRQYTAFLY-------NGRSYTYQVLPFGLKTAVGSFSRAMDVVLGTEVREFV----
VNYIDDLLVASETLNEHLEHLRQVFEKLKQARMTINLEKSNFIQKEVKFLGHILTINGIKADPEKISAIRNFPVPQKTKHV
RAFLGLCNFYRKFCARYA---------------------
>Tor4a
-------------------LENKLIEP--CSSPWNSPSLLVRKK---DGRFRLVIDYR-RLNDATLQMHHPLPNLEDSISYL--
EKSRVYSMCDMIKGFHQIDLDEDSKQKTAFSN-------EFGQFQYTRMPMGCKNAPSFFMRIMDKALLGVKKTEI----
IAYMDDLLCHSQTEVEHIKILEKLFQILAVNNLRINSKKAAFFTDSVNFCGYDISNGKICPSQDKIEAIQKLKIPRTKDAAQ
SLFGALNYHRRFIRNFA---------------------
>Kabuki
```

```
-------------------QELGICRP--SKIAWASPLHIVPKK---NGDLRPCGDYR-QLNAITKPDRYPIPRRHDFTYIL--
DDKKIFSKLDINRAFHCIDVAPEDIEKTAIIT-------PFGLFEFPKMTFGLRNAPQTFQRFMNNTVLHGLE-FL----
FSYLDDVIIASESPTQHNEHLRIVFERFNSYGITINLDKCCFGQPKVEFLGHEVSVNGITPLKNKVEAIINFPKPETVSDLRR
FLGMVNFYRPHLPNSA----------------------
>Cer4
-------------------LDLEVLEPI-THSDWAAPIVVIRKK---DGKVRVCADFKCSLNNSLIEEIHPLPTSDDLFGTL--QGC-
IFSKIDLKDAYLQIALDSESQKLAVINT-------HKGLFKYRRMTFGLKPAPAKFQKIIDKMIAGLPG--V----
AAYLDDVIVSANSLEEHEKVLHELLKRIKDYGFRISPEKCNFAQSEITFLGFIIDKRGRRPDPKKTSVIRSMKAPTDQKQL
MSFLGAICFYGRFVPKMS----------------------
>Mag
-------------------LAAGVIKPV-DHSDWATPLVVVRKA---DGGLRICADYKVTLNKVLAIDRFPVPKMEDLFSNL--
SGNKFFTKLDLSQAYNQIVLSERSSEYTVINT-------HRGLFKYSRLVYGLASSPGIFQKLMVNMFKNVPN--V----
VVFYDDILIRNQDLDSHLKSIKEVLDILERYGLKIKRSKCEFMVTEVRYLGFIIDQNGVRVDPEKVKSIATMPHPNNVTEL
KSFIGMVNFYSKFIQDLA----------------------
>SPM
-------------------GKAGVIKKV-KNSRWAAPIVTVPKS---DGSVRLCGDYKQTVNKVLEEDVYPLPTCEDLFANL--
AGGKVFSKIDLSNAYLQLELTDKSKELLTIHT-------HKGLYQYERLPFGVSTAPAVFQSVMDRILGGMKG--V----
CCYLDDILISSENEEDHIKTLNEVLGRLQEYGVKAKKEKCSFMVSRVIYLGHEIDGKGVHPTKEKVEGICDAKRPENKDE
LRTFIGIVVYYAKFVPNLA----------------------
>SURL
-------------------YELDVITPVDEPTDWVSSLVVVMKK---NGQLRVCLDPR-DLNRAIKREHYQLPSRAEITAHF--
AGAKYFSKLDASSGFWQIQLDDESSKLCTFIT-------PYGRYKFLRLPFGICSAPEVYHKIVHQMFAHIPG--V----
NTMMDDVIVWGTTQQEHDNRLREVLSIARRMNLKLNKDKCEFSVKKLTFIGDLISDQGVQPDPKKVSAILNMERPKCR
KDVQRFLGMINYQGKFIPDLA----------------------
>Athila4-1
----------------------------SDSTWVSPVHCVPKK---ITGHRMCIDYR-KLNAASRKDHFPLPFIDQMLERL--
ANHPYYCFLDGYSGFFQIPIHPNDQEKTTFTC-------PYGTFAYKRMPFGLCNAPATFQRCMTSIFSDLIEEMV----
EVFMDDFSVYGPSFSSCLLNLGRVLTRCEETNLVLNWEKCHFMVKEGIVLDHKISEKGIEVDKGKVEVMMQLQPPKTV
KDIRSFLGHAGFYRRFIKDFA----------------------
>Diaspora
----------------------------SDSQWVSPVQVVPKK---QNNWRVCIHYR-RLNQVTKKDHFPLPFIDQILECL--
AGKSHYCFLDGFSGYMQITIALEDQEKTTFTC-------LFGTFAYRRMSFGLCNAPGTFQRCMISIFSDFLENCI----
EEFMDDFTVYGSSFDGCLDSLEKVLNRRIETNLVLNFEKCHFMVEQGIVLGHIISNKGIEVDPAKISVISQLPYPSCV-
EVRSFLGHAGFYRCFIRDFS----------------------
>RetroSor1
-------------------LDAKVIREV-IYPEWLANVVLVPKK---NGKMRMCIDFT-DLNKACVKDSFPLPRIDTSVDKA--
AGCQRFSLLDCFSGYHQIWLKKEDEGKASFTT-------PFGTYCYTRMPEGLKNAGATFSRMMGKVLGSQLQRNI----
IAYVDDVVVMSKRKEDHIKDLQETFVNLRSAGLKLNPEKCVFGVSKGKMLGYIISSEGIRANPDKTKAIMSMAEPSNKK
EVQRLTGRIAALNRFISRSA----------------------
>Tat4-1
-------------------LGAGSIVEV-KYPEWLANPVVVKKK---NGKLRVCIDFT-DLNKACPKDSFPLPHIDRMVEAT--
TGNELLSFMDAFSGYNQIPMHKDDQEKTSFIT-------DRGTYCYKVMPFGLKNAGARYQRLVNQMFAPQLGKTM----
EVYIDDMLVKSKKSADHIEHLTACFETLNKYNMKLNPAKCSFGVTSGEFLGYIVTKRGIEANPKQIQAILDLKSPRNKKE
VQRLTGRIAALNIFIARST----------------------
>Cigr-1
-------------------LEEEIIEP--SNSPWRAQVLVAREG---SKKPRMVIDYSQTVNVFTELDAYPLPSIESIVNKV--
TQDKVYSSLDLRSAYHQIPLKDSDRPFTAFEA-------EGQLYQYKRLPFGVTNGVSAFQRIINRFIQDNHLNKV----
YAYLDDITVTGETLEEHDRNLKKLLVAAKRHNLTINEEKSKLRMTSIDFLGYRISYGDARPDPERLRPLLNLPPPDTPREL
KRIQGMFAYYAKWISRFS----------------------
>Tor1
-------------------EDAGIIEKA-SGSSFNAPLQLVRKS---SGGYRICVDMR-SLNNRLAESKWPLPSLAETLESL--
AGTAFFSCVDIRQAFFHMALTDESKHLTAFSA-------LNCQYQFRRLPMGLKISPSVYQMAMKETLGNDLGNKA----
VVYLDDVLVTGRTEDEHLEALDVVLDRLRKAGFLLNPDKCILGVKKTTFLGHEVTTEGYYPKTDNLAAIREFPKPTNK
KALRRFIGMTAFYSTLVPKLQ----------------------
>Tor2
-------------------AKDGIIEP--DGKGFHSPLVIVRKK---TGGLRICSDFKSSLNQCLDETTWPLPQMDHLFANI--
HGHRIFTSLDVSKAYWNLLIDPRDRHKTNFTF-------DNKCWMYVRLPFGLKFSGDAFCRSISSMLDRVSKSNF----
CNYVDDVLAYSSDCETHVKVLQQIFSACRHFGARLGAKKCTFGQASTAFMGRIISAEGISIPPENMETILALKPPTTRKQ
VQSLIGNFCWLKSWVSANL----------------------
>HERV-L
QYCIP-GGIAEISATIKDLKDAGVVIP--TTSPFNSPIWPVQKT---DGSWRMTVDYH-
KLNQVVTPIAVPVPDVVSLLEQINTSPGTWYAAIDLANAFFSIPVHKAHQKQFAFSW-------
```

```
QGQQYTFTVLPQGYINSPALCHNLIQRDLDDLDHLPQDITLVHYIDDIMLIGSSEQEVANTLDLLVRHLHARGWEINLTK
IQGPSTSVKFLGVLVQWCACDIVPSKVDKLLHLAPPTTKKEAQCLVGLFGFWRQHIPHLGVLLWPI---------------
>MuERV-L
QYCIP-GEIAEITATIKDLKDAGVVVP--TTSPFNSPIWPVQKT---DGSWRMTVDYR-
KLNQVVTPIAAAVPDVVSLLEQINTSPGTWYAAIDLANAFFSVPVHKDHQKQIAFSW-------
QGQQYTFTVLPQVYINSPALCHNLVRRDLDRLD-
LPQSITLVHYIDDIMLVGPSEQEVATTLDSLVTHMRIRGWEINPTKIQGPSTSVKFLG--
VQWCACRDIPSKVDKLLHLAPPTTKKEAQRLVGLFGFWRQHIPHLGVLLRPI---------------
>SRV-1
QWPLTSEKLAAAQQLVQEQLEAGHITE--SNSPWNTPIFVIKKK---SGKWRLLQDLR-AVNATMVLMGALQPGLPS-
PVAI--
PQGYLKIIIDLKDCFFSIPLHPSDQKRFAFSLPSTNFKEPMQRFQWKVLPQRMANSPTLCQKYVATAIHKVRHAWKQMYI
IHYMDDILIAGKDGQQVLQCFDQLKQELTIAGLHIAPEKIQ-LQDPYTYLGFELNGPKITNQKA----
VIRKDKLQTLNDFQKLLGDINWLRPYLKLTTADLKPLFDTLKGDSNPNS----
>MMTV
QWPLKQEKLQALQQLVTEQLQLGHLEE--SNSPWNTPVFVIKKK---SGKWRLLQDLR-AVNATMHDMGALQPGLPS-
PVAV--
PKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYI
VHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQ-KYDNLKYLGTHIQGDSVSYQKL----
QIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNGDSNPIS----
>JSRV
--------------------------------------------------------------MMHMGALQPGLPT-PSAI--
PDKSYIIVIDLKDCFYTIPLAPQDCKRFAFSLPSVNFKEPMQRYQWRVLPQGMTNSPTLCQKFVATAIAPVRQRFPQLYL
VHYMDDILLAHTDEHLLYQAFSILKQHLSLNGLVIADEKIQ-THFPYNYLGFSLYPRVYNTQLV----
KLQTDHLKTLNDFQKLLGDINWIRPYLKLPTYTLQPLFDILKGDSDPAS----
>LDV
QWPLTAQKLDAVQNIIQDLLKDGRIIP--SRSQWNSPIFVIQKK---DSKFRMLHDLR-AVNALIKDWGALQPGTPW-PGAI--
PSEWPVIAMDISDCFFSIPLAERDSERFAFTIPSPNLREPAKRYQWTVLPQGMKNSPYICQQVVAEVIRPIRERFRDAVIIH
YMDDILIAAAEERQTEVIFEAVKTTCHEKGLKINEAKTQ-RAPEVSYLGWRVMETGIAPLT-----
VQLPETVTNLVQLQRLLGIIQWLKPILSLRPEQLQVFYDLLKGKTRPEE----
>RSV
QWPLPEGKLVALTQLVEKELQLGHIEP--SLSCWNTPVFVIRKA---SGSYRLLHDLR-AVNAKLVPFGAVQQGAPV-LSAL-
-
PRGWPLMVLDLKDCFFSIPLAEQDREAFAFTLPSVNNQAPARRFQWKVLPQGMTCSPTICQLVVGQVLEPLRLKHPSLC
MLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQ-REPGVQYLGYKLGSTYVAPVG-----
LVAEPRIATLWDVQKLVGSLQWLRPALGIPPRLMGPFYEQLRGSDPNEA----
>HIV-1
QWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKK---STKWRKLVDFR-ELNKRTQDFWEVQLGIPH-PAGL-
-
KKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIY
QYMDDLYVGSDEIGQHRTKIEELRQHLLRWGLTTPDKKHQ-KEPPFLWMGYELHPDKWTVQPI----
VLPEKDSWTVNDIQKLVGKLNWASQIYPGIK--VRQLCKLLRGTKALTE----
>FIV
QWPLTNEKIEALTEIVERLEREGKVKRADPNNPWNTPVFAIKKK---SGKWRMLIDFR-ELNKLTEKGAEVQLGLPH-
PAGL--
QIKKQVTVLDIGDAYFTIPLDPDYAPYTAFTLPRKNNAGPGRRFVWCSLPQGWILSPLIYQSTLDNIIQPFIRQNPQLDIYQ
YMDDIYIGSNSKKEHKEKVEELRKLLLWWGFETPEDKLQ-EEPPYTWMGYELHPLTWTIQQK----
QLDIPEQPTLNELQKLAGKINWASQAIPDLS--IKALTNMMRGNQNLNSTRQW
>EIAV
QWPLTKEKLEGAKEIVQRLLSEGKISEASDNNPYNSPIFVIKKR---SGKWRLLQDLR-ELNKTVQVGTEISRGLPH-PGGL--
IKCKHMTVLDIGDAYFTIPLDPEFRPYTAFTIPSINHQEPDKRYVWKCLPQGFVLSPYIYQKTLQEILQPFRERYPEVQLY
QYMDDLFMGSNSKKQHKELIIELRAILLEKGFETPDDKLQ-EVPPYSWLGYQLCPENWKVQKM----
QLDMVKNPTLNDVQKLMGNITWMSSGIPGLT--VKHIAATTKGCLELNQKVIW
>HTLV-2
PISFKPERLQALNDLVSKALEAGHIEP--YSGPGNNPVFPVKKP---NGKWRFIHDLR-ATNAITTTLTSPSPGPPD-LTSLP-
TALPHLQTIDLTDAFFQIPLPKQYQPYFAFTIPQPCNYGPGTRYAWTVLPQGFKNSPTLFEQQLAAVLNPMRKMFPTSTI
VQYMDDILLASPTNEELQQLSQLTLQALTTHGLPISQEKTQQTPGQIRFLGQVISPNHITYEST---
PTIPIKSQWTLTELQVILGEIQWVSKGTPILRKHLQSLYSALHGYRDPRA----
>BLV
QFPLNLERLQALQDLVHRSLEAGYISP--WDGPGNNPVFPVRKP---NGTWRFVHDLR-ATNALTKPIPALSPGPPD-LTAIP-
THLPHIICLDLKDAFFQIPVEDRFRSYFAFTLPTPGGLQPHRRFAWRVLPQGFINSPALFERALQEPLRQVSAAFSQSLLVS
```

YMDDILIASPTEEQRSQCYQALAARLRDLGFQVASEKTRQTPSPVPFLGQMVHNQIVTYQSL---
PTLQISSPISLHQLQAVLGDLQWVSRGTPTTRRPLQLLYSSLKGIDDPRA----

Dataset S4. RT alignment generated by PROMALS3D and used to reconstruct the phylogenetic tree in Fig. 2.

```
>SURL
-----------------------VITPVDTDWVSSLVVVMKK-NGQLRVCLDPRDLNRAIKREHYQLPSRAEITAHFA--
GAKYFSKLDASSGFWQIQLDDESSKLCTFITP-------YGRYKFLRLPFGICSAPEVYHKIVHQMFAHIP------
GVNTMMDDVIVWGTTQQEHDNRLREVLSIARRMNLKLNKDKCEFSVKKLTFIGDLISDQGVQPDPKKVSAILNMERPK
CRKDVQRFLGMINYQGKFIPDLA------------------
>Cigr-1
-----------------------IIEPSNSPWRAQVLVAREG-SKKPRMVIDYSTVNVFTELDAYPLPSIESIVNKVT--
QDKVYSSLDLRSAYHQIPLKDSDRPFTAFEAE-------GQLYQYKRLPFGVTNGVSAFQRIINRFIQDNHL----
NKVYAYLDDITVTGETLEEHDRNLKKLLVAAKRHNLTINEEKSKLRMTSIDFLGYRISYGDARPDPERLRPLLNLPPPDT
PRELKRIQGMFAYYAKWISRFS------------------
>Amn-san
-------------------LQRGFIRPSTSPAGAGFFFVEKK-DGGLRPCIDYRGLNKITVKNRYPLPLISELFDQLK--
GAKIFSKLDLRGAYNLIRIREGDEWKTAFNTR-------DGHYEYLVMPFGLCNAPAVFQEFVNDIFRDLLG----
KSVVVVYLDDILIFSQDLETHRSQVKEALSRLRENFLFAKLEKCTFEVPKISFLGYIISSRGFEMDPAKVSAIQKWPLPQSTK
AIQRFIGFANYYRQFIKDFS------------------
>Sushi-ichi
-------------------LASGIIRPSSSPLAAGFFFVAKE-DGGLRPCIDFRKLNNITVKNKYPLPLMSSTFEPLT--
HARVFTKLDLRNAYHLVQIRKGDEWKTAFNTH-------LGHFEYLVMPFGLSNAPAVFQELVNDVLRDMIN----
VFVVVYLDDILIFSRTMEEHHQHVRLVLQRLLENRLFIKAEKCIFHSASVGYLGYIVEEGRVRADPAKIQAVVEWPRPTD
RTQLRRFLGFAGFIRRFIKGFA------------------
>Reina
-------------------LEQGVIQHSSSPFASPVLLVKKK-DGEWRLCVDYRRLNAHTVKNRYPMPVFDEIVDELC--
GTKIFTKLDHRSGYHQIRIKEGDEFKTAFQTH-------NGHYEYRVMPFGLTGAPATFQDFMNKILTPFLR----
KCVVVFLDDVLIYSRDMEEHVLQVKQVFQKLKDHQLKLKLSKCRFAQTTLEFLGHIISAEGIATDPEKVQVIRDWPIPNN
VKEVRSFLGMAGYYRRFVAHYA------------------
>Del
-------------------LNKGFIRGSTSPWGAHVLFDPKK-DDSKRMCIDY-KLNSVTVKNKYPLPRIDDLFDQLN---
GAYFSKIDLRFRYHQLRIRA-DIPKTAFRTR-------YGHYEFLVMPFGLTNVPTAFMNLMNRVFREYLD----
KFIVVFVDYVLIYSRTQKDHEHHLRISLQLLRNNQLYAKLSKCEFWMEKVKFLGHVVSREGIVVDPVKVKAVMNWELP
KNIFEIRSFLGLAGYYRRFIKGFA------------------
>CRM
-------------------LDKGYVRESLSPCAVPVILVPKK-DGTWRMCVDCRAINNITIRYRHPIPRLDDMLDELS--
GAIVFSKVDLRSGYHQIRMKLGDEWKTAFKTK-------FGLYEWLVMPFGLTNAPSTFMRLMNEVLRAFIG----
KFVVVYFDDILIYSKSMDEHVDHMRAVFNALRDARLFGNLEKCTFCTDRVSFLGYVVTPQGIEVDQAKVEAIHGWPMP
KTITQVRSFLGLAGFYRRFVKDFS------------------
>Galadriel
-------------------LDAGLIQPSKAPYGAPVLFQKKQ-DGTMRMCVDYRALNKATIKNKYSVPLVQDLMDRLS--
KACWFTKLDLRAGYWQVRIAEGDEPKTTCVTR-------YGSYEFLVMPFGLTNAPATFCNLMNNVLFDYLD----
DFVVVYLDDIVIYSRTLEEHVNHLSLVLSQLRKYTLYVKMEKCEFAQQEIKFLGHLVSKNQVRMDPKKVQAIVDWQAP
RHVKDLRSFLGLANYYRKFIAGYS------------------
>Skippy
-------------------IRKGYIRPSKSSAGFPVMFVPKP-NSNLRLVVDYRQLNEITEKDRTSLPLITELKDRLF--
GKKWFTALDLKSAYNLIRIKEADEWKTAFRTK-------YGLFEYLVMPFGLTNAPAVFQRMITNVLREYLD----
IFVVCYLDDILIFSDTEEEHTEHVHKVLKALQDANMLVEPTKSHFHQSQVTYLGHEISHNEIRMDRRKIAAVAEWKVPTS
VKETQSFLGFANYYRRFIKDFA------------------
>Maggy
-------------------LKKGFIRPSSSSVASPVLFVKKQ-GGGLRFCVDYRALNNITVKDRYPLPLVRETLNNLA--
GMKFFSKIDIVSAFNNIRIKKGEEYLTAFRTR-------FGLYESLVMPFGLTGAPATFQRYINDSLREYLD----
VFCTAYLDDILIYSRTRTEHEEHLKLVLEALRKAGLYANAAKCEFFVTETKFLGLLVGVEGVKMDPEKITAVLDWQTPK
KLTDVQAFLGFGNFYRRFIRDFA------------------
>Pyggy
-------------------LDKGFIRVSSLLAFLLVLFAKKL-EGGLRLCIDYWALNAITGKNRYLLPLIRETLNNIS--
KAKWFTKLDVIAAFYKIRVAEGDEWKTAFRTR-------FSLYEWLVTPFGMANSLSTFQRYINWTLREYLD----
EFCSAYLDDVLIYTDSLEQHQDHVRKVLRKLQESGLNVDIKKCEFGVKSTKYLGLIIDAEGIRMDPEKVKAIMEWE-------
--------------------------------------
>TF1
-------------------LKSGIIRESKAINACPVMFVPKK-EGTLRMVVDYKPLNKYVKPNIYPLPLIEQLLAKIQ--
GSTIFTKLDLKSAYHLIRVRKGDEHKLAFRCP-------RGVFEYLVMPYGISTAPAHFQYFINTILGEAKE----
```

```
SHVVCYMDDILIHSKSESEHVKHVKDVLQKLKNANLIINQAKCEFHQSQVKFIGYHISEKGFTPCQENIDKVLQWKQPK
NRKELRQFLGSVNYLRKFIPKTS-----------------
>Kabuki
-------------------QELGICRPSKIAWASPLHIVPKK-NGDLRPCGDYRQLNAITKPDRYPIPRRHDFTYILD--
DKKIFSKLDINRAFHCIDVAPEDIEKTAIITP-------FGLFEFPKMTFGLRNAPQTFQRFMNNTVLHGL-----
EFLFSYLDDVIIASESPTQHNEHLRIVFERFNSYGITINLDKCCFGQPKVEFLGHEVSVNGITPLKNKVEAIINFPKPETVSD
LRRFLGMVNFYRPHLPNSA-----------------
>Cer3
-------------------IQSGRIVESNTPWTSPIVLIKKK-NGSLRVCLDFRKLNEVTVPDNYPLPRIDSIIEQIG--
GSKYFTSLDMANGYLQLRLDAESSDKCGFITE-------NKVYAYTHLPFGLRSAASYFQRDLKQVLGLEK------
EVTVYIDDVLIFSKTFEEHLNSLRKVLERFRKFNLKVSPNKCEFFKQSITFLGHEINADNYKPNRINVDAIVNLPTPRNVG
DVRRFIGMSGFFRKFLPNFS-----------------
>Mdg3
-------------------LENGFIRPSDSEYASPIVLVKKK-TGDLRMCVDFRKLNKMTMKDNYPLPLIDDLLDRMN--
EKTVFTKLDLKNGFFHVHVKKESIKYTSFVTP-------LGQYEWLRMPFGLKNAPSVFQRFVNKIFADMIE----
NKVVVYMDDILLATENINEHLETLKEIFKRLVENKLELRIDKCEFMQSSIKYLGFIINKDGIMPNDKGIEAIKNFPIPNNVH
TVQSFLGLCSYFRRFIKDFS-----------------
>Cer4
-------------------LDLEVLEPITSDWAAPIVVIRKK-DGKVRVCADFKCSLNNSIEEIHPLPTSDDLFGTLQ---
GCIFSKIDLKDAYLQIALDSESQKLAVINTH-------KGLFKYRRMTFGLKPAPAKFQKIIDKMIAGLP------
GVAAYLDDVIVSANSLEEHEKVLHELLKRIKDYGFRISPEKCNFAQSEITFLGFIIDKRGRRPDPKKTSVIRSMKAPTDQK
QLMSFLGAICFYGRFVPKMS-----------------
>Mag
-------------------LAAGVIKPVDSDWATPLVVVRKA-DGGLRICADYKTLNKVLAIDRFPVPKMEDLFSNLS--
GNKFFTKLDLSQAYNQIVLSERSSEYTVINTH-------RGLFKYSRLVYGLASSPGIFQKLMVNMFKNVP------
NVVVFYDDILIRNQDLDSHLKSIKEVLDILERYGLKIKRSKCEFMVTEVRYLGFIIDQNGVRVDPEKVKSIATMPHPNNV
TELKSFIGMVNFYSKFIQDLA-----------------
>SPM
-------------------GKAGVIKKVKSRWAAPIVTVPKS-DGSVRLCGDYKQTVNKVEEDVYPLPTCEDLFANLA--
GGKVFSKIDLSNAYLQLELTDKSKELLTIHTH-------KGLYQYERLPFGVSTAPAVFQSVMDRILGGMK------
GVCCYLDDILISSENEEDHIKTLNEVLGRLQEYGVKAKKEKCSFMVSRVIYLGHEIDGKGVHPTKEKVEGICDAKRPEN
KDELRTFIGIVVYYAKFVPNLA-----------------
>Athila4-1
-----------------------------STWVSPVHCVPKK-GTGHRMCIDYRKLNAASRKDHFPLPFIDQMLERLA--
NHPYYCFLDGYSGFFQIPIHPNDQEKTTFTCP-------YGTFAYKRMPFGLCNAPATFQRCMTSIFSDLIE----
EMVEVFMDDFSVYGPSFSSCLLNLGRVLTRCEETNLVLNWEKCHFMVKEGIVLDHKISEKGIEVDKGKVEVMMQLQPP
KTVKDIRSFLGHAGFYRRFIKDFA-----------------
>Diaspora
-----------------------------SQWVSPVQVVPKK-TNNWRVCIHYRRLNQVTKKDHFPLPFIDQILECLA--
GKSHYCFLDGFSGYMQITIALEDQEKTTFTCL-------FGTFAYRRMSFGLCNAPGTFQRCMISIFSDFLE----
NCIEEFMDDFTVYGSSFDGCLDSLEKVLNRRIETNLVLNFEKCHFMVEQGIVLGHIISNKGIEVDPAKISVISQLPYPSCV-
EVRSFLGHAGFYRCFIRDFS-----------------
>Gypsy
-------------------LKDGIIRPSRSPYNSPTWVVDKK-NPNKRLVIDFRKLNEKTIPDRYPMPSIPMILANLG--
KAKFFTTLDLKSGYHQIYLAEHDREKTSFSVN-------GGKYEFCRLPFGLRNASSIFQRALDDVLREQIG----
KICYVYVDDVIIFSENESDHVRHIDTVLKCLIDANMRVSQEKTRFFKESVEYLGFIVSKDGTKSDPEKVKAIQEYPEPDCV
YKVRSFLGLASYYRVFIKDFA-----------------
>Tom
-------------------LEQGLIRESNSPYNSPTWVVPKK-KAKYRVVIDYRKLNEITIPDRFPIPNMDEILGKLG--
KCQYFTTIDLARGFHQIEMDSESIQKTAFSTK-------RGHYEYVRMPFGLRNAPATFQRCMNNILRPLIN----
KHCLVYLDDMIIFSTSLDEHLNSLQLVFEKLSESNLKLQLDKCEFLKKEATFLGHIVTPDGIKPNPLKVEAIASYPIPTKVK
EIRAFLGMTGYYRKFIPSYA-----------------
>Zam
-------------------LDQQIIRSSHSPWSAPVWVVPKK-NRKWRLVIDYRKLNDKTISDRYPIPNINDILDSIG--
KAKYFSTLDLTSGFHQIEMNPKDIAKTAFTVE-------GGHYEFTRMPFGLKNAPATFQRVMDSVLGDLNG----
TICLFYLDDIIIFSPSLQKHLLDIKMVFEKLRAANFKLQPSKSEFLRKEIEFLGHIVTQDGVKPNPNKISAIKKFPCPTNRRAI
KSFLGLLGYYRKFIRDFA-----------------
>Mdg1
-------------------IKDGIVEQSISEYNSPLLLVPKK-EKRWRLVVDYRQINKKLLADKFPLPRIEDILDQLG--
RAKYFSCLDLMSGFHQIELDERSRNITSFSTS-------TGAYRYTRLPFGLKIAPNSFQRMMTLAFSGLTP----
SQAFLYMDDLVVIGCSEKHMLKNLTDVFKLCRQHNLKLHPEKCTFFMKEVTYLGHKCTDKGILPDDSKYEVIKNYPKP
VNADEARRFVAFCNYYRRFIKNFA-----------------
```

```
>Tor4a
-------------------LENKLIEPCSSPWNSPSLLVRKK-DGRFRLVIDYRRLNDATLQMHHPLPNLEDSISYLE--
KSRVYSMCDMIKGFHQIDLDEDSKQKTAFSNE-------FGQFQYTRMPMGCKNAPSFFMRIMDKALLGVKK----
TEIIAYMDDLLCHSQTEVEHIKILEKLFQILAVNNLRINSKKAAFFTDSVNFCGYDISNGKICPSQDKIEAIQKLKIPRTKDA
AQSLFGALNYHRRFIRNFA------------------
>Gmr1
-------------------LKLGVIEESRSAWASPIVLVPKP-DGTHRFCNDFRRLNEVSDFDSYPMPRVDELIERLG--
PARYLSTLDLTKGYWQVPLAPSSREKTAFATP-------GGLFQYTVLPFGVHGAPATFQRMMDQVLRPHS-----
SYAAAYIDDIIIHSASWDEHVKHVRAVLNGLRAAGLTANPAKCRLGREETAYLGYRVGRGNVRPQEDKVAAIREWPQP
QTKKQVRSFLGLVGYYQRFIPGYA------------------
>Woot
-------------------LDLGVIKREASPYASPMTVGKKK-DGTVRICLDARMINSKMIADCESPPAADELLRRFH--
EIRYMSTIDLRSSYWQIPLSPESRQYTAFLYN-------GRSYTYQVLPFGLKTAVGSFSRAMDVVLGTEVR----
EFVVNYIDDLLVASETLNEHLEHLRQVFEKLKQARMTINLEKSNFIQKEVKFLGHILTINGIKADPEKISAIRNFPVPQKTK
HVRAFLGLCNFYRKFCARYA------------------
>RetroSor1
-------------------LDAKVIREVIPEWLANVVLVPKK-NGKMRMCIDFTDLNKACVKDSFPLPRIDTSVDKAA--
GCQRFSLLDCFSGYHQIWLKKEDEGKASFTTP-------FGTYCYTRMPEGLKNAGATFSRMMGKVLGSQLQ----
RNIIAYVDDVVVMSKRKEDHIKDLQETFVNLRSAGLKLNPEKCVFGVSKGKMLGYIISSEGIRANPDKTKAIMSMAEPS
NKKEVQRLTGRIAALNRFISRSA------------------
>Tat4-1
-------------------LGAGSIVEVKPEWLANPVVVKKK-NGKLRVCIDFTDLNKACPKDSFPLPHIDRMVEATT--
GNELLSFMDAFSGYNQIPMHKDDQEKTSFITD-------RGTYCYKVMPFGLKNAGARYQRLVNQMFAPQLG----
KTMEVYIDDMLVKSKKSADHIEHLTACFETLNKYNMKLNPAKCSFGVTSGEFLGYIVTKRGIEANPKQIQAILDLKSPRN
KKEVQRLTGRIAALNIFIARST------------------
>Tor2
-------------------AKDGIIEPCPKGFHSPLVIVRKK-TGGLRICSDFKSSLNQCTTDIWPLPQMDHLFANIE-
HGHRIFTSLDVSKAYWNLLIDPRDRHKTNFTFD-------NKCWMYVRLPFGLKFSGDAFCRSISSMLDRVSK----
SNFCNYVDDVLAYSSDCETHVKVLQQIFSACRHFGARLGAKKCTFGQASTAFMGRIISAEGISIPPENMETILALKPPTTR
KQVQSLIGNFCWLKSWVSANL------------------
>SRV-1
QWPLTSEKLAAAQQLVQEQLEAGHITESNSPWNTPIFVIKKK-SGKWRLLQDLRAVNATMVLMGALQPGLPSPVAIP---
QGYLKIIIDLKDCFFSIPLHPSDQKRFAFSLPSTNFKEPMQRFQWKVLPQRMANSPTLCQKYVATAIHKVRHAWKQMYII
HYMDDILIAGKDGQQVLQCFDQLKQELTIAGLHIAPEKIQL-QDPYTYLGFELNGPKITNQKAVIR----
KDKLQTLNDFQKLLGDINWLRPYLKLTTADLKPLFDTLKGDSNPNS
>MMTV
QWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKK-SGKWRLLQDLRAVNATMHDMGALQPGLPSPVAVP-
--
KGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIV
HYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQK-YDNLKYLGTHIQGDSVSYQKLQIR----
TDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNGDSNPIS
>JSRV
------------------------------------------------------------MHMGALQPGLPTPSAIP---
DKSYIIVIDLKDCFYTIPLAPQDCKRFAFSLPSVNFKEPMQRYQWRVLPQGMTNSPTLCQKFVATAIAPVRQRFPQLYLV
HYMDDILLAHTDEHLLYQAFSILKQHLSLNGLVIADEKIQT-HFPYNYLGFSLYPRVYNTQLVKLQ----
TDHLKTLNDFQKLLGDINWIRPYLKLPTYTLQPLFDILKGDSDPAS
>LDV
QWPLTAQKLDAVQNIIQDLLKDGRIIPSRSQWNSPIFVIQKK-DSKFRMLHDLRAVNALI-KDWGALQPGTPWPGAIP--
SEWPVIAMDISDCFFSIPLAERDSERFAFTIPSPNLREPAKRYQWTVLPQGMKNSPYICQQVVAEVIRPIRERFRDAVIIHY
MDDILIAAAEERQTEVIFEAVKTTCHEKGLKINEAKTQR-APEVSYLGWRVMETGIAPLTVQL-----
PETVTNLVQLQRLLGIIQWLKPILSLRPEQLQVFYDLLKGKTRPEE
>RSV
QWPLPEGKLVALTQLVEKELQLGHIEPSLSCWNTPVFVIRKA-SGSYRLLHDLRAVNAKL-VPFGAVQQGAPVLSALP--
RGWPLMVLDLKDCFFSIPLAEQDREAFAFTLPSVNNQAPARRFQWKVLPQGMTCSPTICQLVVGQVLEPLRLKHPSLC
MLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQR-EPGVQYLGYKLGSTYVAPVGLVA-----
EPRIATLWDVQKLVGSLQWLRPALGIPPRLMGPFYEQLRGSDPNEA
>HIV
QWPLTEEKIKALVEICTEMEKEGKISKIGNPYNTPVFAIKKK-STKWRKLVDFRELNKRT-QDFWEVQLGIPHPAGLK--
KKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQ
YMDDLYVGSDEIGQHRTKIEELRQHLLRWGLTTPDKKHQK-EPPFLWMGYELHPDKWTVQPIVLP----
EKDSWTVNDIQKLVGKLNWASQIYPGIK--VRQLCKLLRGTKALTE
>FIV
```

```
QWPLTNEKIEALTEIVERLEREGKVKRADNPWNTPVFAIKKK-SGKWRMLIDFRELNKLT-EKGAEVQLGLPHPAGLQ--
IKKQVTVLDIGDAYFTIPLDPDYAPYTAFTLPRKNNAGPGRRFVWCSLPQGWILSPLIYQSTLDNIIQPFIRQNPQLDIYQY
MDDIYIGSNSKKEHKEKVEELRKLLLWWGFETPEDKLQE-EPPYTWMGYELHPLTWTIQQKQLD----
IPEQPTLNELQKLAGKINWASQAIPDLS--IKALTNMMRGNQNLNS
>EIAV
QWPLTKEKLEGAKEIVQRLLSEGKISEASNPYNSPIFVIKKR-SGKWRLLQDLRELNKTV-QVGTEISRGLPHPGGLI--
KCKHMTVLDIGDAYFTIPLDPEFRPYTAFTIPSINHQEPDKRYVWKCLPQGFVLSPYIYQKTLQEILQPFRERYPEVQLYQ
YMDDLFMGSNSKKQHKELIIELRAILLEKGFETPDDKLQE-VPPYSWLGYQLCPENWKVQKMQLD----
MVKNPTLNDVQKLMGNITWMSSGIPGLT--VKHIAATTKGCLELNQ
>MuLV
QYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLP-
PSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWRDPE-
MGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGYRAS
AKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTL
FEW
>KoRV
QYPMSKEAREGIRPHIQRFLDLGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHPTVPNPYNLLSSLP-
PSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPE-
KGNTGQLTWTRLPQGFKNSPTLFDEALHRDLASFRALNPQVVMLQYVDDLLVAAPTYRDCKEGTRRLLQELSKLGYR
VSAKKAQLCREEVTYLGYLLKGGKRWLTPARKATVMKIPTPTTPRQVREFLGTAGFCRLWIPGFASLAAPLYPLTREKV
PFTW
>FLV
QYPMPHEAYQGIKPHIRRMLDQGILKPCQSPWNTPLLPVKKPGTEDYRPVQDLREVNKRVEDIHPTVPNPYNLLSTLP-
PSHPWYTVLDLKDAFFCLRLHSESQLLFAFEWRDPE-
IGLSGQLTWTRLPQGFKNSPTLFDEALHSDLADFRVRYPALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGYRAS
AKKAQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPKNSRQVREFLGTAGYCRLWIPGFAELAAPLYPLTRPGTLFQ
W
>REV
QYPITLEAKRSLRETIHKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLP-
PDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAE-
EGESGQLTWTRLPQGFKNSPTLFDVALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRV
SGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGTIGYCRLWIPGFAELAQPLYAATRGGNDP
LV
>Tor1
-----------------MEDAGIIEKASSSFNAPLQLVRKS-SGGYRICVDMRSLNNRLAESKWPLPSLAETLESLA--
GTAFFSCVDIRQAFFHMALTDESKHLTAFSAL-------NCQYQFRRLPMGLKISPSVYQMAMKETLGNDLGN----
KAVVYLDDVLVTGRTEDEHLEALDVVLDRLRKAGFLLNPDKCILGVKKTTFLGHEVTTEGYYPKTDNLAAIREFPKPT
NKKALRRFIGMTAFYSTLVPKLQ-----------------
>Loki-Xla
QYRLPLAAYESLSEIVKNLEKKGIIRPVHSSFNHPILGVLKP-NGQFRLCSDLRQLNKRVYMSGWPVPYIDQCLAQIQ--
GSKIFTALDCAQGYWTIKIDERDQHKLAFTFG-------KQQYAWTRLPFGYINAGHEFAVFMHKAMPDAAE----
RGTLSYVDDILIKSTTFEEHIAELKYVL--------
KLSLQKAQWCRSKVNFLGHEVTAEGINPQKKKVEAIKNTKSPTNLKELRSFLGMMNYSRKFIDNYAEITKPLLQLLKKG
AKWEW
>Loki-Xtr
QYRLPLAAYESLSEIVKNLEKKGIIRPVHSSFNHPILGVLKP-NGQFRLCSDLRQLNKQVYMSGWPVPYIDQCLAQIQ--
GSKIFTALDCAQGYWTIKIDERDQYKLAFTFG-------KQQYAWTRLPFGYINAGHEFAVFMHKAMPDAAE----
RGTLSYVDDILIKSTTFGEHIAELRYVLNQLKKAGVKLSLQKAQWCRSKVNFLGHEITAEGINPQ----------------------
VMNYSRKFIDNYAEITKPLLQLLKKGAKWKW
>Loki-Npa
QYRLPLACYDSLAEIIRNLKERGVIRSVHSSYNNPILGILKP-NGQWRLCADLRQLNKRVYMSGWPVPYIDQCLVQMQ--
GARIFTTLDCAQGYWTIKVSPEDQYKLAFTFG-------KEQYTWTRMPFGYINSGHEFAVFLHKAMPDATE----
RGNLTYVDDILLKSTSFEKHVQEIRHVLNQLREAGVKISLQKAQWCRTKVNFLGHEVTSEGLNPQKKKVEAVMKSKTP
TNVSELRSFLGMMNYSRKFIDNYAEISKPLLTLLKKDVPWMW
>Loki-Rca
QYRLPLASYESLAEIIRKLHERGIIRPIHSSYNHPILGILKP-NGQWRLCADLRQLNKRVYMSGWPVPYIDQCLAQLQ---
SQVFSALDCAQGYWTIKVSEEDQYKLAFSFG-------NQQYCWVRMPFGYINSGHEFAVFMHKAMPDALE----
RGTLAYVDDILIKSTDMTKHLTELRHILGQLR-
AGVKLSIQKAQWCRTRVNFLGHEVTSEGLNPQKKKVEAVFNCKVPTNLKELRSFLGMVNYSRKFIDNYAEISKPLLHLL
KKDTEWSW
>Loki-Sri
```

```
QYKIPLASYEPVQEIIDNLLEKGIIRPCNSTYSAPLWPVLKP-NGKWRLTIDYRKLNQQVPLSRWPMTQLEQELPKIR--
DAKYFSTLDIASGFWTIPVHVADQHKLAFTFA-------NRQYTFTRCPFGYANSPAEFNIFLNKACPDAND----
RGTLIYVDDILMRSSTLDTHLEEIDHVLEQLTAAGAKISLSKCQWCKTKVNYVGLLVGPRGVEPQINRVQGVTNIKAPT
NVSELRSFLGVCNYSRQFIENYADLAKPLTELLRKDALFVW
>Loki-Ape
QYKIPLAAYESIQEILDKLLEKRIVRECNSTYNSPIWGALKP-TGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVK--
EARFFSTVDVANGFWTMTVDPVDQYKLAFSFC-------NRQYTWNRCPFGYSNSPAEFNIFLHKAMSDASA----
RGNLIYVDDILMRSRSFDDHLAEIRHVLGQLAAAGAKLAIAKGQWCRTKVEYVGLTVGANGIEPKAERIRAIQDIMAPV
NVSKLQSFLGVCNYSRQFIADYAEIARPLTELLGKDKPFLW
>Loki-Pfl
QYKIPLAAYESIQEILDKLLEKQIVRECNSTYNSPIWPVLKP-TGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVK--
EARFFSTVDVANGFWTMRVDPVDQYKLAFSFC-------NRQYTWNRCPFGYSNSPAEFNIFLHKAMSDASA----
RGNLIYVDDILMRSRSFDEHLAEIRHILGQLAAAGAKLAITKGQWCRTKVEYVGLTVGANGIEPKAERMRAIQDIMAPV
NVSKLRSFLGVCNYSRQFIADYAEIARPLTELLGKDKPFLW
>Loki-Pyo
QYKIPLAAYESIQDILDTLLEKRIIRECNSTYNSPIWPVLKP-TGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVK--
GATFFSTVDVANGFWTMKVDQADQYKLAFSFG-------GSQYTWNRCPFGYSNSPSEFNIFLHKAMSDAAT----
RGNLIYVDDILMRSQTFEDHITEIRHVLTQLSNAGAKLALTKGQWCRTKVKYVGLTVGADGIEPQAGRIRAIQDIKAPTR
LTELRSFLGVCNYSRQFIEDYAEIARPLNELLHKDKPFEW
>Loki-Sau
QYRIPLAAYESIQDILDKLLQKQIIRECNSTYNSPIWPVLKP-TGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVK--
GACFFSTVDIANGFWTMKVDPADQYKLAFSFC-------NRQYTWNRCPFGYSNSPAEFNIFLHKAMSDAAA----
RGNLIYVDDILMRSCTFETHLAEIRHVLNQLTRAGAKLAIAKGQWCRTKVEYVGLTVGGSGIEPQTGRIRAIQDIKAPTN
VSELRSFLGVCNYSRQFIEDYAEIARPLTELLCKDKLFQW
>Loki-Ofa
QYRIPLAAYESIQEILDNLLQKQIIRECNSTYNSPIWPVLKP-TGKWRLTIDYRPLNKQVPLSRWPMIHLDQELAKVK--
GACFFSTVDVANGFWTMRVNPADQYKLAFSFG-------NRQFTWNRCPFGYSNSPAEFNIFLHKAMSDAAA----
RGNLIYVDDILMRSRTLE-
HLAEIRHVLNQLSAAGAKLAIAKGQWCRTKVEYVGLTVGADGIEPQAGRIRAIQDIKAPTGVSELRSFLGVCNYSRQFIE
DYAEIARPLTELLRKDKSFEW
>Loki-Bsp
QYRIPLAAYESIQEILDTLLKKEIIRECNSTYNSPIWPVLKP-TGKWRLTIDYRPLNKNVPLSRWPMIHLDQELAKIK--
DAHFFSTVDVANGFWTMRVDPADQYKLAFSFG-------NRQYTWNRCPFGYSNSPAEFNIFLHKAMSDAAA----
RGNLIYVDDILMRSQTFEGHLAEIRHVLSQLLAAGAKLALTKGQWCRTKVKYVGLTVGQHGIEPQADRIQAIRTIKAPN
TLSGLRSFLGVCNYSRQFVEDYAEIARPLTELLRNDTPYIW
>Loki-Cne
QYKIPLAAYGSIQEILNKLLEKNIIRECNSTYNSPIWPVLKP-TGKWRLTIDYRALNKQVPLSRWPMIHLDQELAKVK--
DARFFSTVDVANGFWTMKVDPADRYKLAFSFG-------NRQY--NRCPFGYSNSPAEFNIFLHKAMSDATE----
RGNLIYVDDILMRSRTFKEHLEEI-
HVLGQLAAAGAKLALSKGQWCRKKVDYVGLTVGPNGIEPQADRIRAIQDIAAPTNLSELRSFLGVCNYSRQFIEDYAD
MARPLTELLRKNTPFDL
>Loki-Lch
QYRIPLSSYESVQEILHTLMEREIIRECNSTYNSPIWPVLKP-NGKWRLTIDYRQLNKQVPLSHWPMTYLDQNLTKIA--
GAKYFTTLDIANGFWTVKVDPKDQYKLAFSFG-------KRQYTWNRCPFGYLNSPADFNIFLHKAMPDAEE----
RGNLIYVDDILIKSNSASDHLDEIRYVLTQLKNAGAKLSLTKGQWCRKRVNYLGYDVGPDGLRPQQNRIKAIQAIRSPT
NITELRSFLGICNYSRQFIENYAEISQPLTKLLRKNEPWVW
>Loki-Str
QYKIPLAAYAAVQEIIDSLLAKRIIRECNSTYSAPVWPVLKP-TGKWRLTIDYRQLNKLVPLSRWPMTQLDQELPKVA--
NAKYFSTVDVANGFWTMTVDPRDQHKLAFSFS-------NKLFTFNRCPFGYANSPSEFNIFLHKAMPDAAS----
RGTIIYVDDVLMRSETWSHHLNEMDHVLTQLGTAGAKLAIMKGQWCRTKVNYVGLLVGAEGILPQSNRIQAVRNIKTP
TNLHEVRSFLGVCNYSRQFIENYADLSKPLTHLLQKDTPFVW
>Loki-Tfu
QYKIPLESYEPVQEIINSMLEKGVIRLCNSTCSAPIWPVLTP-NGKWRPTIDYRKLNQQVSLSRWPMTQLEQEIPRVK--
GATIFSTLDVASGFWTIPVHPDDQHKLAFMFC--------NRYTFTRCPFGYANSPAEFNIFLNKACPDARV----
RGNLIYVDDVLMKSTTMTDHLKEIDHVLQQLTKAGAKIALHKGQWCKTKVNYVGPLIAPQGIERQTNRIQAIQNIKPPA
NISELRSFLGVCSYSRQFIKDYSDIARPLMALLKKDSPFTW
>Loki-Lro
QYKIPLESYEPVQDIITSMLEKGVIRPCNSTYSAPIWPVLKP-NGKWRPTIDYRKLNQQVPLSRWPMTQLEQEIPRIR--
GATIFSTLDVASGFWTIPVHPDDQHKLAFRFD-------NRQYTFTRCPFGYANSPAEFNIFLNKACPDARV----
RGNLVFVDDILLKHTTVADHLKEIDHILNQLTTAGAKIALHKGQWCKIKVNYVGLLIGPQGIEPQTNRIQAIQNIKPPAN
VSELRSFLGVCNYSRQFIENYSDIARPLTALLKKDCPFIW
>Loki-Dcl
```

```
QYKIPLESYEPVQEIIDSMLEKGVIRPCNSTYSAPIWPVLKP-NGKWRPTIDYRKLNQQVPLSRWPMTQLEQEIPRIR--
GATIFSTLDVASGFWTIPVHPDDQHKLAFTFS-------NRQYTFTRCPFGYANSPAEFNIFLNKACPDARV----
RGNLIYVDDVLMKSMTVADHLKEIDHVLHQLTAAGAKIALHKGQWGKSKVNYVGLLIGPQGIEPQSSRIQAIQTIKPPV
NVSELRSFLGVCNYSRQFIENYSDIARPLTTLLKKDCPFAW
>Loki-Sme
QYKIPLASYEPVQEIIDDLLHRGIIRPCNSTYSAPVWPVLKP-NGKWRLTIDYRKLNQQVPLSRWPMTQLNQELPKVK--
DAKYFSSMDIASGFWTIPVHIADQHKLAFTFA-------GRQYTFTRCPFGYANSPAEFNIFLNKSCPDASE----
RGTLIYVDDILIRTPTMKGHIDEIEHVLDQLTTAGAKISLSKCQWCKTKVNYVGLLVGSNGVEPQTNRVQGVMNIKTPT
NISELRSFLGVCNYSRQFIPDFADIAKPLTDLLKKDSLFTF
>Loki-Oja
QYKIPLASYEPVQEIIDDLTNKGIIRPCNSTYSAPIWPVRKP-NGKWRLTIDYRRLNQQVPLSRWPMTQLEQELPRVK--
EAKYFSTLDVASGFWTIPVHPDDQHKLAFTFA-------RQQYTFTRCPFGYANSPAEFNIFLNKACPDATQ----
RGTLIYVDDILIRNATLEDHIEEIEHVLKQLTSAGAKISLSKCQWCRTKVNYVGLLVGPTGVQPQVGRVQGLKDIAAPTN
LKELRSFLGVCNYSRQFIEDYADIAKPLTELLKKDVPFRW
>Loki-Ame
QYKIPLASYGPVQEIIDDLLDKGIIRPCNSTYSAPLWPVLKP-NGKWRLTIDYRRLNDQVPLSRWPMTQLEQELPRVR--
DAKYFSTLDVASGFWTIPVHVEDQHKLAFTFA-------GRQFTFTRCPFGYSNSPAEFNIFLNKACPDARE----
RGTLIYVDDVLIRNNSLDAHLEEIDHVLDQLTKAGAKISLAKCQWCKTKVNYVGLLVGPDGVLPQPCRAQGIVDIAEPK
TIHALRSFLGVCNYSRQFIENFAELAKPLYQLLKQDVPFIW
>Loki-Bpe
QYKIPLASYEPIQEIIDDLLTKGVIRPCNSTYSAPIWPVLKP-NGKWRLTVDYRQLNTRVPLSRWPMTQLGQDLPKVR--
DAKYFSTLDIASGFWTIPVHPDDQHKLAFSFA-------NRQYTFTRCPFGLANSPAEFNIFLNKACPNAAS----
RGTLIYVDDILIRTTTLDDHLKEIDHVLDQLTSAGAKISLSKCQWCKTKVNYVGLL--------
RSSRIQGITNLKVPANVSELRSFLGVCNYSRQFVEHYAELAKPLTDLLKKDETFSW
>Loki-Lca
QYKIPLAAYESIQETLENLLKKNIIRECNSTYNSPIWPVLKP-TGKWRLTIDYRQLNKTVPLSRWPMAEIDHGLNQIK--
GAKVLTTMDLANGFWTMPVKETDQYKLAFTFD-------RIQYTWNRCPFGYSNSPADFNIFLHKAMGDAKE----
RGTIVYVDDILVKDSSWQEHLKSLQHTLEQLKEAGAKISIQKGQWARKRVDYLGFQIGTEGLLPQTNRIEALLALKSPNT
VPTLRSFLGICNYLRQFVDDYAGITRPLVKLLQKDEPWEW
>Loki-Pma
QYKIPLAAYESIQETLENLLKKNIIRKCNSTYNSPIWPVLKP-TGKWRLTIDYRQLNKTVPLSRWPMAEIDQNLNQIK--
GAKILTTMDLSN-FWTMPVKEADQYKL--TFN-------GIQYTWNRCPFGYSNSPAD-NIFLHKAMGDAKE----
RGTIVYVDDILVKDSSWEEHLESLQHTLEQLE--AAKISIQKGQWARKRVDYLGFQIGT-
GMLPQTKRLEALLALKPPTTVTTLRSFLGICNYLRQFVDDYAGIARPLVKLLQKDEPWEW
>Loki-Aro
KYRIPIAAHDPVQDTITEILHAGIIRKCNSTYNSPIWPVLKP-SGKQSLTVDYRKLNKQVPLSRCLWPTWMKPLQKS---
RMPHTSQADISNGFCTIPVREQDQHKLAFSFN--------NVQTWTVTPFGYANSPAEFNIFFHKAIPDVNS----
KNIAIYVDDILMWSDSWGEHQQLLSYVLTRLSQAGAKISLDKGQWCKRAVNYLGFQVGGDGIRPRTEAP---
KAIPAPTTLKQLCSFLGICNFSQHFVADYAKISQPLTRLLKKDVPWTW
>ZFERV-2
QYPINKSALPEIKITIEELKRKGIIKEVDAPTNSPIQAVAKP-DGGWRLVTNYKALNKQTVPDTRYLINARDATGAAP--
EGRLLSKIDLANGFWSVPLKPESRARTAFTFE-------NKQYVYNVLPQGFKNSPNAFQAIMMEILKGL-------
PIVVYIDDILIVTNDEEEHLKILDETLSRLTKEGFRLNYKKMEIAKDDVEFLGFQITGNERFLSQTTENKLKELKEITTLKQ
LQSMLGTLNFVRDLIPGYARHAKPLYNATKGGV-LTW
>AciFLERV
QYPLNPGAVKEMDLIVRELLTLGVIRQELPITNSPIQAVKKP-GGGWRPVINFKALNRRTVANRASLINPQGTLKTLK--
LKPFKSCIDLANGFFSLRLTRGSQGKTAFTHK-------GRSYVWERLPQGYKNSPNVFQAAVMDVLDGL-------
GATIYIDDVFIADDTEEQHLDRLQKIVERLTAAGLKLNLKKCQFGRKRVDYLGFQVSD-
DLGLSEVYRKKIEQIEPPSSLNELQKVLGLCNYVRDYVPGYQRYAKPLYAKLKKPDETAA
>AliFLERV
QYPLHPEAVAEMDKIVKELHALGIIREEPPLTNSPIQAVKKP-GGGWRPVINFKALNRRTVANRASLINPQGALKTLR--
VKKYKSCIDLANGFFSLRLARESQGKTAFTHK-------GKSYVWQRLPQGYRNSPNVFQSAVLEILEDV-------
GASVYIDDVFIADDTEEEHLKRLEEVVKRISAAGLKLNLAKCQFGQFKVNYLGFQVTT-
DLGLSDGYREKIQQIQPPRCENDLQKILGLCNYVRDHVPNYQKYAKPLYACLKKEATTEG
>EFV
QYRINPKAKADIQIVIDDLLKQGVLKQQTSPMNTPVYPVPKP-DGRWRMVLDYRAVNKVTPAIATQNCHSASLLNTLY--
RGQYKTTLDLANGFWAHPIQESDQWITSFTWN-------GKSYVWTTLPQGFLNSPALFTADVVDLLKDIP------
NVEVYVDDVYFSNDTEEEHLKTMDLLFQKLQTAGYIVSLKKSKLGQHTVDFLGFQITQTGRGLTDSYKSKLLDITPPNT
LKQLQSILGLLNFARNFIPNYSELITPLYQLIPLAKGIYI
>FFV
QYHINPKAKPDIQIVINDLLKQGVLIQKESTMNTPVYPVPKP-NGRWRMVLDYRAVNKVTPLIAVQNQHSYGILGSLF--
KGRYKTTIDLSNGFWAHPIVPEDYWITAFTWQ-------GKQYCWTVLPQGFLNSPGLFTGDVVDLLQGIP------
```

```
NVEVYVDDVYISHDSEKEHLEYLDILFNRLKEAGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENITAPTTLKQ
LQSILGLLNFARNFIPDFTELIAPLYALIPKSTKNYV
>BFV
QYHINPRAKADIQIVIDDLLRQGVLRQQNSEMNTPVYPVPKA-DGRWRMVLDYREVNKVTPLVATQNCHSASILNTLY-
-RGPYKSTLDLANGFWAHPIKPEDYWITAFTWG-------GKTYCWTVLPQGFLNSPALFTADVVDILKDIP------
NVQVYVDDVYVSSATEQEHLDILETIFNRLSTAGYIVSLKKSKLAKETVEFLGFSISQNGRGLTDSYKQKLMDLQPPTTL
RQLQSILGLINFARNFLPNFAELVAPLYQLIPKAKGQCI
>SFV
QYHINPKAKPSIQIVINDLLKQGVLRQSTSPMNTPVYPVPKP-DGKWRMVLDYRAVNKTIPLIAAQNQHSLGILTNLI--
RHKYKSTIDLSNGFWAHPITEDSQWITAFTWE-------GKQHVWTRLPQGFLNSPALFTADVVDILKEVP------
GVSVYVDDIYISSPTMEEHFQVLDSIFRKLLETGYIVSLKKSALARYEVNFLGFVISETGRGLTSEFRERLQEITPPTTLKQ
LQSILGFLNFARNFVPNFSELVQPLYQLISTASGNFI
>CoEFV
QYPINTKAIPSIQVVINELLEQGVLVKQTSPTNMAVLPVPKP-DGTWRLVLDYRALNKHSEPVRAQNQHSSGILANIE--
RKAYKSSVDLANGFWSHPIREEDWPKMAFTWC-------GFQYLWTRLPQGFLNSPALFSADVISLVGQLP------
GVYCYVDDIYLTHDTEEEHLKILDQVLEILIKAGYVINIKKSKLCRKVVDFLGFSLSDEGRGLSDQYREKLAAIKPPQTLR
QLQSVMGLLNFSRLFVKGFSELAKPLYDL---------
>WDSV
QYPLPKDKTEGLRPLISSLENQGILIKCHSPCNTPIFPIKKAGRDEYRMIHDLRAINNIVAPLTAVVASPTTVLSNLA-
PSLHWFTVIDLSNAFFSVPIHKDSQYLFAFTFE-------GHQYTWTVLPQGFIHSPTLFSQALYQSLHKIKFKIS-
SEICIYMDDVLIASKDRDTNLKDTAVMLQHLASEGHKVSKKKLQLCQQEVVYLGQLLTPEGRKILPDRKVTVSQFQQPT
TIRQIRAFLGLVGYCRHWIPEFSIHSKFLEKQLKKDTAEPF
>Xen-1
QYPLKPAQSESINKQLQTLLENGAIKRQSSPCNTPLFPVKEKGEDKYRLVQDLRAVNEATVMETPLVSNPHTILSGIP-
PSATHFRAVDLTNAFYSIPLREDCQYLFAFTHE-------
RQQYVWTVLPQGAQNSPTHFSLALTSILDSWISSHPEITLLQYVDDLLLCAPDLPTCEASSTDLLSFLADQGCKASKEKL
QWCQTTVVFLGQCISQGTRHITEGRIKTLQDIPLPKGHKPLHAFLGLISYCRSWIPEASLLMQPLYDVLKSDPFTM-
>SnRV
QYPVPDASHASIKETVEALLEQGVLRKCNSTVNSAIWPVGKP-DGSWRLTIDYRPLNSAVSCPYPTVASTPELFAKLE-
KKYQVYSSLDISNGFWSIRLEEECQYLFAFTFD-------
TQQYTWTRLPQGFHASPGIFHQALYNGLASCKTESQGCKLLQYVDDILLMSEDRDHHLRSLAILLQGLKDLGVKINPKK
SHFCKDQVQYLGVNVGADTRSLIDARSQLIRTLDIPLTVQGLRSALGLFNFCRAWIPEFSRKTQSLYDMLKGDCKDKL
>HERV-L
QYCIP-GGIAEISATIKDLKDAGVVIPTTSPFNSPIWPVQKT-
DGSWRMTVDYHKLNQVVIAAAVPVPDVVSLLEQINTSPGTWYAAIDLANAFFSIPVHKAHQKQFAFSWQ-------
GQQYTFTVLPQGYINSPALCHNLIQRDLDDLDHLPQDITLVHYIDDIMLIGSSEQEVANTLDLLVRHLHARGWEINLTKI
QGPSTSVKFLGVLVQWCACDIVPSKVDKLLHLAPPTTKKEAQCLVGLFGFWRQHIPHLGVLLWPIYQVTQKAASFEW
>MuEV-L
QYCIP-GEIAEITATIKDLKDAGVVVPTTSPFNSPIWPVQKT-
DGSWRMTVDYRKLNQVVTPIAAAVPDVVSLLEQINTSPGTWYAAIDLANAFFSVPVHKDHQKQIAFSWQ-------
GQQYTFTVLPQVYINSPALCHNLVRRDLDRLD-
LPQSITLVHYIDDIMLVGPSEQEVATTLDSLVTHMRIRGWEINPTKIQGPSTSVKFLG--
VQWCACRDIPSKVDKLLHLAPPTTKKEAQRLVGLFGFWRQHIPHLGVLLRPIYQVTRKAASFVW
>HTLV-2
PISFKPERLQALNDLVSKALEAGHIEPYSGPGNNPVFPVKKP-NGKWRFIHDLRATNAITTTLTSPSPGPPDLTSLPT--
ALPHLQTIDLTDAFFQIPLPKQYQPYFAFTIPQPCNYGPGTRYAWTVLPQGFKNSPTLFEQQLAAVLNPMRKMFPTSTIV
QYMDDILLASPTNEELQQLSQLTLQALTTHGLPISQEKTQQTPGQIRFLGQVISPNHITYESTPTI---
PIKSQWTLTELQVILGEIQWVSKGTPILRKHLQSLYSALHGYRDPRA
>BLV
QFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKP-NGTWRFVHDLRATNALTKPIPALSPGPPDLTAIPT--
HLPHIICLDLKDAFFQIPVEDRFRSYFAFTLPTPGGLQPHRRFAWRVLPQGFINSPALFERALQEPLRQVSAAFSQSLLVSY
MDDILIASPTEEQRSQCYQALAARLRDLGFQVASEKTRQTPSPVPFLGQMVHNQIVTYQSLPTL---
QISSPISLHQLQAVLGDLQWVSRGTPTTRRPLQLLYSSLKGIDDPRA
```

Dataset S5. RH alignment generated by Mafft 7.402 and used to reconstruct the phylogenetic tree in Fig. 3.

```
>Athila4-1
EIMCDASDYA------------VGAVLGQKID------KKLHVIYYASRTLDD--AQGRYATTEKELLAVVFAFEKFRSYL-VG--------
SKVTVYTDH-ALRHLYAKKD--------------TKPRL--LRW-----ILLL--------QEFDM-EIVDKKGI----------ENGAADHLSR----
>Diaspora
ELMCDASNYA------------LGAVLAQKID------KLPRVIYYASRTLDA--AQANYTTTEKELLAIVFALEKFRSYL-LG--------
THIIVYTDHAALKYLLKKVD--------------SKPRL--IRW-----MLWL--------QEFDL-EIRDRSGA----------QNLVADHLSR----
>Gmr1
TLQTDASEVG------------LGGVLSQARN--E---REEHPVTYLSRKLLP--HERNYSTVEKEALAIKWAVNKLTYYL-LG-------
-HQFVLVTDHAPLKWMATAKD--------------TNARI--TRW-----FLSL--------QPFSF-TVEHRPGR----------EHTNADALSR----
>Mdg1
CITTDASKQA------------CGAVLSQDHN------GQQLPVAYASRSFTK--GESNKSTTEQELAAIHWAINHFRPYV-YG--------
RHFLVQSDHRPLSYLFSMRN--------------PSSKL--TRM-----RLDL--------EEFEF-TVEYLKGK----------DNHVADALFR----
>Zam
ILTTDASNFA------------LGAVLSQGSL------QNDRPVCFASRTLSD--TEVNYSTIEKEMLAIIWAVKYFRPYI-YG--------
VKFTIVTDHKPLIWLMNFKE--------------PNSKI--IRW-----RLQL--------MEYNF-EIIHKKGS----------QNVIADALSR----
>Tom
VLTTDASNLA------------LGAVLSQ----------DNHPISFISRTLND--HELNYSTIEKELLAIVWATKTFRHYL-LG--------
RHFQIASDHQPLRWLHNLKE--------------PNAKL--QRW-----RIRL--------AEFDF-HIEYIKGK----------QNSIADALSR----
>Gypsy
DLTTDASASG------------IGAVLSQ----------EGRPITMISRTLKQ--PEQNYATNERELLAIVWALGKLQNFL-YG--------
SREINIFTDHQPLTFAVADRN--------------TNAKI--KRW-----KSYI--------DQHNA-KVFYKPGK----------ENFVADALSR----
>Woot
YVQTDSSGYG------------LGAELYQIQE--D---GSRGVIAFASRSLRG--PELNYTTTEKELLGVIFALHKFRIYI-QV--------
TKIIIRTDHQALKFLSRCRL--------------LSERL--TRW-----TLIL--------GQYDY-EIELVKGK----------GNVVADILSR----
>Mag
VLTVDASARG------------LGAVLAQRGP--G---CQERVVAYASRALTT--HELHYSQIHKEALAIVFAVEKFHQYL-YG-------
-RKFILRTDHKPLVSIFGPNIGI---------PSAAASRL--QRW-----AIKL--------SAYDF-EIEYVRTD----------KN-VADALSR----
>SPM
VLACDASPYG------------LGVVLSHLME--D---GEERPVAYASRTLTK--SEQNYSQIEKESLAIIYGVTKFHKYL-YG--------
RKFTLLTDHQALTIIFGSKKGI---------PSLAAARL--QRW-----ALIL--------MAHQY-EIKYRKST----------EHANADVLSR---
-
>Cer4
IVAADASKYG------------IGGVILHVNP--D---GVEVPIAHFARSLTE--TEKRYSQIEKEALALIYTVKKSHKFV-FG--------
RRFKLQTDHRPLLALFGDNRDL---------PVHSQNRI--VRW-----ATTL--------MSYDF-ELSYVATE----------
KFAKADWLSR----
>Galadriel
EVHTDASDKA------------IGGVLVQ----------EGHPVAFESRKLND--AEQRYSTHEKEMVAVVHCLQVWRVYL-LG--------
TRFVVRTDNVANTFFKTQKK--------------LSPKQ--ARW-----QEFL--------AEYDF-MWEHKPGK----------HNQVADALSR--
--
>CRM
ELECDASGIG------------LGGVLLQ----------EGKPVAYFSEKLSG--SVLNYSTYDKELYALVRTLETWQHYL-WP--------
KEFVIHSDHESLKHIRSQGK--------------LNRRH--AKW-----VEFI--------ESFPY-VIKHKKGK----------ENIIADALSR----
>Del
VVYTDASLAG------------LEGVLMQ----------DGRVVAYASRQLKV--HENNYPTHDLELAVVIFILKLWRHYL-YG--------
EDFELYCDHKSLKYISTQKD--------------LNLR---QRW-----IEVL--------KDFDF-SIFYHPGK----------ANVVADALSR----
>Cigr-1
QVETDASDYA------------IAAVLSQ----------SGRPVAYMSRTLNT--CERNYPAIEKEATAVVEAVRKWSHFL-KG--------
KSFTLVTDQRSVSFMFDQRN--------------KNSKI--LMW-----RLEL--------SQFTY-DIVHRPGR----------DNFVPDALSR----
>Reina
VVETDASGSG------------IGAVLQQ----------KGHPIAYISKALGP--KNLGLSTYEKEYLAILFAVDHWRPYL-QH--------
GEFFIKTDQQSLTHLEDQKL--------------STIWQ--QKA-----ITKL--------LGLQF-RIIYKKGV----------ENRVADALSR----
>Cer3
HIFTDASAVA------------QGAVLMQQME--D-STKDYAAIAYTSRTLSD--TESRWPAIQTELGAIIFALRQFRPYI-GQ--------
SRTTIHSDHRPPMYLLGKSK--------------VNDNL--ARW-----LIEL--------AQYDT-RIVHIDGK----------KNTVADCLSR----
>SURL
KLSADASKNG------------IGAVLLQQHD------ENWVPIAYASRSMTD--AETRYAQIEKELLAITYACERFHQYI-YG--------
QQVEVETDHKPLIPLFVKSL--------------CPLRI--QRL-----LIRV--------QRYDL-KVMYTPGK----------YMYTADTLSR----
>Sushi-ichi
```

```
IVEVDASDAG------------IGAVLSQRSE-AD---QKIHPCAYFSRRFDP--AERNYDVGNRELLAVYGALVEWKHWL-EG-------
--HPFLVWSDHKNLTYVRTAKR---------------LNPRQ--GRW-----AL-F--------SRFDF-TLTFRPGS----------KNIRADALSR----
>Amn-san
FIEVDASDVG------------AGAILSQRHS-AD---GKLHPCAYFSKKFSS--AEQNYDIGNRELLAVKLALEEWRHLL-EG--------
HPVTIYTDHKNLEFLQSLKR---------------QNPRQ--ARW-----SLFF--------SRFNF-VLTYRPGT----------KNRKADALSR----
>Maggy
ILETDASDYV------------SAGILSQYGD--D---GILRPVAFFSKKHTA--TECNYEIYDKELLAIIRCFEEWRPEL-EG---------
SPVQIITDHRNLEYFTTTKM---------------LNRRQ--ARW-----AEFL--------SRFNF-RITYRPGK----------QGAKPDALTR----
>Skippy
ELETDASDFA------------LGGQIGQRDD--N---GVLHPIAFYSHKMHG--AELNYPIYDKEFLAIVNCFKEFRHYL-RG---------
HPVKVFTDHKNIAYFATTQE---------------LNRRQ--LRY-----AEYL--------CEFDF-TIAHCKGT----------DNGRADAISR----
>TF1
LLETDASDVA------------VGAVLSQKHD--D---DKYYPVGYYSAKMSK--AQLNYSVSDKEMLAIIKSLKHWRHYL-ES------
---EPFKILTDHRNLIGRITNES---------------ENKRL--ARW-----QLFL--------QDFNF-EINYRPGS----------ANHIADALSR----
>Pyggy
ILECDSSGYA------------TGGVLSQYDD--E---GVLRPCAYFSKKNNV--HECNYEIHDKELLAVVRCLEEWDAEL-RS---------
-SFKVITDHKNLDVLYEAKM---------------LTNVK--S----------V--------GQAAE-SIQYGAGK----------QNVRADALSR----
>Mdg3
ELHCDASSSG------------FGAVLMQKKE--D---QKWHPVSFFSKRTTD--IESKYHSFELETLAIVYSLRRFRVYL-HW---------
RTFKIVTDCNSLILTLSKKE---------------LNPRI--ARW-----ALEF--------QGYDF-EIVHRAGS----------RMQHVDALSR----
>Tor4a
VLETDASDKG------------YGGVLYICES--D-NSNCLIPVCYNSGNFSP--VQQNYTIVEKELLSGKLCMEKWAIYL-AF--------
-KKFQWITDNSNIKYVRTLRT---------------NNQKI--ARW-----LTDL--------QSFSF-TISQRPSS----------KMKISDFLSR----
>Tor2
ILLSDASELA------------AGSVLMQNID------KRQRLIAVSSKTFSE--TERKWSATERECYSLLIGCEKFEYYL-KG---------
VGFIALVDHKALLALDKRYL---------------ANSKL--KRW-----QTRL--------AEFKF-TVQYVEGR----------SHVFADMLSR----
>CaMV
IIETDASDDY------------WGGMLKAIKI-NE-GTNTELICRYASGSFKA--AEKNYHSNDKETLAVINTIKKFSIYL-TP---------
VHFLIRTDNTHFKSFVNLNYK--------------GDSKL--IRW-----QAWL--------SHYSF-DVEHIKGT----------DNHFADFLSR----
>PVCV
ILQTDASDQY------------WSAVLLEEHN------GKRKICGFASGKFKV--SEQHYHSTFKEILAVKNGIKKFNFFL-IH---------
TNFLVEMDMRAFPKMIRLNPK--------------PNSQL--LRW-----AQWF--------SPYQF-EVKHLKGK----------DNILADFLSR----
>SbCMV
IVETDASQHS------------WSGCLRALPKKQD-THNELLLCKYVSGTFTD--TETRYPIAELEVLAGVKVLEKWRIDL-LQ------
---TRFLLRTDSKYFAGFCRYNIK--------------RNGRL--IRW-----QLRL--------QAYQP-YVELIKSE----------NNPFADTLTR----
>CSVMV
IIEVDASNNA------------YGSCLKYKPK----NSKIEYLCRYNSGTFKE--NEQKYDINRKELIAVYQGLQSYSLFT-CE---------
GNKLVRTDNSQVYYWIKNDTN-----------KKSIEFRN--IKYL----LAKI--------AVYNF-EIQLIDGK----------TNIIADYLSR----
>Tor1
IVTTDASDTG------------YGGMLSQKIG--D---DPEQPLGFTSGFFRG--PSTRWAINEKELFAFIKTLEVFHHHT-YG---------
RAFTWRTDSRCLAFICAETNG-----------TKRPSAKK--LRW-----LEKL--------GEYDF-SISHVSGT----------EMAVPDGLSR---
-
>RTBV
IIETDASEEG------------WGAVLVCKPDKYS-GKDTEKIAGYASGNFG---EKKTWTSLDYEIEAINEALNKFQIYL------------
KDFTIRTDCEAIVKGIKTEDY--------------KKRSK--TRWIKLRDNLLK--------DGYKP-TFEHIKGN----------KNFLPNFLSR---
-
>ComYMV
IIETDGCMTG------------WGAVCKWKMSKHD-PRSTERICAYASGSFNP-----IKSTIDAEIQAAIHGLDKFKIYY-LD---------
KELIIRSDCEAIIKFYNKTNE--------------NKPSR--VRWLTFSDFLTG--------LGITV-TFEHIDGK----------HNGLADALSR----
>Kabuki
ALMTDASDTC------------MGAVLQQKVN------GVWQPLGYFSRKLSP--AEQKYCTYDRELLAIYKAMIYFRKLF-EG-------
--RPLTVYTDTSPCVMHSVKYVK--------------IKKRL--ARG-----------------SLGEFTTVLQSPGN----------PRPRASAKTI----
>PreDIRS
VFATDASATG------------VGAVKL--------NPKDLTELSSAYREFDE--YGGNDLEHHRELLAVQFALHHYLASK-----------
NTVVTVRTDNQNIPRILAKGSGVQELN----------------ELVLQVTEWCEQ--------RKVEL-MTTWIPRA----------
MNSAADRASR----
>CbrDIRS
FLYTDASARG------------IGGLL----------KNERQDILWKMTEIGD--NFEKQSSAWRELTAVEVASARISSNL-----------
KGRVQVLVDSQAAVSVLRRGSMKPELH----------------SLAERVWHNLE---------NVGES-SFLWIPRK----------
ENVEADEAS-----
>AliFLERV
```

```
VLFTDGSRKGPD---ETAYW---GYIL-----------TQNGKEQHRARGKV------SGSAQAGEVTAVLEGLLELTK---RK--------
IKAARVVTDSYYCAQALKEDLSIWIENGFESARGKTVAHK---ELWEKIAELR----------TQLDL-DVEHQKAHTKE--
GSYWEGNDEVDRFVQQRKI
>AciFLERV
TLFTDGSRKGPD---DTAYW---GFIL-----------KLKGKEQTRQRGKA------PGSAQAGEVTAVLEGLLELEK---RK--------
IKRAKIITDSYYCAQALKEDLTIWEENGYETAKGKKVAHE---DLWKKIAELR----------MNLDL-EVVHQKAHVKE--
GAHWRGNEEVDRFVQMRKI
>ZFERV-2
CYYTDGTTEQ-----GITRW---GWIR-----------KKGRKVTESDKSLE------GLSAQAAEVKAFRKALEKAKQ---LN--------
DLKCFIVTDSDYIYQAVEEHLGAWKNNDFNTSKGKPLKQA---DDWRVIDELIT----------EITP-TVLHQTSHTSQN-
TTAAIGNKEVDSYVRRVRA
>SFV
VFYTDGSAIKSPKT-EQTHSAGMGIVM-VVYT------PEPNITQQWSIPLG------DHTAQYAEISAVEFACKKASL---L---------
QGPVLIVTDSDYVARSANKELPFWRSNGFLNNKKKPLKHI---SKWKNISDSLL---------LKRNI-
TIVHEPGHQPSKTSIHTLGNSLADKLAVQGSY
>BFV
IFYTDGSAIRSPKP-NKTHSAGMGIIQ-AKFE------PDFRIVHLWSFPLG------DHTAQYAEIAAFEFAIRRATG---I---------
RGPVLIVTDSNYVAKSYNEELPYWESNGFVNNKKKTLKHI---SKWKAIAECKN---------LKADI-
HVIHEPGHQPAEASPHAQGNALADKQAVSGSY
>EFV
VFYTDGSSIKSPKK-DKQHSAGMGIIA-VRYQ------PQMNIIQEWSIPLG------DHTAQFAEIAAFEFALKQAIR---K---------
MGPVLIVTDSDYVAKSYNQELDFWVSNGFVNNKKKPLKHV---SKWKSIADCKK---------HKADI-
HVIHEPGHQNDLQSPYAMGNNAADKLAVKASY
>FFV
IFYTDGSAITSPTK-EGHLNAGMGIVYFINKD------GNLQKQQEWSISLG------NHTAQFAEIAAFEFALKKCLP---L---------
GGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKPLKHI---SKWKSVADLKR---------LRPDV-
VVTHEPGHQKLDSSPHAYGNNLADQLATQASF
>CoEFV
VYYSDGSAKRNQQK----NYAGIGIVK-----------THFEPEETKAVPLG------PASAQYAEVMALLDAVKQATD-------------
TGPVLICTDSVYAQRGYTEDLHYWAIRNFHDSRNAKLKYA---DKWKQLDQLKR---------DKPLV-RVIHVPGHTPGT--
VHSCGNGLADSLAQGGTY
>Loki-Xla
CVYYVDGCSFHTEEA-EKMLVAGIGIVW-----------NNIFPEISIGYKIG------PKSSQVAELAAVYKAIQMAIE---YD--------
LKEFVLITDSDYVRNSFVEYFPGWKRSNMLRSNKKPVKHG---KLFCKIDELVTT--------YGLTI-YWKKVRGHSKHP-
GADKDGNDLADSLAKKAAI
>Loki-Xtr
CVYYVDGCSFHTEEA-EKMLVAGIGIVW-----------NNIFPEISTGYKIG------SKSSQVAELAAVYKAVQMAIE---YD--------
IKEFVLITDSDYVRNSFVEYFPGWKRSNMMRSNKKPVKHG---KLFCKIDELVTT--------HGLNI-YWKKVRGHSKHP-
GADKDGNDLADSLAKKAAM
>Loki-Npa
CVYIDGCAYHTTVGDERKLVAGIGITW-----------TSEFPDVSVGYSIG------AKSSQVAELAAVYKTVQIALE---HN--------
ITEFVIVTDSNYVRNSFVEYLPTWKRNNMLRSNNKPVKHA---KLFGAIDELVKS--------NDLTI-YWKKTKGHSKTQ-
NMDKEGNDLADKLAKQGAI
>Loki-Rca
CVYVDGCAYHHKDKDQLQLVAGIGIIW-----------TQGVAPISVGYKLG------PRSSQVAELAAILKAVEMAID---YH--------
LKEFAIITDSNYVRNSFVEYLPRWKRNNMVKSNHKPVKHG---KLFCQIDQLVRE--------HEITL-YFKKTKGHSKVL-
GEEKEGNDQADTLAKEGAL
>Loki-Sme
TAYVDGCSFHH----ETQVRSGAGVLW-----------VHDTPCEPQCFQLG------SQSSQYAEIAAILITLQLAKE---ND--------
VKNLTICTDSNYARLSFTSHLPLWIKNGFMTSNRKPVKHK---ELFIACKQLVDT--------CDMQI-YWKKVRGHSRVP-
GADKDLNDQADKLAKQGAI
>Loki-Sri
TAYVDGCSFHH----ETQIRAGVGVVW-----------VHDTPYKPQMFKLG------SQSSQYAEIAGILITLQLAVE---KE--------
IKTLVICTDSNYARLSFSCHLPLWQRNGFITSNRKPVKHK---ELFVACNHLVDT--------QDMQV-YWKKVRGHSRVP-
GQDKELNDLTDTLAKRGAV
>Loki-Oja
TAYVDGCSFRK----DKELLAGVGIAW-----------LHDDPCEPKSFHLG------SQTSQYAEIAAVLIVLQLAVS---LN--------
IHTLVICTDSNYARLSFSCHLDLWKQNNFQTSNRKPVKNK---ELFVACDHLITS--------HNIQI-YWKKVRGHSRIP-
GQDKVMNDLSDSMAKQGAL
>Loki-Bpe
```

```
TVYYVDGCSFHH----EATVRAGAGTVW-----------LNNDPCEPEKFQLG------AQTSQYAEISAVLIALNQAAA---HD--------
VDTFVICTDSNYARLSFSCHLPLWKSNGFLTAAKKSVKHK---ELFIACDRLVET--------KNMKV-YWKKVRGHSREP-
GPDKELNDRADALAKQGAL
>Loki-Ame
TVYYVDGCSFRH----DQELKAGAGIVW-----------LDDNPCEPQQFKLG------SQTSQYAEIAAILITIQLAID---QG--------
VKTLVICTDSNYARLSFTCHLPIWKTKGFLTSGRKAVKHT---ELFTAADYLVVR--------HDMLV-YWKKVRGHSRVP-
GTDKTYNDQADSLAKRGAL
>Loki-Lro
TAYYVDGCSYNH----EGLLKAGAGVLW-----------LNDRPCPPQHFKLG------PQSSQYAEVAAILITLQIASA---HN--------
IRDLLICTDSNYARLSFTCHLPNWKQNGFKTANNKPVKHQ---HLFQICDDLTSK--------HDMII-YWKKVKGHSKQP-
GTDKDLNDQTDALAKAGAL
>Loki-Dcl
TAYYVDGCSYNQ----EGKLKAGAGVLW-----------LDNRPCPPQHFKLG------PQSSQYAEIAAILITLQIAST---HN--------
IRELLICTDSNYARLSFMCHLPNWKQNGFKTANKKPVKHQ---QLFQACDNITRE--------HDIMV-YWKKVKGHSKQP-
GLDKDLNDQTDALATAGAL
>Loki-Tfu
TAYYVDGCSYNH----EGNLKAGAGVLW-----------M--XPCQPQHFRLG------PQSSQYAEIAAVLIALQIASA---HN--------
IRDLLICTDSNYARLSFTCHLPSWKKNGFRTANNKQVKHK---HLFLACEDITNE--------HDMTI-YCKKVKGHSKQP-
GTNKDLNDQTDALAKAGAL
>Loki-Str
MAFVDGCSYRH----LDHLQAGVGLVW-----------HNDIPCKPLQFQLG------NKTSQFAEVAGVLITLQTAVK---HE--------
LAELAICTDSNYARLTFLCHLPFWKQKHMTTSSGKEVKNK---ELILACDDLITK--------HDIQV-YWKKVKGHSKSP-
GRDKLGNDHADSMAKSGAI
>Loki-Sau
KVYVDGCSFHH----ESQLRAGVGVVW------------DHKNDEPDHYHLG------PKLSQYAEIAAVLIVLQQAAK---LD--------
VVQLVICSDSNYARHSFISHFPTWKVNDMKNARNKPVKHA---ELFQACDLLVTD--------KGIMV-YWKKVKGHSQIP-
GPEKDGNDEADRLAKLGAE
>Loki-Ofa
RVYVDGCSFHH----ESQTRAGVGIVW-----------VNHKVDEPNHYQLG------PKTSQYAEIAAVLITLQQAAK---LA--------
IVQLVICSDSNYARHSFISHFPVWKENGMKNARNKEVKHS---ELFLACDQLVTD--------RGMTV-YWKKVKGHSQTA-
GPDKDGNDEADRLAKLGAE
>Loki-Pyo
KVYVDGCSFHH----QSHLRAGVGIAW-----------VDYNCSEPTHYQLG------PKTSQYAEIAAILIALQQAVK---LA--------
ITQFVICSDSNYARHSFISHFPLWKEKGMKNARNKEVKHS---ELFLTCDQLVTE--------KGLIV-YWKKVKGHSQIA-
GPDKDGNDEADRLAKLGAE
>Loki-Bsp
EVYVDGCSFRC----DSQLRAGVGIVW-----------ANRQFEEPNHYLLG------PKTSQYAEIAAALIAIQQAAR---QA--------
QSQLVICSDSNYARHSFISHFPTWKASGMKNARNKPVKHA---ELFLACDQLVTD--------KGMTV-YWRKVKGHSQTS-
GPDKEGNDEADRLARLGAE
>Loki-Pfl
EVYVDG-SFHH----ESQTQAGVGIYW-----------VKYRSEEPSHYQLG------PKTSQYAEIAAGFIALQQAAR---QK--------
ISQLVICSDSNYARYSFIAHFPAWKEKGMLNARNKEIKHA---KLFLACDRLVSQ--------QGMTV-YWRKVKGHSQIS-
GGDKEGNDEADRLAKLGAS
>Loki-Ape
EVYVDGCSFHH----ESQTQAGVGIYW-----------VNYRSEEPSHYQLG------PKTSQYAEIAAGFIALQQAAR---QK--------
ISQLVICSDSNYARYSFVAHFPAWKEKGMLNARNKEIKHA---KLFLACDRLVSQ--------QGMTV-YWRKVKGHSQIS-
GGDKEGNDEADRLAKLGAS
>Loki-Cne
RVYTDGCAFMH----EHQLRAGVGIVW-------------------------------------TGIDSILITIQQAAA---LN--------
QKQLVICSDSNYARLSFISHLPVWKTNGMKNARNKEVKHS---NIFLACEKLTTE--------NDMTI-YWKKVKGHSRTL-
GQDKDGNDEADRLARLGAE
>Loki-Lch
KAYVDGCSFHR----EGQLQAGIGIFW-----------EGNFPCTPLQYKLG------PKSSQFSEIAAVFVAIETAVK---NK--------
LSDIVICTDSDYARHSFCKFLPTWKANGMKNSRGKTVKHA---ELILAVDQLVLE--------NEITV-YWKKVKGHAKDS-
SQDKYGNDLADKLAKEGAI
>Loki-Pma
MVYIDGCSKMK----ENKLHAGSGIVW-------E---TGPLTGIQEGFQLG------PKSNQYAELAGVHIAVQMASN---NQ--------
IQTMVVCTDSNYVQHSFLHHLPIWKKNGMKNHRNKPIHHQ---ELFEAIDTMVQN--------KDMKI-FWKKVKGHSKIP-
GQEKTGNDQADALAKVGSL
>Loki-Lca
```

```
RVYIDGCAKMK----ENKLHAGAGIMW-------E---SGPLTGIQEGFQLG------PKSNQYAELAGVHIALQMAAD---NK--------
IQTMVVCTDSDYVQYSFLNHLPIWRKNGMKNHRNRPIHHR---ELFEAIDTMIQQ--------NDMKI-FWRKVKGHSKIP-
GQEKAGNDRADALAKLGSL
>Loki-Aro
MVYMDGSAHRK----TGTPGAGVGLLW------------TNCDLPPQKFKLG------PKSSQYAELAGTLIALXAAQN------------
YHDLVICTDSDYTVRSFLDFLPNWLANNMIKSDGKLVAHT---ELFTAVQHLLDT--------HNLTV-YWKKVPGHSKLD-
SPNKHGNDQADTLVKEGVS
>Homo_sapiens_RNase_H
VVYTDGCCSSNGRRR---PRAGIGVYW-----------GPGHPLNVGIRLPG------RQTNQRAEIHAACKAIEQAKT---QN--------
INKLVLYTDSMFTINGITNWVQGWKKNGWKTSAGKEVINK---EDFVALERLT----------QGMDI-QWMHVPGHSGFI------
GNEEADRLAREGAK
>Caenorhabditis_elegans_RNase_H
VVYTDGACSSNGTKN---AKAGWGVYW-----------GDDSEDNEFGPVYG------APTNNRGELIAVQKAIEKAIE---KR--------
LPKVVIKTDSNLLVQSMNIWIHGWKRKGWKTSTGSEVLNQ---DVLMKIDNLR----------QKLKV-KFLHVRGHAGID------
GNEKADELARKGAQ
>Trypanosoma_brucei_RNase_H
VVYVDGACRNNGSRSERRPRAGFGGFY-----------GDGDSRNFKFPLPA------PQTNQRAELSALIHVLRVALD---SH--------
CYNLCVYSDSKYTVMGVNSYLHRWERNGFKTAGGGDVANI---DLWSQFTKLRNRHLS--RCAERVAL-
QLKHVPGHAGVY------GNEMADRLAVEACE
>Candida_orthopsilosis_RNase_H
RIYVDGASRGNGQS--KSAPSGYGVYY-----------GPGDSRNAAVGLHE------TPTNQRAELHGMKHALTNIKDDSSTS-------
-SPRYDIYSDSKYAIQSINEWSDKWQSNGWKTSTGQPVANS---DLIKDAVQLKNEINQLYRERDWKPL-
EFHHVKGHSGDV------GNEAADRLANQGAD
>Komagataella_phaffii_RNase_H
IIYTDGATRGNGKFG---AKGGYGVFF-----------KENDLRNISKPLTG------VQTNQRAELSAVRDALIIIEKDLASG--------
KKHYTIRTDSRYTINSITKWVTKWKQNDWRTATG-PVQNR---DLIEPCTNLIESVNSKYKELGMPPL-NIEYVKGHSGDY-
-----GNEMADKLATAGAD
>Helicobacter_pylori_RNase_H
EIFCDGSSLGNP------GPGGYAAIL-----------RYKDKEKTISG------GENFTTNNRMELRALNEALKVLKR------------
PCHITLYSDSQYVCQAINVWLTNWQKKNF-----SKVKNV---DLWKEFLEVS----------KGHSI-MAVWIKGHNGHA------
ENERCDSLAKLEAQ
>Buchnera_aphidicola_RNase_H
KIFSDGSCLGNP------GPGGYSFII-----------QHLEYENISSS------GFYLTTNNRMELMGIIVATESLKQ------------
PCCITISTDSQYVQKGILYWIKNWKTKGWKTSRKTYVKNV---DLWLRLEKSL----------NLHQV-TWKWIKSHSGNK------
KNEQCDHLARESAK
>Shewanella_denitrificans_RNase_H
NIYTDGSCLGNP------GPGGYGIVM-----------QYKQHSKEIAD------GFALTTNNRMELLAPIIALEALME------------
PCIVTLTSDSQYMRQGITQWIHGWKKKGWMTSNKQAVKNV---DLWKRLDSVS----------QRHNI-DWRWVKGHTGHK---
---QNERCDKLARDAAE
>KoRV
AWYTDGSSFIM----DGRRQAGAAIVD------------NKRTVWASNLPE------GTSAQKAELIALTQALRLAE-------------
GKSINIYTDSRYAFATAHVHGAIYKQRGLLTSAGKDIKNK---EEILALLEAIH---------LPKRV-AIIHCPGHQRGT-
DPVATGNRKADEAAKQAAQ
>MuLV
TWYTDGSSFLQ----EGQRKAGAAVTT------------ETEVVWAKALPA------GTSAQRAELIALTQALKMAE-------------
GKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNK---DEILALLKALF---------LPKRL-SIIHCPGHQKGN-
RAEARGNRMADQAAREVAT
>FLV
TWYTDGSSFIR----NGEREAGAAVTT------------ESEVIWAAPLPP------GTSAQRAELIALTQALKMAE-------------
GKKLTVYTDSRYAFATTHVHGEIYRRRGLLTSEGKEIKNK---NEILALLEALF---------LPKRL-SIIHCPGHQKGD-
SPQAKGNRLADDTAKKAAT
>REV
TLFTDGSSYVR----DGKRYAGAAVVT------------LDSVIWAEPLPI------GTSAQKAELIALTKALEWSK-------------
DKSVNIYTDSRYAFATLHVHGMIYRERGLLTAGGKAIKNA---PEILALLTAVW---------LPKRV-AVMHCKGHQKDD-
APTSTGNRRADEVAREVAI
>WDSV
TLFSDGSYTT--------GRGGAAVVV------------TDDFIIIHQQPG------GASAQTAELLALAAACHLAT-------------
DKTVNIYTDSRYAYGVVHDFGHLWMHRGFVTSAGTPIKNH---KEIEYLLKQIM---------KPKQV-SVIKIEAHTKGV-
SMEVRGNAAADEAAKNAVF
>Xen-1
```

```
TLFVDGSRFSD----SGKFHTGYAVTT-------------TDSVLEAQPLPA------SCSAQEAELKALTAACKLAA-------------
GKRANIFSDSRYAQGVALDFGTIWKTRGYLTATGSPIKNG---RSVADLMEALT---------LPEQV-AVLKVKAHGRLT-
SPEAIGNHLADTTAKEIAV
>SnRV
KWWIDGSRYWD-----GGYVTGWAALE-------------DKKNQLGGALEG------HVSAQVAELVALREALRL---------------
QRPLTLYTDSTYVLGICTKYLAVWKRRGMVNADGSQISNQ---NILQEIWQLIE---------STQTL-GIVKVKAHTQRK-
STHEQQNNDVDQPAKQYAK
>HERV-L
AWFTDGSADMQ----ATTWQKWTAAAL--------------QPLSRTSLKDS------GESSQWAELAVHLVVLHFAWK-------------
KWPVYLLTDSWAVANGLAGWSGTWKKHDWKIGDDKEIWGRG--GMWMDLS-------------KTVKI-FVSHVSAHQSVT-
SAEEDFNNQVDRMTHSVDT
>MuERV-L
AWFTDGSARYA----GT-TQKWTAAAL--------------QPLSGTTLKDT------RKSSQWAELRAVHMVLQFVCK-------------
KWPVRLFTDSWAVANGLAGWSGTWKDHNWKIG-EKDIWGR---SMWIDISKWA----------KDVKI-FVSHVNAHQKVT-
SAEEEFNNQVDKMTRSVDS
>JSRV
LIFTDGSSNGT-----------AALII------------NHQTYYAQTS---------FSSAQVVELFAVHQALLTV---------------
PTSFNLFTDSSYVVGALQMIETVP-------------------NLFTLIQQVL----------RQHPC-FFGHIRAHSTLP-
GALVQGNHTADVLTKQVFF
>BLV
CLFSDGATGR------------GAYCLW------------KDHLLDFQAVPA------PESAQKGELAGLLAGLAAAP--------------
PEPLNIWVDSKYLYSLLRTLVLGA--------------------YALLYKSLL----------RHPAI-FVGHVRSHSSAS-
HPIASLNNYVDQLLPLETP
>HTLV-2
CLFSDGSPQK------------AAYVLW------------DQTILQQDI---------HESAQKGELLALICGLRAAK--------------
PWPLNIFLDSKYLIKYLHSLAIGA-------------------TLQAALPPLL----------QGKTI-YLHHVRSHTNLP-
DPISTFNEYTDSLILAPLV
>SRV-1
LVFTDGSSTGM-----------AAYTL------------ADTTIKFQTN---------LNSAQLVELQALIAVLSAFP--------------
NQPLNIYTDSAYLAHSIPLLETVA-------------------KLFLQCQQLIYN--------RSIPF-YIGHVRAHSGLP-
GPIAHGNQKADLATKTVAS
>MMTV
VIFTDGSANGR-----------VTYIQ------------GREPIIKENT---------QNTAQQAEIVAVITAFEEV---------------
SQPFNLYTDSKYVTGLFPEIETAT-------------------TELKHLQRLIHK--------RQEKF-YIGHIRGHTGLP-
GPLAQGNAYADSLTRILTA
>RSV
TVFTDASSSTH-----------KGVVVWR-------EGPRWEIKEIADL---------GASVQQLEARAVAMALLLW---------------
TTPTNVVTDSAFVAKMLLKMGQEG-------------------AAAFILEDALSQ--------RSAMA-AVLHVRSHSEVP-
GFFTEGNDVADSQATFQAY
>LDV
TVFTDASSQTA-----------MAVAVWQ-------ENGRWKKVTECSP---------GSSVQQLEARAVQLASDAF---------------
TEALNVVTDSLYVARLVSRMATPG-------------------LASPWLYNSLSR--------RTAPL-FICHVQAHQSLE-GLHT-
GNDAADRAAK-GLW
>HIV-1
TFYVDGAANRE-------KLGKAGYVT--------------NRGRQKVVTLT------DTTNQKTELQAIYLALQDS--------------
GLEVNIVTDSQYALGIIQAQPDQS-------------------ELVNQIIEQLI---------KKEKV-YLAWVPAHKGIG------
GNEQVDKLVSAGIR
>FIV
TWYIDGGRKL--------GKAAKAAYW-------------TDTGKWQVM---------EGSNQKAEIQALLLALKAG--------------
SEEMNIITDSQYVINIILQQPDMM-------------------GIWQEVLEELE---------KKTAI-FIDWVPGHKGIP------
GDGILDKRSEDAGY
>EIAV
TIYTDGGKQN--------GEGIAAYVT--------------SNGRTKQKRLG------PVTHQVAERMAIQMALEDTR-------------
DKQVNIVTDSYYCWKNITEGLGLE-------------------SPWWPIIQNIR---------EKEIV-YFAWVPGHKGIY------
GNQLADEAAKDAGF
>Lian-Aa1
VCYTDGSLLE--------GRAGAGVYS--------------RELRLNQFSLG------NCTVFQAEIFALMCGVQSALQ---QR--------
GKVIYFCSDSQAAIKALASANSRS-------------------KLVIACRTQIEELN------SVNSV-NLVWVPGHSSIA------
GNELADELARDGAS
>TRAS1
```

```
RIFTDGSKIE--------GRVGAALSIW----------DGEVEIRSLKLALA------YCTVYQAELLALSYAVKEAQL---RN--------
GSTFGVFSDSKAALLTVINHGSLH-------------------PLAVDIRKMLKQCAL-----QNKTV-ALYWIKAHAGLE------
GNERADQLAKEAA-
>CATS
-IFTDGSKIE--------GRVGAALSIW----------EGTGEIKTKKLKLG------YCTVYQAELLALLKATEEVLS---GG--------
AATYNIFCDARSTLDVIASGESLH-------------------PLAFKITKNLKTITE-----RNQEI-RLFWIKAHIGLE------
GNERADVLAKEAAL
>You
-IYTDGSKVT--------GATTFAVVD------------SNRKIIAGGRLPS------YNSIFTAEAFAILKACQFASK-------------
AGKSVICTDSLSSLSAIRNWNHND-------------------PTTQEVRHILSS--------HPKKI-TLLWVPSHQGIH------
GNELADKAAQEMRL
>I
-IFTDGSKIN--------YTISFAITT-------------ETDVLKYGILPP------YSSVLTSETIAILEAIELTKN-------------
RGKFIICSDSLSAVDSIQNTNNNS-------------------FYPSRIRSLITQ--------HAPKI-KIMWIPGHSGIK------
GNELADQAAKSASS
>CgT1
-VYSDGSKAPN-------GATGFGFVIY----------RGSRRIAQGCGRLG------IAEVFDGEAEGARAGLRRALL---TS--------
GQPIHICIDNTSVIQGIRSNIPDSSQA---------------------IQAVA----------RIYNI-QTHWSPGHQGIK------GNEEADILAKEGTT
>MGR583
VVYTDGSEKHEN------VQIGYGWAAF----------RAGLEFAARAASIT------ESHVFDAEAIGALKGLQAAAK---AQ--------
GARIWICVDSTSVIWGLRGDAPRSSQW----------------LEFHNLVDLLRK--------QSTEV-RVRWCPGHQGIP------
GNDRADELAKAGSA
>Pisum_sativum_RNase_H
TLMFDGAVNMNG--------NGVGAVL--------INPK-GAHMPFSARLTF------DVTNNEAEYEACIMGIEEAID---LR--------
IKTLDIFGDSALVVNQVNGDWNTNQPH--------------LIPYRDYTRRILT---------FFKKV-KLYHVPRD----------
ENQMADALATL---
>Arabidopsis_thaliana_RNase_H
TIEFDGASKGNP------GKAGAGAVL--------RASD-NSVLFYLREGVG------NATNNVAEYRALLLGLRSALD---KG--------
FKNVHVLGDSMLVCMQVQGAWKTNHPK--------------MAELCKQAKELMN---------SFKTF-DIKHIARE----------
KNSEADKQANSAI-
>Oryza_sativa_RNase_H
-LHFDGASKGNP------GKAGAGAVL--------MTED-GRVISRLREGLG------IVTNNVAEYRGLILGLRYAIR---HG--------
FKKIIVYGDSQLVCYQVKGTWQTKNQN--------------MMELCKEVRKLKE---------NFVSF-EINHIRRE----------
WNAEADRQANIAI-
>Ostreococcus_tauri_RNase_H
VLEFDGASRGNP------GEAGAGALL--------RRKRDDRVVEELLEYLG------ERTVNEAEYAALCLGLRKAIE---LG--------
ITKIEVRGDSKLIVNQVDGSFKLKSEN--------------LRSMHAEAVSLKK---------KFAEF-KISHVKRE----------
FNKHADHLANMAVD
>Haloarcula_marismortui_RNase_H
--YFDGASRGNP------GPASVGYVL--------VNDS--GIVTEGGETIG------TATNNQAEYKALIRAVEVARD---YG--------
FDDVHIRGDSELIVKQVRGEWDTNDPE--------------LREHRVRVRELLT---------DFDDW-QIEHVPRE----------
INDRADELANDALD
>Natrinema_pellirubrum_RNase_H
-VYFDGGSRGNP------GPAAIGWVI--------VTGD--GIVAEDGETIG------TATNNQAEYEALIAGLEAARD---YG--------
YDEVHVRGDSELIVKQVRGEYDTNDPD--------------LRENRVTVHELLR---------AFDEW-TLEYVPRE----------
VNDRADGLVNEALD
>Halorhabdus_utahensis_RNase_H
-VYVDGASRGNP------GPAAIGWVI--------LTGD-GGIVTEGGKRIG------STTNNRAEYEALIHALEIAAD---YG--------
FDSVEVRSDSELAVRQVRGEWDTNDPD--------------LRERRVRVRELFR---------EFDDW-SIEHVPRE----------
INERADALANEA--
>Halorubrum_lacusprofundi_RNase_H
-VYFDGASRGNP------GPGAVGWCL--------VTAD--GIVAEGGERIG------RVTNNQAEYAALIRALEAADE---YD--------
FDEIDVRGDSQLIVKQVRGEWNANDPE--------------LRERRVRVRELLE---------RFDRW-SIGHVPRE----------
INERADDLANEALD
>Mycolicibacterium_gilvum_RNase_H
IVECDGGSRGNP------GPAGYGAVV--------WSADRESVLAEARSAIG------TATNNVAEYRGLIAGLESAAQ---TG--------
ATEVEVRMDSKLIVEQMAGRWRVKHPD--------------LAALHRQAADLAR---------RFEHI-TYSWIPRE----------
QNSYADRLANEAMD
>Mycobacteriaceae_RNase_H
LVEADGGSRGNP------GPAGYGSVV--------WSEDRSSVLAEAKQSIG------RATNNVAEYRGLIAGLEEAAN---VG--------
ATEVAVSMDSKLIVEQMAGRWRVKHPD--------------LIPLYQRARELAE---------GFDRV-TYSWIPRA----------
ANARADRLANEAMD
```

```
>Micromonospora_aurantiaca_RNase_H
VIEADGGSRGNP------GPAGYGAVV--------RDPETGEVLAERSESIG------TATNNVAEYRGLIAGLEAAAE---LG--------
AAEVEARMDSKLVVEQMCGRWQIKHPG--------------LRPLAAQAAGLVS---------RFAAV-RFTWVPRE----------
RNRHADALANAAMD
>Ignicoccus_hospitalis_RNase_H
ELYFDGLCEPVNPG----GVATYGFVV----------KEGGKVLCSGKGLVG------DVTNNVAEYTALIKALECLLE---KG--------
GAEVVVKGDSQLAIRQLRGEYKVRSPR--------------IAPLYKRAKELLS---------KFKA--ELQWVPRE----------
LNEEADALSREAFR
>Vulcanisaeta_moutnovskia_RNase_H
-VYFDGACEPKNPG----GVGTYGFAV--------YNDN-DTISEGYGIACE------NCTNNVAEYTGLIKALECLLL---YG--------
ESSIVVHGDSQLVIKQLTGIYNVRTEH--------------LKPLFEKAHELLS---------HF-RV-RLEWIPRE----------
LNSKADELSKRAY-
>Ferroglobus_placidus_RNase_H
VLYFDGACLPVNPG----GIATYGFVI--------ITPS-GEIVKEKGIAAE------KGTNNIAEYTALIRGLEKALE---LG--------
IDELIVRGDSQLAIYQMNGVYAVKSPN--------------IIPLWQRAMELAG---------KFRKI-RFEWVPRE----------
QNSAADELSTQAY-
>Mars1
VFYSDGSKLD--------NKIGAGVYL---------RPSRAIPAESHSFSLG------TSEVFDAELQACYYACVRAKA---RN-------
TITDCWVFLDNTSVIHRLRHTSPTAG----------------------------------------------------------------------------
>Bel
HGFADASSHA------------YGAVVYARVA-VG-CSFQVTLVAAKTR----VAPIKPVSIPRLELNAALLLSRLLSIVK-TSLTI----
PLFSTSCWTDSEIVLHWLSAPPRRWNTY--VCNRTSEILSDF----------------------PRS---CWNHVRTEDNP-----------
ADCASR----
>Tamy
HTFVDASETA------------YAAACYWRAE---TEVIHVALIAGKAR----VSPAKPVTIPRLELQAALLGARLARTVT-EEIDL----
QVTDRYFWSDSSTVLQWLKADPRRFKPF--VAHRLAEIKDLS----------------------KPH---EWRWVPTRDNP-----------
ADIATR----
>Zebel
HIFCDASERA------------YGSVAYLRAS-DDQGHVAVTFVLARSR----VAPRKCLSMPRLELSAALTGAQLAQVIQ-NELTL-
---PIDSVTLWSDSTTVLYWLTAESCRYKVF--VGGRVAEIQTLT----------------------ETA---EWRYVDSANNP-----------
ADHITR----
>Purbel
CIFCDASEIA------------YGACAYLRWK-TDDDKYEVRFVSAKSK----VAPLKALTIPRLELLAGVLAARMHEAIS-NEMRL-
---QVEKVVFFTDSMIVLQWIKSSARTYKAF--VSSRVGEIQTLT----------------------NPA---DWKHIPGEVNI-----------
ADKVSR----
>Kobel
HAFADASEIG------------YGCVIYLRQI-NSSENIHCSFVFAKAR----VTPLKKITIPRLELTAATLAVRLVSIVQ-RELDF----
KIDKAIYWTDSTAVLRYIRNDRARYHTF--VANRVQVIREAT----------------------VPE---QWHHVDTKRNP-----------
ADLASR----
>Cer10-1
IMFSDASTDI------------YATAVYVQYS-YKSRKPVTRLLTSKSK------KNEMWTIPKLELLGIECSSNLARTII-AELGTL---
KIKKVRLFTDRACALYWILSHKTTRV---WVANRVSQIQANQ----------------------CGIET-SLHHCPTKENP-----------
ADLATR----
>Koala
SGYSDADWAGNGDD----RRSTGGFAVFL----------GNSLVSWSAR----KQPTVSRSSTESEYKAMANATAEIMWIQ-
TLLKEIHVSNSPTARLWCDNLSAKYLSSNPIFHARTK-HIEVDYHFVRERV----------------------QQKLL-EVGFVPTGDQI----
-------ADGFTK----
>SIRE1-4
VGYCDADWAGSADD----RKSTSGGCFYL----------GNNLISWFSK----KQNCVSLSTAEAEYIAAGSSCSQLVWMK-
QMLKEYNV-EQDVMTLYCDNMSAINISKNPVQHSRTK-HIDIRHHYIRDLV----------------------DDKVI-TLKHVDTEEQI---
--------ADIFTK----
>Tnt-1
KGYTDADMAGDIDN----RKSSTGYLFTF----------SGGAISWQSK----LQKCVALSTTEAEYIAATETGKEMIWLK-
RFLQELGL-HQKEYVVYCDSQSAIDLSKNSMYHARTK-HIDVRYHWIREMV----------------------DDESL-KVLKISTNENP--
---------ADMLTK----
>Vitico1-1
VGYTDSDWAGCLED----RKNTSSYMFSL----------G------------------------------------SVWLR-
RILADISQEHEESTIIYCDNKAAIAMTKNPAYHGRTK-HVDIRVHFIRDLV----------------------VEGKV-VLQYCNTNEQV---
--------ADVLTK----
>Copia
```

```
IGYVDSDWAGSEID----RKSTTGYLFKM----------DFNLICWNTK----RQNSVAASSTEAEYMALFEAVREALWLK-
FLLTSINIKLENPIKIYEDNQGCISIANNPSCHKRAK-HIDIKYHFAREQV---------------------QNNVI-CLEYIPTENQL---------
--ADIFTK----
>Hydra1-2
VAYCDADFASDLND----RRSTTGYCFSL----------NGPLISWKSK----KQSTVALSTCEAEYMALTVTTQESMYLV-
QLLDSMNDCMYMPVQIFEDNQGAIALSKNPVCRQRCK-HVDVRYHFIRSAL---------------------SDGKV-
TVEYCPTEDMV-----------ADVMTK----
>pCretro6
VGYSDADWGANPDD----QKSISGYVFLL----------GGAPVCWASR----KQKSVALSSMEAEYMAGSTAASQALWCR-
MLLEELGFAQPNPTLLYMDNQSALALARNTGTQGRAK-HIDIRYHFLRDKI---------------------SSKEI-SVAHCPGEDNP---
--------ADIFTK----
>1731
TGFVDADWGGDRLD----RKSYTGYVFFL----------SGGPVSWRSE----KQQSVALSSTEAEYMALTTACKEAIALR-
RLIVEICGDLKTPTVMHGDNKCAAQLAKNPVHHSRTK-HIDIRYH--REVM---------------------KEGHV-VLEYTSTNEMI---
--------ADINTK----
>CoDi6.3
VCYSDSDYAGDPDT----RRSVSGYILYV----------KGVPICWRSK----AQRSITLSSSEAEWIALSEATKEIMFVL-
QLLESLYIKVQLPITVRVDNIGAIWMSQNVNTSSRTK-HVDIRTKYVNEYC---------------------EDGVL-KIIFVKSADND----
-------SDIMTK----
>CoDi6.7
EVYADADFSGDPDT----ARSRTGFAIFY----------AGCPVTWQSK----LQTEISLSTTESEFVSLSTALRTALPLI-
GLAKEMDISVTVPTVAFDDNMGAIEIALVPKMRPRTK-HINVKYHHFRQHV---------------------DNGDI-TIQHVDSEDQI--
---------ADFLTK----
>Ty1B
VVISDASYGNQPY-----YKSQIGNIYLL----------NGKVIGGKST----KASLTCTSTTEAEIHAISESVPLLNNLS-YLIQEL---
DKKPITLLTDSKSTISIIISNNEEKFRNRFFGTKAMRLRDEV---------------------SGNHL-HVCYIETKKNI-----------
ADVMTK----
>Ty4
IAITDASVGSEYD-----AQSRIGVILWY----------GMNIFNVYSN----KSTNRCVSSTEAELHAIYEGYRDSETLK-
VTLKELGEGDNNDIVMITDSKPAIQGLNRSYQQPKEK-FTWIKTEIIKEKL---------------------KRS---ITVKITGKGNI---------
--ADLLTN----
>Zeco1
VVFSDASLGNLPD-----GGTQGGTLIGL----------KFSPLCWQSK----KIRRVVRSTLAGETLALSDGIDNAIFLT-
TLFSELAELNTPPLICVTDNHSLFDALKSTKQVSEKR--LRLEISSIKDLM---------------------QSKKIKKVLWSDTKTQL------
-----ADCLTK----
```

Dataset S6. RH alignment generated by PROMALS3D and used to reconstruct the phylogenetic tree in Fig. 3.

```
>PreDIRS
VFATDASATG-----------VGAVKLNPK------LTELS-SAYREFDEYGGNDLEHHRELLAVQFALHHYLAS---------
KNTVVTVRTDNQNIPRIL-AKGS----------------ELVLQVTEW-CEQ-----------RKVELMTTWIP-RA---------
MNSAADRASRETD-
>CbrDIRS
FLYTDASARG-----------IGGLLKNER-----QDILWK-MTEIGDGN-FEKQSSAWRELTAVEVASARISSN---------
LKGRVQVLVDSQAAVSVL-RRGS-----------------LHSLAERV-WHNL---------ENVGESSFLWIP-RK----------ENVEADEAS---
--
>Mars1
VFYSDGSKLDNKIG-------AGVYLRPSR-----AIPAES-HSFSL----GSTSEVFDAELQACYYACVRAKAL---------
TITDCWVFLDNTSVIHRL-RHTSPTAGQ-----------ALAIGVHEV--RTP---------LTTWE---------------------------------
>RSV
TVFTDASSSTH----------KGVVVWREG-----RWEIKE----IA----DLGASVQQLEARAVAMALLLWP------------
TTPTNVVTDSAFVAKML-LKMGT------------------AAAFILEDALS---------QRSAMAAVLHVR-SHSEVPG-
FFTEGNDVADSQATFQAY
>LDV
TVFTDASSQTA----------MAVAVWQEN-----RWKKVT----EC----SPGSSVQQLEARAVQLASDAFP------------
TEALNVVTDSLYVARLV-SRMAD------------------LASPWLYNSLS---------RRTAPLFICHVQ-AHQSLE--
GLHTGNDAADRAAKGLW-
>BLV
CLFSDGATR------------GAYCLWKDH-----L-LDFQ----AV----PAPESAQKGELAGLLAGLAAAP------------
PEPLNIWVDSKYLYSLL-RTLVLGAW-----------LPDPVPSYALLYKSL---------LRHPAIFVGHVR-SHSSASH-
PIASLNNYVDQLLPLETP
>HTLV-2
CLFSDGSP-------------AAYVLWDQT-----LQQDIT----PL----PSHESAQKGELLALICGLRAAK------------
PWPLNIFLDSKYLIKYL-HSLAIGAF-----------LTSAHQTLQAALPPL---------LQGKTIYLHHVR-SHTNLPD-
PISTFNEYTDSLILAPLV
>JSRV
LIFTDGSSNG-----------TAALIINHQ-----TYYAQ-----------TSFSSAQVVELFAVHQALLTV-------------
PTSFNLFTDSSYVVGAL-QMIETVPI-----------IEVLNLFTLIQQVLH---------CRQHPCFFGHIR-AHSTLPG-
ALVQGNHTADVLTKQVFF
>SRV-1
LVFTDGSSTG-----------MAAYTLADT-----TIKFQ-----------TNLNSAQLVELIAVLSAFPNQ---------------
PLNIYTDSAYLAHSI-PLLETVAQ-----------IETAKLFLQCQQLIY---------NRSIPFYIGHVR-AHSGLPG-
PIAHGNQKADLATKTVAS
>MMTV
VIFTDGSANG-----------SVTYIQGRE-----PIIKE-----------NTQNTAQQAEIVAVITAFEEV-------------
SQPFNLYTDSKYVTGLF-PEIETAKI-----------YTELK--HLQ-RLIH---------KRQEKFYIGHIR-GHTGLPG-
PLAQGNAYADSLTRILTA
>FLV
TWYTDGSSFIRNGERE-----AGAAVTTES-----EVIWAA----PL----PPGTSAQRAELIALTQALKMAE------------
GKKLTVVYTDSRYAFATT-HVHGEIYRRRGLLTSEGKEIKNKNEILAL-LEAL---------FLPKRLSIIHCP-GHQKGDS-
PQAKGNRLADDTAKKAAT
>KoRV
AWYTDGSSFIMDGRRQ-----AGAAIVDNK-----RTVWAS----NL----PEGTSAQKAELIALTQALRLAE------------
GKSINIYTDSRYAFATA-HVHGAIYKQRGLLTSAGKDIKNKEEILAL-LEAI---------HLPKRVAIIHCP-GHQRGTD-
PVATGNRKADEAAKQAAQ
>MuLV
TWYTDGSSFLQEGQRK-----AGAAVTTET-----EVVWAK----AL----PAGTSAQRAELIALTQALKMAE------------
GKKLNVYTDSRYAFATA-HIHGEIYRRRGLLTSEGKEIKNKDEILAL-LKAL---------FLPKRLSIIHCP-GHQKGNR-
AEARGNRMADQAAREVAT
>REV
TLFTDGSSYVRDGKRY-----AGAAVVTLD-----SVIWAE----PL----PIGTSAQKAELIALTKALEWSK------------
DKSVNIYTDSRYAFATL-HVHGMIYRERGLLTAGGKAIKNAPEILAL-LTAV---------WLPKRVAVMHCK-GHQKDDA-
PTSTGNRRADEVAREVAI
>Xen-1
TLFVDGSRFSDASGKFH----TGYAVTTTD-----SVLEAQ----PL----PASCSAQEAELKALTAACKLAA------------
GKRANIFSDSRYAQGVA-LDFGTIWKTRGYLTATGSPIKNGRSVADL-MEAL---------TLPEQVAVLKVK-AHGRLTS-
PEAIGNHLADTTAKEIAV
```

```
>WDSV
TLFSDGSYTTGRGGA---------AVVVTD-----DFIIIH----QQ----PGGASAQTAELLALAAACHLAT------------
DKTVNIYTDSRYAYGVV-HDFGHLWMHRGFVTSAGTPIKNHKEIEYL-LKQI---------MKPKQVSVIKIE-AHTKGVS-
MEVRGNAAADEAAKNAVF
>SnRV
KWWIDGSRYVTG---------WAALEDKKN-------QLGG----AL----EGHVSAQVAELVALREALRLQ--------------
RPLTLYTDSTYVLGIC-TKYLAVWKRRGMVNADGSQISNQNILQEI-WQLI---------DSTQTLGIVKVK-
AHTQRKCSTHEQQNNDVDQPAKQYAK
>EIAV
TIYTDGGKQNGEG--------IAAYVTSNG-------RTKQ----KR----LGPVTHQVAERMAIQMALEDTR------------
DKQVNIVTDSYYCWKNI-TEGLGLEGPQ-----------SPWWPII-QNIR----------EKEIVYFAWVP-GHKGI------
YGNQLADEAAKDAGF
>HIV-1
TFYVDGAANRETKLGK-----AGYVTNRGR--------QKV----VT----LTDTTNQKTELQAIYLALQDS-------------
GLEVNIVTDSQYALGII-QAQPD--------------SESEVNQII-EQLI----------KKEKVYLAWVP-AHKGI------GGNEQ---------
>FIV
TWYIDG-GRKLGKAAK-----AAYWTDTGK--------WQV----ME-----LEGSNQKAEIQALLLALKAG-------------
SEEMNIITDSQYVINII-LQQPD--------------MEGIWQEVL-EELE----------KKTAIFIDWVP-GHKGI------PGNEE---------
>Lian-Aa1
VCYTDGSLLEGRAG-------AGVYSRERL-------NQFY----SL----GRNCTVFQAEIFALMCGVQSALQQ----------
MGKVIYFCSDSQAAIKAL-ASANSRL--------------VIACRTQ-IEEL---------NSVNSVNLVWVP-GHSSI------
AGNELADELARDGAS
>TRAS1
RIFTDGSKIEGRVGAA-----LSIWDGEVE-----IRSLKL----AL----APYCTVYQAELLALSYAVKEAQLR----------
NGSTFGVFSDSKAALLTV-INHGS--------------LAVDIRKML-KQCA---------LQNKTVALYWIK-AHAGL------
EGNERADQLAKEAA-
>CATS
-IFTDGSKIEGRVGAA-----LSIWEGTGE-----IKTKKL----KL----GSYCTVYQAELLALLKATEEVLSG----------
GAATYNIFCDARSTLDVI-ASGES--------------LAFKITKNL-KTIT---------ERNQEIRLFWIK-AHIGL------
EGNERADVLAKEAAL
>Caenorhabditis_elegans_RNase_H
VVYTDGACSSNGTKNAKA--GWGVYWGDDS-----EDNEFG----PV----YGAPTNNRGELIAVQKAIEKAIEK----------
RLPKVVIKTDSNLLVQSM-NIWIHGWKRKGWKTSTGSEVLNQDVLMKI-DNLR----------QKLKVKFLHVR-GHAGI------
DGNEKADELARKGAQ
>Candida_orthopsilosis_RNase_H
RIYVDGASRGNGQSKSAPS-GYGVYYGPGD-----SRNAAV-GLHEVDNINKYTPTNQRAELHGMKHALTNIKDD----------
SSPRYDIYSDSKYAIQSI-NEWSDKWQSNGWKTSTGQPVANSDLIKDA-VQLKNEINQLYRERDWKPLEFHHVK-
GHSGD------VGNEAADRLANQGAD
>Komagataella_phaffii_RNase_H
IIYTDGATRGNGKFGAKG--GYGVFFKEND-----LRNISK----PL----TGVQTNQRAELSAVRDALIIIEKD----------
SKKHYTIRTDSRYTINSI-TKWVTKWKQNDWRTAT-GPVQNRDLIEPC-TNLIESVNSKYKELGMPPLNIEYVK-GHSGD---
---YGNEMADKLATAGAD
>MGR583
VVYTDGSEKHENNCVQ---IGYGWAAFRAG-----LEFAAR---AASI---TPESHVFDAEAIGALKGLQAAAKA----------
PGARIWICVDSTSVIWGL-RGDAPRSS----------QWAFLEFHNLV-DLLR---------KQSTEVRVRWCP-GHQGI------
PGNDRADELAKAGSA
>Homo_sapiens_RNase_H
VVYTDGCCSSNGRRRP----RAGIGVYWGG----HPLNVGI----RL----PGRQTNQRAEIHAACKAIEQAKTQ----------
NINKLVLYTDSMFTINGI-TNWVQGWKKNGWKTSAGKEVINKEDFVAL-ERLT----------QGMDIQWMHVP-GHSGF------
IGNEEADRLAREGAK
>CgT1
-VYSDGSKAPNGATG------FGFVIYRGS-----RIAQGC----GR----LGIAEVFDGEAEGARAGLRRALLT----------
QGQPIHICIDNTSVIQGI-RSNIPDS--------------SQAAFLE-IQAV---------ARIYNIQTHWSP-GHQGI------
KGNEEADILAKEGTT
>You
-IYTDGSKVTGATT-------FAVVDSNKI-------IAGG----RL----PSYNSIFTAEAFAILKACQFASKN-----------
AGKSVICTDSLSSLSAI-RNWNHND--------------PTTQEVR-HILS---------SHPKKITLLWVP-SHQGI------
HGNELADKAAQEMRL
>I
-IFTDGSKINYTIS-------FAITTETVL--------KYG----IL----PPYSSVLTSETIAILEAIELTKNR-----------
RGKFIICSDSLSAVDSI-QNTNNNS--------------FYPSRIR-SLIT---------QHAPKIKIMWIP-GHSGI------
KGNELADQAAKSASS
```

```
>Ostreococcus_tauri_RNase_H
VLEFDGASRGNPGEAG-----AGALLRRKR-----DRVVEE-LLEYL----GSERTVNEAEYAALCLGLRKAIEL----------
GITKIEVRGDSKLIVNQV-DGSFKLKS-----------ENLRSMHAEA-VSLK---------KKFAEFKISHVK-RE----------
FNKHADHLANMAVD
>Helicobacter_pylori_RNase_H
EIFCDGSSLGNPGPGG-----YAAILRYKD-----KEKTIS-GGE-------NFTTNNRMELRALNEALKVLKR-----------
PCHITLYSDSQYVCQAI-NVWLTNWQKKNFS-----KVKNVDLWKEF-LEVS----------KGHSIMAVWIK-GHNGH------
AENERCDSLAKLEAQ
>Buchnera_aphidicola_RNase_H
KIFSDGSCLGNPGPGG-----YSFIIQHLE-----YENISS-SGF-------YLTTNNRMELMGIIVATESLKQ-----------
PCCITISTDSQYVQKGI-LYWIKNWKTKGWKTSRKTYVKNVDLWLRL-EKSL----------NLHQVTWKWIK-SHSGN------
KKNEQCDHLARESAK
>Shewanella_denitrificans_RNase_H
NIYTDGSCLGNPGPGG-----YGIVMQYKQ-----HSKEIA-DGF-------ALTTNNRMELLAPIIALEALME-----------
PCIVTLTSDSQYMRQGI-TQWIHGWKKKGWMTSNKQAVKNVDLWKRL-DSVS----------QRHNIDWRWVK-GHTGH------
KQNERCDKLARDAAE
>Arabidopsis_thaliana_RNase_H
TIEFDGASKGNPGKAG-----AGAVLRASD-----NSVLFY--LREG----VGNATNNVAEYRALLLGLRSALDK----------
GFKNVHVLGDSMLVCMQV-QGAWKTNH-----------PKMAELCKQA-KELM---------NSFKTFDIKHIA-RE----------
KNSEADKQANSAI-
>Oryza_sativa_RNase_H
-LHFDGASKGNPGKAG-----AGAVLMTED-----GRVISR--LREG----LGIVTNNVAEYRGLILGLRYAIRH----------
GFKKIIVYGDSQLVCYQV-KGTWQTKN-----------QNMMELCKEV-RKLK---------ENFVSFEINHIR-RE----------
WNAEADRQANIAI-
>Haloarcula_marismortui_RNase_H
--YFDGASRGNPGPAS-----VGYVLVNDS-----GIVTEG---GET----IGTATNNQAEYKALIRAVEVARDY----------
GFDDVHIRGDSELIVKQV-RGEWDTND-----------PELREHRVRV-RELL---------TDFDDWQIEHVP-RE----------
INDRADELANDALD
>Halorhabdus_utahensis_RNase_H
-VYVDGASRGNPGPAA-----IGWVILTGD-----GIVTEG---GKR----IGSTTNNRAEYEALIHALEIAADY----------
GFDSVEVRSDSELAVRQV-RGEWDTND-----------PDLRERRVRV-RELF---------REFDDWSIEHVP-RE----------
INERADALANEA--
>Natrinema_pellirubrum_RNase_H
-VYFDGGSRGNPGPAA-----IGWVIVTGD-----GIVAED---GET----IGTATNNQAEYEALIAGLEAARDY----------
GYDEVHVRGDSELIVKQV-RGEYDTND-----------PDLRENRVTV-HELL---------RAFDEWTLEYVP-RE----------
VNDRADGLVNEALD
>Halorubrum_lacusprofundi_RNase_H
-VYFDGASRGNPGPGA-----VGWCLVTAD-----GIVAEG---GER----IGRVTNNQAEYAALIRALEAADEY----------
DFDEIDVRGDSQLIVKQV-RGEWNAND-----------PELRERRVRV-RELL---------ERFDRWSIGHVP-RE----------
INERADDLANEALD
>Mycolicibacterium_gilvum_RNase_H
IVECDGGSRGNPGPAG-----YGAVVWSAD-----RESVLA-EARSA----IGTATNNVAEYRGLIAGLESAAQT----------
GATEVEVRMDSKLIVEQM-AGRWRVKH-----------PDLAALHRQA-ADLA---------RRFEHITYSWIP-RE----------
QNSYADRLANEAMD
>Mycobacteriaceae_RNase_H
LVEADGGSRGNPGPAG-----YGSVVWSED-----RSSVLA-EAKQS----IGRATNNVAEYRGLIAGLEEAANV----------
GATEVAVSMDSKLIVEQM-AGRWRVKH-----------PDLIPLYQRA-RELA---------EGFDRVTYSWIP-RA----------
ANARADRLANEAMD
>Micromonospora_aurantiaca_RNase_H
VIEADGGSRGNPGPAG-----YGAVVRDPE-----TGEVLA-ERSES----IGTATNNVAEYRGLIAGLEAAAEL----------
GAAEVEARMDSKLVVEQM-CGRWQIKH-----------PGLRPLAAQA-AGLV---------SRFAAVRFTWVP-RE----------
RNRHADALANAAMD
>Vulcanisaeta_moutnovskia_RNase_H
-VYFDGACEPKNPGGVG----TYGFAVYND-----NDTISE----GY----GINCTNNVAEYTGLIKALECLLLH----------
DYGSIVVHGDSQLVIKQL-TGIYNVRT-----------EHLKPLFEKA-HELL----------SHFRVRLEWIP-RE----------
LNSKADELSKRAY-
>Ignicoccus_hospitalis_RNase_H
ELYFDGLCEPVNPGGVA----TYGFVVKEG-----GKVLCS----GK----GLDVTNNVAEYTALIKALECLLEK----------
GLEEVVVKGDSQLAIRQL-RGEYKVRS-----------PRIAPLYKRA-KELL----------SKFKAELQWVP-RE----------
LNEEADALSREAFR
>Ferroglobus_placidus_RNase_H
```

```
VLYFDGACLPVNPGGIA----TYGFVIITP-----SGEIVK----EK----GIKGTNNIAEYTALIRGLEKALEL---------
GIDELIVRGDSQLAIYQM-NGVYAVKS----------PNIIPLWQRA-MELF-----------RKIRFEWVP-RE---------
QNSAADELSTQAY-
>Bel
HGFADASSHA----------YGAVVYARVAV-GCSFQVTL-VAAKTRVAPIKPVSIPRLELNAALLLSRLLSIVKTS----
LTIPLFSTSCWTDSEIVLHWL-SAPPRRW--------------NTYVCNRTSEIL---------SDFPRSCWNHVR-TE---------DN-
PADCASR----
>Tamy
HTFVDASETA----------YAAACYWRAET--EVIHVAL-IAGKARVSPAKPVTIPRLELQAALLGARLARTVTEE----
IDLQVTDRYFWSDSSTVLQWL-KADPRRF--------------KPFVAHRLAEIK---------DLSKPHEWRWVP-TR---------DN-
PADIATR----
>Zebel
HIFCDASERA----------YGSVAYLRASD-DQGHVAVTFVLARSRVAPRKCLSMPRLELSAALTGAQLAQVIQNE----
LTLPIDSVTLWSDSTTVLYWL-TAESCRY--------------KVFVGGRVAEIQ---------TLTETAEWRYVD-SA---------NN-
PADHITR----
>Purbel
CIFCDASEIA----------YGACAYLRWKT-DDDKYEVRFVSAKSKVAPLKALTIPRLELLAGVLAARMHEAISNE----
MRLQVEKVVFFTDSMIVLQWI-KSSARTY--------------KAFVSSRVGEIQ---------TLTNPADWKHIP-GE---------VN-
IADKVSR----
>Kobel
HAFADASEIG----------YGCVIYLRQIN-SSENIHCSFVFAKARVTPLKKITIPRLELTAATLAVRLVSIVQRE----
LDFKIDKAIYWTDSTAVLRYI-RNDRARY--------------HTFVANRVQVIR---------EATVPEQWHHVD-TK---------RN-
PADLASR----
>Cer10-1
IMFSDASTDI----------YATAVYVQYSY-KSRKPVTRLLTSKSKIRSAKMWTIPKLELLGIECSSNLARTIIAT----
LATKIKKVRLFTDRACALYWI-LSHKTTR--------------VSQIQANQKILE---------ECGIETSLHHCP-TK---------EN-
PADLATR----
>Tor1
IVTTDASDTGY----------GGMLSQKIGD--DPEQPLG-FTSGFFRGPSTRWAINEKELFAFIKTLEVFHHHT--------
YGRAFTWRTDSRCLAFIC-AETNGK--------------PSAKKLRW-LEKL---------GEYDFSISHVS-G---------TEMAVPDGLSR-
---
>HERV-L
AWFTDGSADMQATTTKWQK-WTAAALQPLSR----TSLKDS----GE----DSGESSQWAELAVHLVVLHFAWKE----------
KWPVYLLTDSWAVANGL-AGWSGTWKKHDWKIGDDKEIWGRGMVGMW-MDLS---------KTVKIFVSHVS-
AHQSVTS-AEEDFNNQVDRMTHSVDT
>MuERV-L
AWFTDGSARYAGTT-QKW---TAAALQPLSG----TTLKDT---------GERKSSQWAELRAVHMVLQFVCKK----------
KWPVRLFTDSWAVANGL-AGWSGTWKDHNWKIG-EKDIWGR---SMW-IDIS---------AKDVKIFVSHVN-AHQKVTS-
AEEEFNNQVDKMTRSVDS
>Loki-Npa
CVYIDGCAYHTTVGDERKL-VAGIGITWTSE---FPDVSVG----YS----IGAKSSQVAELAAVYKTVQIALEH---------
NITEFVIVTDSNYVRNSF-VEYLPTWKRNNMLRSNNKPVKHAKLFGAI-DELV--------KSNDLTIYWKKTK-GHSKTQN-
MDKEGNDLADKLAKQGAI
>Loki-Xla
CVYVDGCSFHTEEAEKML--VAGIGIVWNNI---FPEISIG----YK----IGPKSSQVAELAAVYKAIQMAIEY---------
DLKEFVLITDSDYVRNSF-VEYFPGWKRSNMLRSNKKPVKHGKLFCKI-DELV--------TTYGLTIYWKKVR-GHSKHPG-
ADKDGNDLADSLAKKAAI
>Loki-Xtr
CVYVDGCSFHTEEAEKML--VAGIGIVWNNI---FPEISTG----YK----IGSKSSQVAELAAVYKAVQMAIEY---------
DIKEFVLITDSDYVRNSF-VEYFPGWKRSNMMRSNKKPVKHGKLFCKI-DELV--------TTHGLNIYWKKVR-GHSKHPG-
ADKDGNDLADSLAKKAAM
>Loki-Rca
CVYVDGCAYHHKDKDQLQL-VAGIGIIWTQG---VAPISVG----YK----LGPRSSQVAELAAILKAVEMAIDY---------
HLKEFAIITDSNYVRNSF-VEYLPRWKRNNMVKSNHKPVKHGKLFCQI-DQLV--------REHEITLYFKKTK-GHSKVLG-
EEKEGNDQADTLAKEGAL
>Loki-Ofa
RVYVDGCSFHHESQTR-----AGVGIVWVNH---KVDEPNH----YQ----LGPKTSQYAEIAAVLITLQQAAKL---------
AIVQLVICSDSNYARHSF-ISHFPVWKENGMKNARNKEVKHSELFLAC-DQLV--------TDRGMTVYWKKVK-GHSQTAG-
PDKDGNDEADRLAKLGAE
>Loki-Sme
```

```
TAYVDGCSFHHETQVR-----SGAGVLWVHD---TPCEPQC----FQ----LGSQSSQYAEIAAILITLQLAKEN----------
DVKNLTICTDSNYARLSF-TSHLPLWIKNGFMTSNRKPVKHKELFIAC-KQLV--------DTCDMQIYWKKVR-GHSRVPG-
ADKDLNDQADKLAKQGAI
>Loki-Sri
TAYVDGCSFHHETQIR-----AGVGVVWVHD---TPYKPQM----FK----LGSQSSQYAEIAGILITLQLAVEK----------
EIKTLVICTDSNYARLSF-SCHLPLWQRNGFITSNRKPVKHKELFVAC-NHLV--------DTQDMQVYWKKVR-GHSRVPG-
QDKELNDLTDTLAKRGAV
>Loki-Oja
TAYVDGCSFRKDKELL-----AGVGIAWLHD---DPCEPKS----FH----LGSQTSQYAEIAAVLIVLQLAVSL----------
NIHTLVICTDSNYARLSF-SCHLDLWKQNNFQTSNRKPVKNKELFVAC-DHLI--------TSHNIQIYWKKVR-GHSRIPG-
QDKVMNDLSDSMAKQGAL
>Loki-Bpe
TVYVDGCSFHHEATVR-----AGAGTVWLNN---DPCEPEK----FQ----LGAQTSQYAEISAVLIALNQAAAH----------
DVDTFVICTDSNYARLSF-SCHLPLWKSNGFLTAAKKSVKHKELFIAC-DRLV--------ETKNMKVYWKKVR-GHSREPG-
PDKELNDRADALAKQGAL
>Loki-Ame
TVYVDGCSFRHDQEGLK----AGAGIVWLDD---NPCEPQQ----FK----LGSQTSQYAEIAAILITIQLAIDQ----------
GVKTLVICTDSNYARLSF-TCHLPIWKTKGFLTSGRKAVKHTELFTAA-DYLV--------VRHDMLVYWKKVR-GHSRVPG-
TDKTYNDQADSLAKRGAL
>Loki-Lro
TAYVDGCSYNHEGLLK-----AGAGVLWLND---RPCPPQH----FK----LGPQSSQYAEVAAILITLQIASAH----------
NIRDLLICTDSNYARLSF-TCHLPNWKQNGFKTANNKPVKHQHLFQIC-DDLT--------SKHDMIIYWKKVK-GHSKQPG-
TDKDLNDQTDALAKAGAL
>Loki-Dcl
TAYVDGCSYNQEGKLK-----AGAGVLWLDN---RPCPPQH----FK----LGPQSSQYAEIAAILITLQIASTH----------
NIRELLICTDSNYARLSF-MCHLPNWKQNGFKTANKKPVKHQQLFQAC-DNIT--------REHDIMVYWKKVK-GHSKQPG-
LDKDLNDQTDALATAGAL
>Loki-Tfu
TAYVDGCSYNHEGNLK-----AGAGVLWM-----XPCQPQH----FR----LGPQSSQYAEIAAVLIALQIASAH----------
NIRDLLICTDSNYARLSF-TCHLPSWKKNGFRTANNKQVKHKHLFLAC-EDIT--------NEHDMTIYCKKVK-GHSKQPG-
TNKDLNDQTDALAKAGAL
>Loki-Str
MAFVDGCSYRHLDHLQ-----AGVGLVWHND---IPCKPLQ----FQ----LGNKTSQFAEVAGVLITLQTAVKH----------
ELAELAICTDSNYARLTF-LCHLPFWKQKHMTTSSGKEVKNKELILAC-DDLI--------TKHDIQVYWKKVK-GHSKSPG-
RDKLGNDHADSMAKSGAI
>Loki-Sau
KVYVDGCSFHHESQLR-----AGVGVVWDH----KNDEPDH----YH----LGPKLSQYAEIAAVLIVLQQAAKL----------
DVVQLVICSDSNYARHSF-ISHFPTWKVNDMKNARNKPVKHAELFQAC-DLLV--------TDKGIMVYWKKVK-GHSQIPG-
PEKDGNDEADRLAKLGAE
>Loki-Pyo
KVYVDGCSFHHQSHLR-----AGVGIAWVDY---NCSEPTH----YQ----LGPKTSQYAEIAAILIALQQAVKL----------
AITQFVICSDSNYARHSF-ISHFPLWKEKGMKNARNKEVKHSELFLTC-DQLV--------TEKGLIVYWKKVK-GHSQIAG-
PDKDGNDEADRLAKLGAE
>Loki-Bsp
EVYVDGCSFRCDSQLR-----AGVGIVWANR---QFEEPNH----YL----LGPKTSQYAEIAAALIAIQQAARQ----------
AQSQLVICSDSNYARHSF-ISHFPTWKASGMKNARNKPVKHAELFLAC-DQLV--------TDKGMTVYWRKVK-GHSQTSG-
PDKEGNDEADRLARLGAE
>Loki-Ape
EVYVDGCSFHHESQTQ-----AGVGIYWVNY---RSEEPSH----YQ----LGPKTSQYAEIAAGFIALQQAARQ----------
KISQLVICSDSNYARYSF-VAHFPAWKEKGMLNARNKEIKHAKLFLAC-DRLV--------SQQGMTVYWRKVK-GHSQISG-
GDKEGNDEADRLAKLGAS
>Loki-Pfl
EVYVDG-SFHHESQTQ-----AGVGIYWVKY---RSEEPSH----YQ----LGPKTSQYAEIAAGFIALQQAARQ----------
KISQLVICSDSNYARYSF-IAHFPAWKEKGMLNARNKEIKHAKLFLAC-DRLV--------SQQGMTVYWRKVK-GHSQISG-
GDKEGNDEADRLAKLGAS
>Loki-Cne
RVYTDGCAFMHEHQLR-----AGVGIVWTG-------------------------------IDSILITIQQAAAL----------
NQKQLVICSDSNYARLSF-ISHLPVWKTNGMKNARNKEVKHSNIFLAC-EKLT--------TENDMTIYWKKVK-GHSRTLG-
QDKDGNDEADRLARLGAE
>Loki-Lch
```

```
KAYVDGCSFHREGQLQ-----AGIGIFWEGN---FPCTPLQ----YK----LGPKSSQFSEIAAVFVAIETAVKN---------
KLSDIVICTDSDYARHSF-CKFLPTWKANGMKNSRGKTVKHAELILAV-DQLV--------LENEITVYWKKVK-GHAKDSS-
QDKYGNDLADKLAKEGAI
>Loki-Pma
MVYIDGCSKMKENKLH---AGSGIVWETGPL---TGIQEGF-----Q----LGPKSNQYAELAGVHIAVQMASNN---------
QIQTMVVCTDSNYVQHSF-LHHLPIWKKNGMKNHRNKPIHHQELFEAI-DTMV--------QNKDMKIFWKKVK-GHSKIPG-
QEKTGNDQADALAKVGSL
>Loki-Lca
RVYIDGCAKMKENKLH---AGAGIMWESGPL---TGIQEGF-----Q----LGPKSNQYAELAGVHIALQMAADN---------
KIQTMVVCTDSDYVQYSF-LNHLPIWRKNGMKNHRNRPIHHRELFEAI-DTMI--------QQNDMKIFWRKVK-GHSKIPG-
QEKAGNDRADALAKLGSL
>Loki-Aro
MVYMDGSAHRKTGTP-----GAGVGLLWTNC---DLPPQKF-----K----LGPKSSQYAELAGTLIALXAAQNN----------
YHDLVICTDSDYTVRSF-LDFLPNWLANNMIKSDGKLVAHTELFTAV-QHLL--------DTHNLTVYWKKVP-GHSKLDS-
PNKHGNDQADTLVKEGVS
>AliFLERV
VLFTDGSRKGPDE-----TAYWGYILTQNGK-----EQHRA-RGK-------VSGSAQAGEVTAVLEGLLELTKR---------
KIKAARVVTDSYYCAQAL-KEDLSIWIENGFESARGKTVAHKELWEKIAELRT-----------QLDLDVEHQK-AHTKEG--
SYWEGNDEVDRFVQ----
>AciFLERV
TLFTDGSRKGPDD-----TAYWGFILKLKGK-----EQTRQ-RGK-------APGSAQAGEVTAVLEGLLELEKR---------
KIKRAKIITDSYYCAQAL-KEDLTIWEENGYETAKGKKVAHEDLWKKIAELRM-----------NLDLEVVHQK-AHVKEG--
AHWRGNEEVDRFVQ----
>ZFERV-2
CYYTDGTTEQGIT-----RWGWGWIRKKGRK---VTESDKG-SLE-------AGLSAQAAEVKAFRKALEKAKQL---------
NDLKCFIVTDSDYIYQAV-EEHLGAWKNNDFNTSKGKPLKQADDWRVI-DELI----------TEITPTVLHQT-SHTSQNT-
TAAIGNKEVDSYVRRVRA
>CoEFV
VYYSDGSAKRNQQ-----KNYAGIGIVKGKF---HFEPEET-KAVP-----LGPASAQYAEVMALLDAVKQATD-----------
TGPVLICTDSVYAQRGY-TEDLHYWAIRNFHDSRNAKLKYADKWKQL-DQLK---------RDKPLVRVIHVP-GHTPGT--
VHSCGNGLADSLAQGGTY
>EFV
VFYTDGSSIKSPKKDKQHSAGMGIIAVRYQP---QMNIIQE-WSIP-----LGDHTAQFAEIAAFEFALKQAIRK----------
MGPVLIVTDSDYVAKSY-NQELDFWVSNGFVNNKKKPLKHVSKWKSI-ADCK---------KHKADIHVIHEP-
GHQNDLQSPYAMGNNAADKLAVKASY
>SFV
VFYTDGSAIKSPKTEQTHSAGMGIVMVVYTP---EPNITQQ-WSIP-----LGDHTAQYAEISAVEFACKKASLL----------
QGPVLIVTDSDYVARSA-NKELPFWRSNGFLNNKKKPLKHISKWKNI-SDSL---------LLKRNITIVHEP-
GHQPSKTSIHTLGNSLADKLAVQGSY
>BFV
IFYTDGSAIRSPKPNKTHSAGMGIIQAKFEP---DFRIVHL-WSFP-----LGDHTAQYAEIAAFEFAIRRATGI----------
RGPVLIVTDSNYVAKSY-NEELPYWESNGFVNNKKKTLKHISKWKAI-AECK---------NLKADIHVIHEP-
GHQPAEASPHAQGNALADKQAVSGSY
>FFV
IFYTDGSAITSPTKEGHLNAGMGIVYFINKDG--NLQKQQE-WSIS-----LGNHTAQFAEIAAFEFALKKCLPL----------
GGNILVVTDSNYVAKAY-NEELDVWASNGFVNNRKKPLKHISKWKSV-ADLK---------RLRPDVVVTHEP-
GHQKLDSSPHAYGNNLADQLATQASF
>Trypanosoma_brucei_RNase_H
VVYVDGACRNNGSRSRERR-GFGGFYGDGDS---RNFKFPL----PA----HEPQTNQRAELSALIHVLRVALDS---------
PCYNLCVYSDSKYTVMGV-NSYLHRWERNGFKTAGGGDVANIDLWSQF-TKLR---------AERVALQLKHVP-GHAGV----
--YGNEMADRLAVEACE
>Pisum_sativum_RNase_H
TLMFDGAVNG-----------VGAVLINPKG------AHMP-FSARLTF----DVTNNEAEYEACIMGIEEAIDL---------
RIKTLDIFGDSALVVNQV-NGDWNTL---------------IPYRDYT-RRIL---------TFFKKVKLYHVP-RD---------
ENQMADALATL---
>Cer3
HIFTDASAVA-----------QGAVLMQQMED-STKDYAAIAYTSRTLSDTESRWPAIQTELGAIIFALRQFRPYI---------
GQSRTTIHSDHRPPMYLL-GKSKV----------------NDNLARW-LIEL----------AQYDTRIVHID-GK----------KNTVADCLSR---
-
>Tor2
```

```
ILLSDASELA-----------AGSVLMQNIDK---RQRLIA-VSSKTFSETERKWSATERECYSLLIGCEKFEYYL--------
KGPVGFIALVDHKALLALD-KRYLA-----------------NSKLKRW-QTRL----------AEFKFTVQYVE-GR----------
SHVFADMLSR----
>Ty1B
VVISDASYGNQPYY-KSQ---IGNIYLLNGK-------VIG-GKSTKAS--LTCTSTTEAEIHAISESVPLLNNLSYI----
QELDKKPITLLTDSKSTISII-ISNNEEFRN----------RFFGTKAMRL-RDEV----------SGNHLHVCYIE-TK----------KN-
IADVMTK----
>Ty4
IAITDASVGSEYDA-QSR---IGVILWYGMN-------IFN-VYSNKST--NRCVSSTEAELHAIYEGYRDSETLKVL----
GEGDNNDIVMITDSKPAIQGL-NRSYQQPKE----------KFTWIKTEII-KEKL------------KRSITVKIT-GK----------GN-
IADLLTN----
>CSVMV
IIEVDASNNA-----------YGSCLKYKPKN--KIEYLCR-YNSGTFKENEQKYDINRKELIAVYQGLQSYSLFT---------
CEGNKLVRTDNSQVYYWI-KNDTNK--------------IEFRNIKYL-LAKI----------AVYNFEIQLID-GK----------TNIIADYLSR----
>SbCMV
IVETDASQHS-----------WSGCLRALPKG--NELLLCK-YVSGTFTDTETRYPIAELEVLAGVKVLEKWRIDL---------
LQTRFLLRTDSKYFAGFC-RYNIKT--------------YRNGRLIRW-QLRL----------QAYQPYVELIK-SE----------NNPFADTLTR--
--
>CaMV
IIETDASDDY-----------WGGMLKAIKIN--NTELICR-YASGSFKAAEKNYHSNDKETLAVINTIKKFSIYL---------
TPVHFLIRTDNTHFKSFV-NLNYKG--------------SKLGRNIRW-QAWL----------SHYSFDVEHIK-GT----------DNHFADFLSR-
---
>PVCV
ILQTDASDQY-----------WSAVLLEEHN---GKRKICG-FASGKFKVSEQHYHSTFKEILAVKNGIKKFNFFL---------
IHTNFLVEMDMRAFPKMI-RLNPKI--------------VPNSQLLRW-AQWF----------SPYQFEVKHLK-GK----------
DNILADFLSR----
>Kabuki
ALMTDASD-------------MGAVLQQKV-----VWQPLG-YFSRKLSPAEQKYCTYDRELLAIYKAMIYFRKL----------
EGRPLTVYTDTSPCVMHS-VKYVI-----------------KKRLDARGICFS-----------LGEFTTVLQSPGN-----------------------
>Tor4a
VLETDASDKG-----------YGGVLYICESD--NCLIPVC-YNSGNFSPVQQNYTIVEKELLSGKLCMEKWAIYL---------
AFKKFQWITDNSNIKYVR-TLR-----------------TNNQKIARW-LTDL----------QSFSFTISQRP-SS----------KMKISDFLSR----
>RTBV
IIETDASEEG-----------WGAVLVCKPDK--KDTEKIA-GYASGNFGEKKTWTSLDYEIEAINEALNKFQIYL----------
DKDFTIRTDCEAIVKGI-KTED----------------YKKRSKTRW-IKLK----------DGYKPTFEHIK-GN----------KNFLPNFLSR----
>ComYMV
IIETDGCMTG-----------WGAVCKWKMSKHRSTERICA-YASGSFNPIK---STIDAEIQAAIHGLDKFKIYY--------
LDKKELIIRSDCEAIIKFY-NKTN-----------------ENVRWLTF-SDFL--------TGLGITVTFEHID-GK----------HNGLADALSR---
-
>Skippy
ELETDASDFA-----------LGGQIGQRDDN--GVLHPIA-FYSHKMHGAELNYPIYDKEFLAIVNCFKEFRHYL-------
RGSKHPVKVFTDHKNIAYFA-TTQE-----------------LNRRQLRY-AEYL----------CEFDFTIAHCK-GT----------
DNGRADAISR----
>Amn-san
FIEVDASDVG-----------AGAILSQRHSA--GKLHPCA-YFSKKFSSAEQNYDIGNRELLAVKLALEEWRHLL-------
EGASHPVTIYTDHKNLEFLQ-SLKR-----------------QNPRQARW-SLFF----------SRFNFVLTYRP-GT----------
KNRKADALSR----
>Sushi-ichi
IVEVDASDAG-----------IGAVLSQRSEA--QKIHPCA-YFSRRFDPAERNYDVGNRELLAVYGALVEWKHWL-------
EGAKHPFLVWSDHKNLTYVR-TAKR-----------------LNPRQGRW--ALF----------SRFDFTLTFRP-GS----------
KNIRADALSR----
>TF1
LLETDASDVA-----------VGAVLSQKHDD--DKYYPVG-YYSAKMSKAQLNYSVSDKEMLAIIKSLKHWRHYL-------
ESTIEPFKILTDHRNLIGRITNESE-----------------ENKRLARW-QLFL----------QDFNFEINYRP-GS----------ANHIADALSR----
>Maggy
ILETDASDYV-----------SAGILSQYGDD--GILRPVA-FFSKKHTATECNYEIYDKELLAIIRCFEEWRPEL-------
EGTSSPVQIITDHRNLEYFT-TTKM-----------------LNRRQARW-AEFL----------SRFNFRITYRP-GK----------QGAKPDALTR-
---
>Pyggy
ILECDSSGYA-----------TGGVLSQYDDE--GVLRPCA-YFSKKNNVHECNYEIHDKELLAVVRCLEEWDAEL--------
RSVKSFKVITDHKNLDVLY-EAKM-----------------LTNVVGQA------------------AESIQYGA-GK----------QNVRADALSR----
>Mdg3
```

```
ELHCDASSSG-----------FGAVLMQKKE---DQKWHPVSFFSKRTTDIESKYHSFELETLAIVYSLRRFRVYL---------
HWRTFKIVTDCNSLILTL-SKKE-----------------LNPRIARW-ALEF----------QGYDFEIVHRA-GS----------RMQHVDALSR----
>Mdg1
CITTDASKQA-----------CGAVLSQDHN---GQQLPVA-YASRSFTKGESNKSTTEQELAAIHWAINHFRPYV---------
YGRHFLVQSDHRPLSYLF-SMRN-----------------PSSKLTRM-RLDL----------EEFEFTVEYLK-GK----------DNHVADALFR---
-
>SPM
VLACDASPYG-----------LGVVLSHLMED--GEERPVA-YASRTLTKSEQNYSQIEKESLAIIYGVTKFHKYL---------
YGRKFTLLTDHQALTIIF-GSKKG---------------IAAARLQRW-ALIL----------MAHQYEIKYRK-ST----------EHANADVLSR--
--
>Mag
VLTVDASARG-----------LGAVLAQRGPG--CQERVVA-YASRALTTHELHYSQIHKEALAIVFAVEKFHQYL---------
YGRKFILRTDHKPLVSIF-GPNIG---------------IAASRLQRW-AIKL----------SAYDFEIEYVR--T----------DKNVADALSR----
>Cer4
IVAADASKYG-----------IGGVILHVNPD--GVEVPIA-HFARSLTETEKRYSQIEKEALALIYTVKKSHKFV---------
FGRRFKLQTDHRPLLALF-GDNRD---------------LSQNRIVRW-ATTL----------MSYDFELSYVA-TE----------
KFAKADWLSR----
>SURL
KLSADASKNG-----------IGAVLLQQHD---ENWVPIA-YASRSMTDAETRYAQIEKELLAITYACERFHQYI---------
YGQQVEVETDHKPLIPLF-VKSLG---------------DIQRLLIRV---------------QRYDLKVMYTP-GK----------YMYTADTLSR----
>Zam
ILTTDASNFA-----------LGAVLSQGSL---QNDRPVC-FASRTLSDTEVNYSTIEKEMLAIIWAVKYFRPYI---------
YGVKFTIVTDHKPLIWLM-NFKE-----------------PNSKIIRW-RLQL----------MEYNFEIIHKK-GS----------QNVIADALSR----
>Cigr-1
QVETDASDYA-----------IAAVLSQS-------GRPVA-YMSRTLNTCERNYPAIEKEATAVVEAVRKWSHFL---------
KGKSFTLVTDQRSVSFMF-DQRNRG---------------KNSKILMW-RLEL----------SQFTYDIVHRP-GR----------DNFVPDALSR-
---
>Gmr1
TLQTDASEVG-----------LGGVLSQARNE--REEHPVT-YLSRKLLPHERNYSTVEKEALAIKWAVNKLTYYL---------
LGHQFVLVTDHAPLKWMA-TAKD-----------------TNARITRW-FLSL----------QPFSFTVEHRP-GR----------EHTNADALSR--
--
>Tom
VLTTDASNLA-----------LGAVLSQ-------DNHPIS-FISRTLNDHELNYSTIEKELLAIVWATKTFRHYL---------
LGRHFQIASDHQPLRWLH-NLKE-----------------PNAKLQRW-RIRL----------AEFDFHIEYIK-GK----------QNSIADALSR----
>Diaspora
ELMCDASNYA-----------LGAVLAQKID---KLPRVIY-YASRTLDAAQANYTTTEKELLAIVFALEKFRSYL---------
LGTHIIVYTDHAALKYLL-KKVD-----------------SKPRLIRW-MLWL----------QEFDLEIRDRS-GA----------QNLVADHLSR---
-
>Athila4-1
EIMCDASDYA-----------VGAVLGQKID---KKLHVIY-YASRTLDDAQGRYATTEKELLAVVFAFEKFRSYL---------
VGSKVTVYTDH-ALRHLY-AKKD-----------------TKPRLLRW-ILLL----------QEFDMEIVDKK-GI----------ENGAADHLSR---
-
>Woot
YVQTDSSGYG-----------LGAELYQIQED--GSRGVIA-FASRSLRGPELNYTTTEKELLGVIFALHKFRIYI---------
QVTKIIIRTDHQALKFLS-RCRL-----------------LSERLTRW-TLIL----------GQYDYEIELVK-GK----------GNVVADILSR----
>Gypsy
DLTTDASASG-----------IGAVLSQEG-------RPIT-MISRTLKQPEQNYATNERELLAIVWALGKLQNFL--------
YGSREINIFTDHQPLTFAV-ADRN-----------------TNAKIKRW-KSYI----------DQHNAKVFYKP-GK----------ENFVADALSR-
---
>Reina
VVETDASGSG-----------IGAVLQQKG-------HPIA-YISKALGPKNLGLSTYEKEYLAILFAVDHWRPYL---------
QHGEFFIKTDQQSLTHLE-DQKL-----------------STIWQQKA-ITKL----------LGLQFRIIYKK-GV----------ENRVADALSR----
>Galadriel
EVHTDASDKA-----------IGGVLVQEG-------HPVA-FESRKLNDAEQRYSTHEKEMVAVVHCLQVWRVYL---------
LGTRFVVRTDNVANTFFK-TQKK-----------------LSPKQARW-QEFL----------AEYDFMWEHKP-GK----------
HNQVADALSR----
>CRM
ELECDASGIG-----------LGGVLLQEG-------KPVA-YFSEKLSGSVLNYSTYDKELYALVRTLETWQHYL---------
WPKEFVIHSDHESLKHIR-SQGK-----------------LNRRHAKW-VEFI----------ESFPYVIKHKK-GK----------ENIIADALSR----
>Del
VVYTDASLAG-----------LEGVLMQDG-------RVVA-YASRQLKVHENNYPTHDLELAVVIFILKLWRHYL---------
YGEDFELYCDHKSLKYIS-TQKD------------------LNLRQRW-IEVL----------KDFDFSIFYHP-GK----------ANVVADALSR----
```

```
>Koala
SGYSDADWAGNGDDRRST---GGFAVFLGN-------SLVS-WSARKQP--
TVSRSSTESEYKAMANATAEIMWIQTLLKEIHVSNSPTARLWCDNLSAKYLS-SNPIFHARTK----------HIEVDYHFV-
RERV----------QQKLLEVGFVP-TG-----------DQIADGFTK----
>Hydra1-2
VAYCDADFASDLNDRRST---TGYCFSLSA-------PLIS-WKSKKQS--
TVALSTCEAEYMALTVTTQESMYLVQLLSMENDCMYMPVQIFEDNQGAIALS-KNPVCRQRCK----------HVDVRYHFI-
RSAL----------SDGKVTVEYCP-TE-----------DMVADVMTK----
>pCretro6
VGYSDADWGANPDDQKSI---SGYVFLLGG-------APVC-WASRKQK--
SVALSSMEAEYMAGSTAASQALWCRMLLEELGFAQPNPTLLYMDNQSALALA-RNTGTQGRAK----------HIDIRYHFL-
RDKI----------SSKEISVAHCP-GE----------DN-PADIFTK----
>Vitico1-1
VGYTDSDWAGCLEDRKNT---SSYMFSLG-----------------------------------------
SVWLRRILADISQEHEESTIIYCDNKAAIAMT-KNPAYHGRTK----------HVDIRVHFI-RDLV----------VEGKVVLQYCN-
TN-----------EQVADVLTK----
>Tnt-1
KGYTDADMAGDIDNRKSS---TGYLFTFSG-------GAIS-WQSKLQK--CVALSTTEAEYIAATETGKEMIWLKRFL-
QELGLHQKEYVVYCDSQSAIDLS-KNSMYHARTK----------HIDVRYHWI-REMV----------DDESLKVLKIS-TN----------EN-
PADMLTK----
>SIRE1-4
VGYCDADWAGSADDRKST---SGGCFYLGN-------NLIS-WFSKKQN--CVSLSTAEAEYIAAGSSCSQLVWMKQML-
KEYNVEQDVMTLYCDNMSAINIS-KNPVQHSRTK----------HIDIRHHYI-RDLV----------DDKVITLKHVD-TE-----------
EQIADIFTK----
>Copia
IGYVDSDWAGSEIDRKST---TGYLFKMFD-------NLIC-WNTKRQN--
SVAASSTEAEYMALFEAVREALWLKFLLTSINIKLENPIKIYEDNQGCISIA-NNPSCHKRAK----------HIDIKYHFA-REQV-
---------QNNVICLEYIP-TE-----------NQLADIFTK----
>CoDi6.7
EVYADADFSGDPDTARSR---TGFAIFYAG-------CPVT-WQSKLQT--
EISLSTTESEFVSLSTALRTALPLIGLALGFDISVTVPTVAFDDNMGAIEIA-LVPKMRPRTK----------HINVKYHHF-RQHV-
---------DNGDITIQHVD-SE-----------DQIADFLTK----
>1731
TGFVDADWGGDRLDRKSY---TGYVFFLSG-------GPVS-WRSEKQQ--
SVALSSTEAEYMALTTACKEAIALRRLIEIVCGDLKTPTVMHGDNLKAAQLA-KNPVHHSRTK----------HIDIRYH---
REVM----------KEGHVVLEYTS-TN-----------EMIADINTK----
>CoDi6.3
VCYSDSDYAGDPDTRRSV---SGYILYVKG-------VPIC-WRSKAQR--SITLSSSEAEWIALSEATKEIMFVLQLL--
LYIKVQLPITVRVDNIGAIWMS-QNVNTSSRTK----------HVDIRTKYV-NEYC----------EDGVLKIIFVK-SA----------DN-
DSDIMTK----
>Zeco1
VVFSDASLGNLPDGGTQ----GGTLIGLCW-------------QSKKIR--
RVVRSTLAGETLALSDGIDNAIFLTTLFTIGNAELNTPPLICVTDNHSLFDA--LKSTKQR-----------LRLEISSI---KDLQ-------
---SKKIKKVLWSD-TK-----------TQLADCLTK----
```

Dataset S7. IN alignment generated by Mafft 7.402 and used to reconstruct the phylogenetic tree in Fig. 4.

```
>Loki-Xla
HDA-PTAGHRGEKITYELLRDYAYWPHMLQDVRTYCQGCLVCPQFQPQG-PNHRAPLMKR----GISMPWSDIQIDFIGPV-
TTS----SKGNRYML-TVTCLFTKWVECLPCKTC-SSSVCASLLINHIF--SRFGLPQRIESDRGSHFTSEV-----------------
MTKMWEIL-GVKRKLHIAYRPASSGGVERYNQSIVNILKKFV------KE---SGKDWDVKLPLVLMAIRATPSAA-T-------
KLSPFEMITGRKMVLPQHL
>Loki-Xtr
HDA-PTAGHRGEKITYELLRDYAYWPHMLQDVRTYCQGCLICPQFQPQG-PNHRAPLMKR----GISMPWSDLQIDFIGPV-
TTS----SKGNRYML-TVTCLFTKWVECLPCKTC-SSSVCASLLINHIF--SRFGLPQRIESDRGSHFTSEV-----------------
MTKMWEIL-GVKRKLHIAYRPASSGGVERYNQSIINILKKFV------KE---SGKDWDVKLPLVLMAIRATPSAA-T-------
KISPFEMITGRKMVLPQHL
>Loki-Npa
HDA-PTAGHRGEKITYELLRDYAYWPHMLKDVQTYCQGCLVCAQFQPQG-PTHRAPLQKK----
GFSLPWSDIQVDFIGPV-TRS----SRGNKYML-TVTCVFTKWVECLPAHNN-TAETCATLLINHVF--
SRYGLPLRIDTDRGSHFISEV-----------------MSKMWKIL-GVKRKLHIAYRPASSGSVERYNQSIVNILKKFV------KE---
SGKDWDMKLPLVLMAIRATPSAA-T-------KMSPFELMTGRKMVLPQHL
>Loki-Rca
HDA-PTSGHRGEKITYEILRDYAYWPHMLRDVKAYCQGCLICPQFQPAS-PTHRAPLQKR----GICMPWSDLQMDFIGPV-
TKS----SRGNRYML-TVTCLFTKYV-CIPAPNN-SSDTCAALLINHIF--SRFGLPQRIESDRGTHFTSEV-----------------
MTKLWKIL-GVKRKLHIAYRPASSGGVERYNQTIVNILKKFV------KE---TGKDWDVKLPLVLMALRATPNTV-T-------
KMSPFELMTGRRMVLPQHL
>Loki-Sme
HDS-PCSGHRGIKATLKSLQEVAYWPHMQKDVLDYIKGCLVCCQFQPAS-TLHRAPLQRR----
GVSFPWSDLQIDWVGPL-TKS----TRGNKYFL-TVVCAFTKWVECLPAPND-TAQTTAYLLMNHIF--
SRFGLPGRVNSDRGTHFTAEM-----------------MQQLWQLL-GVKANFHISYRPQSSGQVERANRTVVNILRKYV------SS---
NHKDWDVKLPLVLMAIRATPHES-T-------GLSPFELMTGRQMTLPLHL
>Loki-Sri
HDS-PCAGHKGSKATYNALRQVAYWPQMRKDVADYIKGCLVCCQFQPAN-PVHRAPLQRR----
GISFPWSDLQIDWVGPL-TKS----TRGNKYFL-TVTCAFTKWLECLPAPND-TAQTTAYLLMNHIF--
SRFGLPSRVNSDRGTHFTADV-----------------MQELWQIL-GVKANFHISHHPQASGQVERANRTVVSVLKKYV------TS---
NHRDWDVKLPLVLMAIRATPHEA-T-------GFSPFEMMTGRQMTLPLHL
>Loki-Oja
HDL-PSSGHKGINATYSALKQIAFWPHMKKDVTDYVNGCLVCCQFQPTN-PTHRAPLQKK----GISFPWSDIQIDWVGPL-
TKS----SRGNKYFL-TVTCAFTKWVECLPAPND-TAQMTALLLMNHVF--SRFGLPMRVNSDRGTHFTAEV-----------------
MQQVWNTL-GVKANFHVSHHPQSSGQVERANRTVVNILKKYV------AA---SHKDWDIKLPLVMMAIRATPHES-T-------
GTSPFEMMTGRQMTLPLHL
>Loki-Ame
HDS-PSAGHKGIKETYRTLKQVAFWPRMREHVASYIKGCLVCCQFQPAN-PLHRAPLQRK----
GLTFPWSDLQMDWVGPL-TKS----ARGNKYFL-TVTCAFTKWVECLPAAND-TAITTACLLVNHIF--
TRFGLPCRVDSDRGTHFTAEV-----------------MQQMWHVL-GVKAKLHVSYHPQSSGQVERANRTVVSILKKFV------AT--
-HHRDWDVKLPLVLMAIRATPHDS-T-------GVSPFEVMTGRQMTLPLHL
>Loki-Lro
HDA-ACAGHHGTRATYETLKQVAYWPGMQQDVAEYVKGCLVCCQFQPAN-PNHRAPLQRR----
GVTFPWSDLQIDWVGPL-PRS----TRGNKYFL-TVVCQFTKWVECLPAPND-TAQTTAYLLMNHIF--
SRFGLPLKVNSDRGTHFTAEI-----------------MQQIWKLL-GIHAQLHISHHPISSGQVERSNQTVVSMLRKYV------SA---
NQKDWDVKLPLVLMAIRATPNES-T-------GVSPFELMTGRLMTLPLHL
>Loki-Tfu
HDA-PCAGHHGTKATYETLKQVVYWPGMQQDVTEYVKGCLRCCQFQPAN-PNYRAPLQRR----
GVTFPWSDLQIDCVGPL-PRS----TRGNKYFI-TVVCQFTKWVECLPAPND-TAHTTAYLLMNHIF--
NRFGLPLRVNSDRGTHFTAEV-----------------MQQIWKLL-GVKAKLHISHHPISSGQVERTNRTVVSMLRKYV------AA---
NQXDWDVKLPLVLMAIKATPQES-T-------GISPFELMTGRQMTLPLHL
>Loki-Bpe
HDT-PVAGHRGVQATYQALQQVAYWPHMRKDVQEYTKGCLVCCQFQPSR-PLHRAPLHKR----
GITFPWSDLQIDWVGLL-TKS----ARGNKYIL-MVTCAFTKWVEGLPAPND-TAQTTAFLLMNHVF--
SRFGLPCKIDSDQGTHFTAAV-----------------MSELWQTL-GVKATFHISHHPMASGQVERTNRTVINILKKYV------SA---
NQRDWDVKLPLVLMAIRATPNKS-T-------GISPFEMMTGRQMTLPLHL
>Loki-Dcl
HDT-PSARHHGTSTTYETLEQAAYWPSMQQDLAEYVKEFFPCYQFQPTK-TNHRASLQRR----
GVTSPWSDLQSDWVGPL-PRS----TRGYKYFL-TVVFQFTKWVECLPAPND-TAQTTAYLLMNHIF--
```

```
SRFGLPLRVNSDRGTHFTAEI-----------------MQQTWKLL-GIQARLHISQHPIASRQVERMNRTVVSMLRKFV------SP---
NQKDWDVKLPLVLMAIRATPNEA-T-------GVSPFELMTGRQMTLPLHL
>Loki-Str
HDE-PCGGHRGVKATCETLRQVAHWPHMEQDVARYVRGCLVCCQFQPTK-PLHRAPLQRK----
GVSYPWSSIQIDWVGPV-ARS----ARGNKYLL-TVTCAFTKWVECLPATND-TAETTAVLLLNHVF--
SRFGLPGEVDSDRGTHFSAAV-----------------MTELWKLL-GVKAKLHIAYHPRSSGGVERSNQSIVRILRKYV------AA---
NHKDWDLKLPLVLMAIRATRNRS-T-------GMTPFEMMTGRQMTLPLHL
>Loki-Sau
HDT-PIGGHRSYKATLKTLQQVAYWPFMARDTKLYVQGCLICCQFQPSR-PLNRAPLQER----GITFPWSHLQVDWIGPV-
PKS----SRGNKYLL-TVTCSFTKWVECLPAPND-TAVTTAVLLMNHVF--SRWGLPLSIDSDRGTHFTSSV-----------------
TTALYDIL-GVEAKFHIAYHPQSSGQVERANRTIVNMLKKYV------GS---NGKDWDMKLPLVLMAIRSTPHRS-T-------
GVTPFEMMTGRQMTLPIHL
>Loki-Ofa
HDS-PVGGHRGYRATLQTLQQVAYWPLMARDTKSYVQGCLVCCQFQPSR-PLGRAPLQKR----
GVTFPWSHLQVDWVGPV-PKS----SRGNKYLL-TVTCAFTKWVECLPAPND-TAMTTATLLLNHVF--
SRWGLPLSVDSDRGTHFTSSV-----------------MTTLYDIL-GVEVKFHISHHPQSSGQVERANRTIVSMLKKYV------SS---
NGKDWDVKLPLVLMAIRSTPHRS-T-------GVTPFEMMTGREMTLPLHL
>Loki-Bsp
HDS-PVGGHRGYKATLQTLQQVAYWPSMSKDVKIHVQSCLVCCQFNPSR-PLNRAPLQKR----
GMTFPWSHIQVDWIGPV-PKS----SRGNKYLL-TVTCAFTKWVECLPAPND-TAVTTAVLLLNHVF--
SRWGLPLSVDTDRGTHFTSHV-----------------MTELYHIL-GVNVRFHISHHPQSSGQVERANRTIVQMLKKYV------SS---
HGKDWDMKLPLVLMAIRATPHRA-T-------GVTPFEMMTGRQMTLPLHL
>Loki-Pfl
HDS-PVGGHRGPQP-XETLRQVAYWPSTNRDVKAYVRGCLVCCQFQPSQ-PLGRAPLQKR----
GVTFPWSHLQVDWVGPV-PRS----SRGNKYLL-TVTCAFTKWVECLPAPND-TATTTAVLLLNHVF--
SRWGLPLSVDSDRGTHFTSGV-----------------MTAMYDIL-GVKTRFHIPYHPQSSGQVERANRTIVNMLKKYV------NG---
SARDWDVKLPLVLMAIRSTPHRT-T-------GITPFEMMTGREMTLPLHL
>Loki-Ape
HDS-PVGGHRGPQPTLETLRQVAYWPSMSRDVKAYVKGCLVCCQFQPSR-PLGRAPLQKR----
GVTFPWSHLQVDWVGPV-PRS----SRGNKYLL-TVTCAFTKWVECLPAPND-TATTTAVLLLNHVF--
SRWGLPLSVNSDQGTHFTSGV-----------------MTAMYDIL-GVKARFHIPYHPQSSGQVERANRTIVNMLKKYV------NG---
SARDWDVKLPLVLMAIRSTPHRT-T-------GITPFEMMTGREMTLPLHL
>Loki-Pyo
HDS-PVGGHRGTRLLXSTLQQVGYWPSMSNDTMSYVRGCLTCCQFQPSQ-PLNRAPLQGR----
GVSFPWSHLQIDWIGPV-PKS----SRGNKYLL-TVTCAFTKWVECLPASNG-TAVTIAVLLLNHVF--
SRWGLSISIDSDRGTHFTSHV-----------------MTSLCKIL-GVEVRFHISHHPQSSGQVERANRTIIQMLKKYV------GS---
NGKDWDIKLPLVLMAIRLTPHRS-T-------GVTPFEMMTGRQMTLPLHL
>Loki-Cne
HDS-PVGGHRGYKATLHTLQQVAYWPLMAND--------------------------XRR----GVTFPWSHIQIDWVGPV-PRS----
ARGNKYFL-TVTCSFTKWVECLPAPND-TATTTAVLLLNHVF--SRWGLPLSVDSDRGTHFTSNV-----------------
MTALFEML-GVEVRFHISYHPQSSGQVERMNRTVVNMLRKYV------DS---HAKDWDVKLPLVLMAIRFTPQTS-T-------
GVTPFEMMTGRQMTLPLHL
>Loki-Lch
HDE-ECGGHRADRATYETLKQVAYWPNMYQDVQDYTKGCLVCCRFQSSS-PKHRAPLQTK----
GITMPWSDLQIDWIGPV-TRS----SRGNKYLL-TVTCLFTKWIECLPAPND-TAETTAMLLINHVF--
SCWDLPMRIESDRGSHFTGKV-----------------IENVWNTL-GVQRKLHVSHHPQSSGQVERANKTIVNILKKYI------ST---
SGKNWDVKLPLVLMAIKATPNKS-T-------GVTPFEMMTGREMTLPLHL
>Loki-Pma
HDE-PCSGHRCEMNTGQMLRQVAYWPKMGEDIHKHVSNCLVCCKYQPTT-SQHRAPLQKR----
GTDYPWSNLQIDWIGPV-NRS----TKGNKYML-TVTCAFSKWIECLPAPNN-TAETTAILLINQIF--
SRWGLPSTIDSDKGSHFTAEV-----------------ITEVWKML-GVKRQLHVAYHPQSSGQVERANRTIVSLLKKYV------ST---
TGRDWDTKLPLVLMAMRATPNRA-T-------GRTPFEIMTGRLMTLPVHL
>Loki-Lca
HNE-PCSGHRCEMNTGQMLRQVAYWPKMGEDVHKHVSNCLVCCKYQPTN-SQHRAPLQKR----
GTDFPWSNLQIDWIGPV-NRS----TKGNKYML-TVTCAFSKWVECLPAPNN-TAETTATLLVNQIF--
SRWGLPSTIDSDKGSHFTADV-----------------IAEVWKTL-GVKRQLHVAYHPQSSGQVERANRTIVSLLKKYV------ST---
TGCDWDTKLPLVLMAIRATPNRA-T-------GRTPFEIMTGRQMTLPVHL
>Loki-Aro
HDE-PCGGHRGEQPTIHMLWQVAYWFKIWTDVKDYVNGCTICHKFQPVL-PRHRAPLQSA----
PSAQPWSHIQIDWIGPV-TRS----TWGNXYAL-TITDRFTKWTKCLTALAD-TAETTAVLLMKHVF--
THWGLPQSVDSDGGTHFTATI-----------------MEEVWKTL-GVKRQLHIAYQPQSSGQVERANQTVINMLKKQL------DA---
NQKDWDIKLPLVLMTIXAMPSKA-T-------QVSLFEMMTDRQIVLPLHL
```

```
>Gmr1
RDA-LLGAHLGMDKTRERVVARFYWPGVRRDVARYCQECPDCQRVRPRA-VERSPLIPMP----IIETPFERIALDIVGPL-
PRT----SRGHRYLL-VILDYATRYPEALPLRAA-TSKAYAR--ISATL--SRVGLPKEILTDQGSCFMSRV----------------
VKELLKLL-QVSQLRTSVYHPQTDGLVERFNQTIKRMLKKSI------EA---DGKNWDQLLPHVLFAIREVPQAS-T-------
GFSPFELLYGRRPRGILDL
>ApGIN7a
HDS-LLGGHAGVSRTVKRLKLNYNWRHMKKDIKHYIKHCEICQKNKSHL-KTKKPMLITS----TIVKPFEKNCLDIVGPL-
PKT----LTGNMYIL-TLQDELSRFSLAIVLSTT-DAQTVAQAFVE-----VSYGIPESILTDRGTNFLFDV----------------
FKSMCKLL-DIKKSKTTPWHPQGNCFLERSHKTLKTYLRSFVD----------KDSNWDKLLCYATFCYNTTVHTS-T-------
NFTP---------------
>Mdg1
DDP-SEGGHSGISRTLRKMKNCCCWPRMTKAISEYVETCLKCQQAKTTK-HTKTPLTITE----TPATAFDKVLIDTIGPL-
PRS----ENGNEYAV-TIICDLTKYLVTVPIPNK-SAKSVAKAIFENFI--LKYGPMKTITTDMGTEYKNQI----------------
IDDLCKYM-KIKNITSTAHHHQTLGTVERSHRTFNEYVRSYI------SV---DKTDWDIWIQYFTYCFNTTPSVV-H-------
EYCPYELVFGRLPRQFIDF
>Cer3
HES-LGGGHFGYRKTLHKIKRNYYWPNMRSDVLKWTLQCKICQQKRNPH-PSTRELQKIV----ITTKVFEKVGVDLTGPL-
RMT----ASGNKY----MVCWFTKFVISVPLPNA-STETVADAMMKELL--LKFGTPSQLVSDRASTFTSEA----------------
FRAFCRKL-EIQQHLAIPYHSKGNGATERTFRTFHNMVSKHV------NK---THTDWDQILPYMTFVYNTTVHDT-T-------
GETPFFLIFGRDPVFAIDK
>DrGIN2-2
HFN-DIGHHLGQKKTVHRIQSKYYWLGIVKDVVDWIKVCETCQHAERSK--NLARTVRPI----KVDGPWEILAIEIIGPF-
PGT----HGGNTHIV-IITDYYSKWVEAFPVQKK-DGLCVARCISS-SV--CRFGPSKTIFCSQNADFCEEV----------------
MKQLCERW-NITVRVLSVDQPQRNALYDRSCNLLRDTIKQMV------VE---KQVEWDDFLDPILAIFRTSVNPT-T-------
KFTP---------------
>HmGINA1
YDS-NHAAHVGLNNTRAKLKGSFYWLGMVNDITKCVKECDKCQRMEKIR--TIAPELKSI----KVNGLWDFLGIDLIRPL-
PIT----KLGNKYIL-TITDLWSKYIEAFPIPDK-PAFYVSKCLTT-LF--YRFGPPKKILSDQGREFVNSL----------------
NEQLFFLF-QIKHLITSAYYPQTIGQDERTNQTIKKSLSKLS------ND---TQDNWDELLEAVLFGLRTCVQKS-T-------KFTP-
--------------
>HmGINO1
HND-AGGAHQGIVRTQNKIKNLYYWMSITSVVEAWIKTCKQCQSFEKIK--TIAPQLHPI----QVTERWAVLGVDLIGPL-
PET----VNGNKYVL-TITDLFTKWVVARVLPDK-SAAAVATALVN-NF--HTYGPPKKVITDQGREFVNEV----------------
KTKIFQLM-GIKHSVTSAYHPQSNGQDERTNQTFKCALSKYT------NN---EMNDWDRFSSXIAYGLNISCQKS-T-------
KVTP---------------
>HmGINNY1
HSAILGGGHFGRDKTLAKISERYYWKGMVNDVRSFCKYCDKCQRANRAF-EKHSAELHPI-K--
VKDEVWSTVGIDLIGPL-PLT----EKGNKYII-TATCLFSKWPEAASLSDK-TATSAAEFLYT-CF--
TRHGCCEVQISDQGREFVNEV-----------------NHELNKMM-GTKCNVTSAYHPQSNGEDERFNQTLQRQLLKYV------DE--
-KQNTWDLYIESILFSYRVSVQDS-T-------KQTP---------------
>HmGINGER1
EKAVALASHLGINAVRKQISSRFFWHSIVEDITKHVAQCERCQKTSNRN-LKVSPALKPV-K--VEQQVMKQVGVDLI-KL-
PES-----NGFNYAI-VLIDYFSKWTEAEPLIDK-TAVSVAAFLYR-VI--CRHGCFQIQINDQGREFVNSV----------------
SIALHDMS-GVQQRVTSAYHPQANGLVERQNQTIKKAIVKVL------NE---NVKSRASVLDGILFALRVKVQDS-M-------
GYSP---------------
>HsGIN1
HEN-DSGAHHGISRTLTLVESNYYWTSVTNDAKQWVYACQHCQVAKNTV--IVAPKQHLL----
KVENPWSLVTVDLMGPF-HTS----NRSHVYAI-IMTDLFTKWIVILPLCDV-SASEVSKAIIN-IF--
FLYGPPQKIIMDQRDEFIQQI-----------------NIELYRLF-GIKQIVIS-HTSGTVNPMESTPNTIKAFLSKHC------AD---
HPNNWDDHLSAVSFAFNVTHLEP-T-------KNTP---------------
>Amn-san
HSS-KQAGHPGSEKTLELLRRLVWWPTIRKDVRDFVAACTVCATTKASH-SRPCGLLHPL-P--IPSRPWTHLGMDFIVEL-
PPS-----CGNTVIW-VVIDRFSKMAHFIPLRKLPSAVELAHLFIQHIF--RLHGFPVEIVSDRGSQFVSRF----------------
WRSLCKSL-GVSLQFSSAYHPQTNGAAERVNQALEQFLRNHV------SL---CQDDWSDLLPWAEFAHNNASHSS-T-------
GRSPFLSVYGQHPLAFPQD
>Sushi-ichi
HSS-CFACHPGVRRTAEFVQRRFWWPNLQEDVREFVGACTVCARSKASH-RSPAGLLHPL-P--
VPSRPWSHVALDFVTGL-PVS-----QGNDTIL-TIVDRFSKGVHFVALPKLPSAAETAELLVSHVV--
RLHGIPLDVVSDRGPQFTSRV-----------------WQAFCKGI-GATVSLSSGYHPQSNGQAERANQAMEAALRCVT------TS---
NPASWSKFLPWVEYSLNAMESSA-T-------GMSPFQCFLGYQPPLFPQQ
>Maggy
```

```
HDS-PVAGHPGKAKTYDLLSREYYWPGMLHYVSLWVKKCQTCRRINPSR-EGHQGLLRPL-P--
TPERSWQHLSMDFITHL-PQS-----NGHDAIL-VVVDRLTKMRHFVPCKGTCNAEDTANLYLHHVW--
KLHGLPLTIVSDRGTQFVSKF-----------------WKHLTTRL-KIDSLLSTAHHPETDGQTERFNASLEQYLRAYV------AY---
LQDDWESWLPLAEFTANSHKSET-T-------GTSPFYATYGFHPRMGFE-
>Galadriel
YDS-AWAGHPGVERMLALLSRVYFWPKMEDDIEAYVKTCHVCQVDKTER-KKEAGLLQPL-P--
IPERPWLSVSMDFISGF-PKV-----DGKASIM-VVVDRFSKYSVFIAAPELCSSEVAAELFYKHVI--
KYFGVPADIVSDRDTRFTGRF-----------------WTALFNMM-GTELKFSTANHPQTDGQTERINHLLEEYLRHYV------TA---
SQRNWVELLDTAQFCYNLHKSSA-T-------EMSPFEIVLGKQPMTPLDV
>Reina
HDS-PIGGHSGFPVTYHRIKKLFYWAGMKGQIKEFVQSCEICTKAKADR-NRYPGLLLPL-P--IPDQAWQVISLDFISGL-
PTS-----RRFNCIL-VVVDKFSKYAHFLAMSHPFTALSVAKLFLSQVY--KLHGLPLSIISDRDPIFTSNL----------------
WQELFKLV-GTKLCLSSAYHPQSDGQTEWVNQCVEAYLRCFV------HG---CPKQWSNWLSLAEFWYNTCFHTA-L-------
GQSPFEVLYGHTPSQLGLS
>CRM
HGG-GLMGHFGAKKTEDILAGHFFWPKMRRDVVRLVARCTTCQKAKSRL--NPHGLYLPL-P--
VPSAPWEDISMDFVLGL-PRT----RKGRDSVF-VVVDRFSKMAHFIPCHKTDDATHIADLFFREIV--
RLHGVPNTIVSDRDAKFLSHF-----------------WRTLWAKL-GTKLLFSTTCHPQTDGQTEVVNRTLSTMLRAVL------KK---
NIKMWEDCLPHIEFAYNRSLHST-T-------KMCPFQIVYGLLPRAPIDL
>Del
HRS-KFTIHPGSTKMYRNLKINFWWSGIKREVVEYVSRCLICQQVKADH-HHHSGLLQPL----PVSEKWEHILMDFIIGF-
PLS----KRCHDSIW-VIVDRFTKSAHFIPIHTTISGKDLA-LYIKEII--RLHGIPTTIVTDRDTKFTSRF----------------WGSL-
KSL-GTELFFSTAFHPQTDG-SERTIQILEDMLRSCS------LD---FKGNWEEHLPLVEFAYNNSYQSS-I-------
GMAPFEALYGRPCRSPTCW
>Skippy
HAH-PLHGHQGVTKTMKRLQELGYRHFKKGQVEKVIKQCDLCAKTKAQR-HKPYGQLQPL-P--
VAQRPWDSITMDFITKL-PLS----STGYDSIM-VIVDRLTKFSYYLPYREATDAEELSYVFYRHIV--
SIHGLPTEILSDRGPTFAATF-----------------WQSLMARL-GLNHRLTTAFRPQVDGQTERMNQVLEQYLRCYI------NY---
EQNDWVEKLPIAQLAYNTAYNES-T-------KLTPAYANFGFTPNAYHNA
>TF1
HEE-GKLIHPGIELLTNTILRRFTWKGLRKQIQEYVQNCHTCQINKSRN-HKPYGPLHPI-P--PPERPWESLSMDFITAL-
PES-----SGYNALL-VVVDRFSKMAILLPCTKSITAEQTARMFDQRVI--AYFGNPKEIIADNDHIFTSQT----------------
WKDFAHKY-NFVMKFSLPYRPQTDGQTERTNQTVEKLLRCVC------ST---HPNTWVDHISLVQQSYNNAIHSA-T-------
QMTPFEIVHRYSPALSPLE
>Athila4-1
HGS-AYGGHFATFKTVKILQAGFWWPTMFKDAQEFVSKCDSCQRKGNIN-RRNEMPQNPI----LEVEIFDVWGIDFMGPF-
PSS-----YGNKYIL-VAVDYVSKWVEAIASPTN-DAKVVLKLFKTIIF--PRFGVPRVVISDGGKHFINKV----------------
FENLLKKH-GVK-------------QVEISNREIKTILEKTV------GI---TRKDWSAKLDDALWAYRTAFKTP-I-------
GTTPFNLLYGKLCHLPVEL
>Diaspora
HSS-APGGHLGVQRTAKVLDCGFYWPTIFKDAWKICSTCEQCQRAGNLT-WRQQMPQQPM----
LFCEVFDVWGIDFMGPF-PVS-----FGYVVYIL-LAVDYVSKWVEAKPTRTN-DAKVVADFVRSNLF--
CRFGVPKAIVSDQGTHFCNKT-----------------MHALLKNY-GVVHRVSTPYHPQTNGQAEISNREIKRILEKIV------QP---
SRKDWSTRLDDALWAHRTAYKAP-I-------GMSPYRVVFGKACHLPVEI
>RetroSor1
HSG-MCGSHIGPRALSKALRQGFYWPTHIRDAEEIVKTCKACQTFSPIQ-SGPSALTQLI----PASWPLQRWGMDLVGPM-
PTA----QGGNKFAV-VAIEYFTRWIEAKPLTTI-TSETIRKFFWQNIV--CRFGVPRLLTVDNGKQFDSDN----------------
FKEFCHLI-GTKIAFASVYHPESNGAVERANRTIFSAISKTL------LNL--RKGKWVEELPRVVWSHNTTVSRA-T-------
GFTPFKLLYGEEAMLPEEI
>Tat4-1
HEG-AGGNHSGGRSLAAKIKKHQYWPTMITDCIRFVAKCEPCQRHASII-HSPTEALTTS----TAPYPFMRWAMDIVGPL-
PSS-----NGKRFML-IMTDYFTKWVEAEAYNKI-QAGEVQRFVWKNII--CRHGLPYEIITDNGSQFISRQ----------------
FEDFCAKW-RIRLNKSTPRYPQGNGQAEATNKIILDGLKKRL------EA---KKGAWADELDGVL-SHRTTPRRS-T-------
GATPFSLTYGIEAMAPAEV
>KoRV
HQL----THLGPDKLLQLVGRTFHIPNLQSVVREITSKCQVCAVTNAVT-TYREPGRRQ-----RGDRPGVYWEVDFTEVK-
PGR-----YGNRYLL-VFIDTFSGWVEAFPTKTE-TALTVCKKILEEIL--PRFGIPKVLGSDNGPAFVAQV----------------
SQGLATQL-GIDWKLHCAYRPQSSGQVERMNRTIKETLTKLA------LET--GGKDWVTLLPLALLRARNTPGQF---------
GLTPYEILHGGPPPVLASG
>MuLV
HQL----THLSFSKTKALLERNYYMLNRDRTLKDITETCQACAQVNASK-SAVKQGTRV-----RGHRPGTHWEIDFTEVK-
PGL-----YGYKYLL-VFIDTFSGWVEAFPTKKE-TAKVVTKKLLEEIF--PRFGMPQVLGTDNGPAFVSKV----------------
```

```
SQTVADLL-GVDWKLHCAYRPQSSGQVERMNRTIKETLTKLT------LAT--GSRDWVLLLPLALYRARNTPGPH--------
GLTPYEILYGAPPPLVNFP
>FLV
HKL----THLSARKMKTLLEREFYLPNRDLHLRQVTESCRACAQINAGK-IKFGPDVRA-----RGRRPGTHWEVDFTEIK-
PGM-----YGYKYLL-VFIDTFSGWAEAYPAKHE-TAKVVAKKLLEEIF--PRYGIPQVLGSDNGPAFISQV----------------
SQSVATLL-GINWKLHCAYRPQSSGQVERMNRSIKETLTKLT------LET--GSKDWVLLLPLVLYRVRNTPGPH--------
GLTPFEILYGAPPPMAHFF
>REV
HRA----THLGESKLTELVRKHYLICGIYRAARDITTRCVACAQVNPGA-APVEKGLNSR---IRGAAPGEHWEADFTEMI-
TAK-----GGYKYLL-VLVDTFSGWVEAYPAKRE-TSQVVIKHLIHDII--PRFGLPVQIGSDNGPAFVAKV----------------
TQQLCEAL-NVSWKLHCAYRPQSSGQVERMNRTLKETIAKLR------KE---TGGDWVSLLPQALLRARCTPGRE--------
GLSPFEILYGLKPPVVP--
>Xen-1
HGP----THVSKNLMNNLISKLYFAPGITTLTRNYTAACTICAQCNPGR-MEKPPVLNLA----KPLYPFQRIQIDHIQM--PRC--
---GRFEYVL-VVVDMFSGWPEAFPVANM-TAKTTAKKLLSEIV--CRYGVPEVIESDQGPVFTASV----------------
TKDIWTAL-GVTLHFHTPYHPQSSGKVERMNGTLKTKMLKMS------QD---SGMLWPDSLPIALFSVRYTPRGV-N-------
NLSPFEILFGCAPRL----
>WDSV
HGV----SHTHKGGIMSYFSKFWTHPKASQTIDLILGHCQICLKHNPKY-KSRLQGHRP-----LPSRPFAHLQIDFVQMC-------
-VKKPMYAL-VIIDVFSKWPEIIPCNKE-DAKTVCDILMKDII--PRWGLPDQIDSDQGTHFTAKI----------------SQELTHSI-
GVAWKLHCPGHPRSSGIVERTNRTLKSKIIKAQ------EQL--QLSKWTEVLPYVLLEMRATPKKH--------
GLSPHEIVMGRPMKTTYLS
>Kabuki
HNL----SHPGIRTTRKMITNTFFWPNMNKDIGMWAKPCIRCQESKIYR-HTISDLGNF-----VDAGRFEHIHVDIVGPL-PTS--
--PQGFRYCI-TIIDRCTRWPEAFPVHDI-TADKVAKTIYDGWI--TRFGCPVRITSDQGRQFESNL----------------FLKLTKLL-
GINKICTTPYHPQSNGAVERWHRSLKVALMTRL-----------NDTSWVDELSTAMFGLRASVRTD-S-------
GVSAAQLTYGKNIRLPGDF
>FFV
HNI----SHAGREAVLLKIQENYWWPKMKKDISSFLSTCNVCKMVNPLN-LKPISPQAIV----HPTKPFDKFYMDYIGPL-
PPS-----EGYVHVL-VVVDAATGFTWLYPTKAQ-TSKATIKVLNH--L--TGLAIPKVLHSDQGSAFTSEE----------------
FAQWAKER-NIQLEFSTPYHPQSSGKVERKNSEIKKLLTKLL------VG---RPLKWYNLISSVQLALNNTHVVS-T-------
KYTPHQLMFGIDCNLPFAN
>SFV
HEL----AHTGREATLLKLQTTHWWPNMRKDIITVLRQCKPCLQTDSTN-LTPIPPVSQP----RPVKPFDKFFIDYIGPL-PPS---
--HGFSYVL-VIVDARTGFTWLYPTKAP-STNATINSLNL--L--LGTAIPKVLHSDQGSAFTSAT----------------FAEWAKDR-
SIQLEFSTPYHPQSSGMVERKNKEIKRLITKLL------VG---RPTKWYPLLPTVQLALNNTYSPR-T-------
KLTPHKLLFGVDGNVPFAN
>BFV
HDG-PGSAHLGRSALLLKLQQKYWWPRMHIDASRIVLNCTVCAQTNSTN-QKPRPPLVIP----
HDTKPFQVWYMDYIGPL-PPS-----NGYQHAL-VIVDAGTGFTWIYPTKAQ-TANATVKALTH--L--
TGTAVPKVLHSDQGPAFTSSI----------------LADWAKDR-GIQLEHSAPYHPQSSGKVERKNSEIKRLLTKLL------AG---
RPTKWYPLIPIVQLALNNTPNTR-Q-------KYTPHQLMYGADCNLPFEN
>EFV
HKG-PGSGHLGKNTMYIKILNKYWWPNLIKDISKYIRTCTNCIITNTDN-VPNKSYIVQE----KTGLPFQKYYMDYIGPL-
PPS-----DGYYHVL-VIVDEGTGYTWLYPTKAQ-TANATVKALNH--L--TGTAIPKVLHSDQGSAFTSAT----------------
LVAWAKDK-GIQMEYSSPYHPQSSGKVERKNSEIKRLLTKLL------VG---RPTKWYPLIPTVQLALNNTPNAK-I-------
GKTPHQLMYGVDCNLPFQD
>CoEFV
HCG-IGHIHLGGNNTGKTLRRLYWWPGLYTECKRYVANCPECLAVNVTP-TTRPPNLRP-----PRGALFDKIFLDFVGPL-
PRS-----NGYTAIL-ILVESLSSFTWLLPCRDQ-SASTTVTALSA-FF--PASIEPKCFHSDQGGDFTSQL----------------
FKKMCSER-NIRVEYSTPHHPQSAGVVERKNRGLKAALTKLV------RN---RPRKWFQVLDIVQTGLNNTPIAR-N------
HGATPFFLMYGQHMNTPLTN
>ApGIN5c
D-------------------------LKKIGKIMNNCVTCILANRKQ-GKKEGFLHSI-D--KSDTPLAMWHIDFLGPL-TPT----PKGYRHIL-
SVIDGFTKFCWLFPTKSV-TAQEVIEKLTV-ME--ATFGNPGKIVCDRGTAFTSKD----------------FGDYCSGR-
EIRHILITIGVPRANGQVERLNTIIINVLTKLS------LN---KTDQWYRQYQSIKEAV--------------------------------
>Mdg3
HNE---LGHVGRDKMIEAIMKNYWFPNLKQKCSTHISNCLKCISFSPKT-GKTEGFLHNI-P--KGNKPFEIIHIDHYGPV-------
--RPKKHIL-VIVDAFTKFVRLYATKTT-NTKEVIQSLND-YF--RAYSRPKCIISDRGACFTSGD----------------FDSFLKEC-
NVKHIKIATGSPQANGQVERINRSLGPMISKLI------EP---DGLHWDLVLEKVEYTLNNTLHRS-I-------
KQYPSIMLFGLQQKGQIMD
>Tor2
```

```
HDL---KGHLGVDRTAQFLS--CWWPHKMDDIRTYVASCANCLKQKGDM-QPSRPDRKHL-Y--RATRPHEIIYCDFITLP-
TSS----RSGKRHAL-TVMCGFSRWLQVYAVHRC-RSIDAARGLMN-YF--LQFDFPRTLSSDRGRHFENEL----------------
LADLCKLL-QIRQNLHCAYRPESSGVIERCHKTLKSSLWAMV------RD---NRLDWELALPSVVSAMNRSTNAA-T-------
KVSPYKCIFGRDPSFNGIA
>SnRV
HLY----GHPSEESLRKVLTKRFVWEDMGQHCKEITNTCLTCAKYKVLR-AGPPMGVGR-----SAEGPCQKLQVDHVGPL-
GPG----THGYRYLT-TMVDVYTGWFWAKPCRGP-TTGATIAALEE-HI--SIWGVPYSIQSDNGTAFTSKA----------------
MQEWANTY-GIEWKVGAIYHPQSQGKVERKHRLLKDRLKRATH----------EGKNWVQALPSILLFINSHPRDQ----------
FSAYELMTGRVPHL----
>ApGIN4c
HDY---YAHTGPLKTAHALKNICYFPSFNKTIRCIVQACELCQKCKPKT-TRIAGPLQPI----LSNQPLDKLLVDFYGPL-PTG-
---IFQFSYIF-VVVDNFTRFVKLYPLQSA-NAKICIKKLTT-YF--PNYGIPKNIVSDHGRQFISNY----------------WQTSLKKY-
KVQVSHTSVYHPQSNP-AERVMRELGRMFRTYC------HE---QHSLWPQYVPYIEWTLNNVRHES-T-------HNTP------------
--
>Woot
HQL---LGHFGATKVYNSMKREYYWPNMYRTIKKRLRSCDLCQKTKSSN-RPHQGPLTPI----LYDHIGDLVCVDFYGPL-
PTG----RLGASYVF-VVIDVFSKFLKLYPLRKA-TAKIAAKRLIE-DF--SGYIKPKCVLSDHGTQFISNT----------------
WQNSLRAA-DIQPTLSSIRHPESNP-SERVMRELGRIFRAYC------RE---NHASWVNHLSNIEDCLNYVPHIS-T-------
GFSPYEILYGRTPPNPLDA
>ApGIN3c
-----------MVEYESIARSYFWWPSIDKEIEDMARNCNECINARPNP---PKSVLTPW-K--WPQRQWTRVHCDFLGPY-----------
NKTFL-IIVDATTKWLEVFQVNSM-TAQTVITKLSE-LI--ARFGIPRTITTDNAKCFNGDE----------------FQSYCKTL-
GIRHLTGAPFHPASNGFAESGVKIVKQFFNKCS------TS--------LHHLNKFLLMYRNTVHST-T-------NETP--------------
>Mag
HDP-----HMGIVKTKSLARSYVWWPGID-EAETECRACTVCAAVADAP---STHAPRSW-P--WPSRPWSRLHLDFLGPI-----
-----GVTYL-VVVDSCSKWIEAIKMQRT-TAQAVISVLRD-LW--SKFGLPKQTVSDNGPPFSSSD----------------FQKFLIHN-
GIKHIYSAPYHPASNGAAENAVKICKRAIKKALK----------QNLNVDTALCRFLLAYRNTEHAT-T-------
GDSPANILQGRSLRMRLDN
>SPM
HQE-----HTGIVRMKAVARSYFWYPKLDADIEKLSQSCEACLRMRNDP---PKVPFVPW-S--NARKPWERVHVDFFELE----
-------GKDYL-VLVDTYSKWVNVELMTST-TSAKTIETLRS-WF--AIFGLPETLVSDNGPQFTSEE----------------
MEVFLSKN-GVKHVLVPPYHPASNGAAERTVQIVKRTLQKYF------LS---DHVSIRHRLDNFLFSYRTTPQTV-T-------
GVSPYELVFKFAPRTRFSL
>Cer4
HE-----GHPGVVRMKQKACSFVFWTGIDKDVEKLVRGCENCQESAKMP---RVAPLRPW-P--EPQKAWSRVHIDFAGPV---
--------GHWFL-VIVDAKSKYAEVKMTKTI-SASATVSLLEE-VF--ATHGYPELLVSDNGTQFTSNQ----------------
FKLMCQEY-GMEHKTSAVYYPRSNGAAERFVDSLKRGLAKITR----------SGVVTQQALNKFLICYRNTPHSALA-------
GATPAECHFGRKIRTKMSL
>SURL
-------GHLGIEKCKKRAREVIYWPRINADIAEMVQSCTSCLMYKPKQ---QAESLHPH-A--VPSRPWEKVAVDLF----------
TLNKREYM-VIVDYYSQFIEVCTMTST-SSKAVINHMKA-IF--ARHGTPCELMSDNGPQFASQE----------------
FKSFAKEW-DFHHTTSSPHYPQSNGLAENAVKIVKNLLLKS-------------QHCGQDIHRALQVYRSSPLAC---------
GKSPAELLYNRRLRSNLPM
>Tor1
HLP----GHARTTNMLGSMRDTVIFPKMAKKIKQKIMTCPTCLATGSTE-RPRFTPITPP----KESHPYMTVTVDLLGPL-PST-
---PTRKKYIL-AAIDNLTRWIELRCIPDK-TAAHVAKALMD-IF--FLRGPPRAVSCDNGREFSNTL----------------
LRELLASF-GTCINYGTPYRPQGQGLIERANREIVKHLKGLNI----------EEHKWDTFIPSIQLSMNLTYHSA-L-------
GSSPFQAMHGWTLAEPLFS
>MmCGIN1
HDS-PIGEHQGLEDTYKTVRLLGWWPGMQDHVRDYCRSCLFCIPRNLIG-GELKVIESPW-PL-RSTAPWSSLQIEVVGPV-
TVS----EEGHKHVL-IVADANTRWVEAFPLKPY-THVAVAQVLLQ-VF--ARWGVPIRLEAAQGPQFARHV----------------
LVSCGLAL-GAQVTTLSRALQFPCLMSSEAYWEFKRALKEFIFL---------YGKKWAASLPLLHLAFRASTTEA-----------TP-
--------------
>HsCGIN1
HDI-PLGAHQRPEETYKKLRLLGWWPGMQEHVKDYCRSCLFCIPRNLIG-SELKVIESPW-PL-RSTAPWSNLQIEVVGPV-
TIS----EEGHKHVL-IVADPNTRWVEAFPLKPY-THTAVAQVLLQ-VF--ARWGVPVRLEAAQGPQFARHV----------------
LVSCGLAL-GAQVASLSRDLQFPCLTSSGAYWEFKRALKEFIFL---------HGKKWAASLPLLHLAFRASSTDA-----------TP--
-------------
>ZFERV-2
-----------------------------------CPNCRKAIDSR-FNDYGSI-------KTDHENDEISMDFAGPIRPQT----KKKNCYFL-
LITDNASEYTQIFPCKQA-TEEVVIESLQQ-WF--KFRGVPNRVRADGAFNLSGRK----------------VEQFIKDS-
GAKLIKSVKYQPSSNGIAERRIKEVKNYFCKN-----------SDLPWDEQLIDCLKFVNY--------------------------------
```

```
>AciFLERV
HEA---LGHAGADPTRRELKKQLWIP--VREIRRVLKDCVECGRYNAGR-RGQRMGGLTI----KSTVPWGSICMDVAGPL-
GIT---GMKGEKYLL-VLVDSMSGYVAVKPVRQA-NGSSVVSMLDQ-AC--CYMGIPKELRTDNGTHFRNAK----------------
VDQWCQEN-GVMRVYSPPYTPQANGVVERTIGLVKSWIGKNA-----------NGKEWSTKVLELVKTLNDSPAEE-L-------
NQRPFTSIQGVEQLKPVPD
>AliFLERV
HEA---LGHAGLLPTRRGLVKQLWIP--MKQIQQVLKRCTVCGKYNAGR-RGKRTEGLTI----KSTIPWGSICMDVAGPM-
GVT---GTKGEKYLI-VLVDSMSGFVVVRPVRKA-NGNSVVSMLDY-TC--ACLGIPRELRTDNGTHFKNVQ----------------
VNRWCEQH-GVMRVYSPPYTPQANGVAECTIGLVKNWLAKNA-----------NSKEWSLKSVDIGRALNDSPSDE-L-------
NQRPFVTIQGVEQLKPALG
>Cigr-1
HNN---LGHPGFSRLYHFIKTRNL-PFSSEEAKIICRQCKTCAEVKPRF-YKPASRNLI-----KAMRPWERLSMDFKGPV--------
KGRNSYLL-IIVDEYSRFPFVFPCKRM-TSAVVIQCLTN-LF--CLFGFPAYVHSDRGLSFISHE-----------------VKVFLTER-
GIATSKSTPYHPQGNSQCERVNQTIWRTVKLLL------HGK--LEEQWEEVLPQALHAIRSLVCLA-T-------
NETPHERFFNFSRRS----
>JSRV
HDL----HHQNSHSL------RLQFKISREAARQIVKSCSTCPQFFVLP----QYGVNP-----RGLRPNHLWQTDVTHIP-------
QFGRLKYVH-VSIDTFSNFLMASLHTGE-STRHCIQHLLF-CF--STSGIPQTLKTDNGPGYTSRS-----------------FQRFCLSF-
QIHHKTGIPYNPQGQGIVERAHQRIKHQLLKQK------KGN--ESPSPHNALNHALYVLNFLTLDT-E-------
GNSAAQRFWGERSSC----
>SRV-1
HTL----HHLNAQTL------KLMFNIPREQARQIVRQCPICATYLPVP----HLGVNP-----RGLLPNMIWQMDVTHYS-------
EFGNLKYIH-VSIDTFSGFLLATLQTGE-TTKHVITHLLH-CF--SIIGLPKQIKTDNGPGYTSKN-----------------FQEFCSTL-
QIKHVTGIPYNPQGQGIVERAHLSLKTTIEKIK------KGE--WKGTPRNILNHALFILNFLNLDD-Q-------
NHSAADRFWHSNPRK----
>MMTV
HAL----HHQNAAAL------RFQFHITREQAREIVKLCPNCPDWGHAP----QLGVNP-----RGLKPRVLWQMDVTHVS-------
EFGKLKYVH-VTVDTYSHFTFATARTGE-ATKDVLQHLAQ-SF--AYMGIPQKIKTDNAPAYVSRS-----------------
IQEFLARW-KISHVTGIPYNPQGQAIVERTHQNIKAQLNKLQ------KAG--KYYTPHHLLAHALFVLNHVNMDN-Q-------
GHTAAERHWGPISAD----
>RSV
HTA----LHIGPRALSKAC------NISMQQAREVVQTCPHCNSAPALE-----AGVNP-----RGLGPLQIWQTDFTLEP-------
RMAPRSWLA-VTVDTASSAIVVTQHGRV-TSVAVQHHWAT-AI--AVLGRPKAIKTDNGSCFTSKS-----------------
TREWLARW-GIAHTTGIPGNSQGQAMVERANRLLKDRIRVLA------EGD--GTSKQGELLAKAMYALNHFERGE-N-------
TKTPIQKHWRPTVLT----
>LDV
HAF----LHLGAKALAQEC------NIPLTEVERLVKSCSHCQKTPLWM-----AGVNP-----RGLEPNQIWQTDFTMYP-------
PFRPNHWLA-VTVDTYSGVIVATAHRGT-KSRHAIHHWTA-AI--AYLGRPSVIKTDNGSAFTSKA-----------------
ATAWAQRW-HITLKHGIAYNSQGQAIVERANRTLKARLEALG------EGE--GPEHQPELLYRALYSMNFFARST---------
EQTSVARHWSHQRLP----
>BLV
HKL----THCN---------SRSRWPNSAWDPRSPATLCETCQRLNPTG-GGKMRTIQ------RGWAPNHIWQADITHYK--------
YKQFTYALHVFVDTYSGATHASAKRGL-TTQMTIEGLLE-AI--VHLGRPKKLNTDQGANYTSKT-----------------
FVRFCQQF-GISLSHHVPYNPTSSGLVERTNGLLKLLLSKYH------LDE--PHLPMTQALSRALWTHNQINLLP-I-------
LKTRWELHHSP--------
>HTLV-2
HGL----THCN---------QRVSFGATPREAKSLVQTCHTCQTINSQH-HMPRGYIR------RGLLPNHIWQGDVTHYK--------
YKKYKYCLHVWVDTFSGAVSVSCKKKE-TSCETISAVLQ-AI--SLLGKPLHINTDNGPAFLSQE-----------------FQEFCTSY-
RIKHSTHIPYNPTSSGLVERTNGVIKNLLNKYL------LDC--PNLPLDNAIHKALWTLNQLNVMN-S-------GKTRWQIHHSP-
-------
>HERV-L
HEQ---SGHGGRDGGYWAQNQQHGLPLTKADLATATAECPICQQQRPTL-SPRYGTI----P--QGDQPATWWQVDYIGPL-
PSW-----KGQRFVL-TGIDTYSGYGFAYPACNA-SAKTTIHGLTE-LI--HHHGIPHSIASDQGTHFTAKE-----------------
VWQWASVQ-GIHWSYHVPHHPEAAGLIERWNGLLKSQLQCQL------GDN--TLQGWGKVLQKAVYALNQHPIYG---------
TVSPIARIHGSQTMP----
>MuERV-L
HEQ---SGHGGRDGGYPWA-QQHGLPLTKADLATAAADCQICQQQKPTL-SPRYGTI----P--RGDQPATWWQVDYVGPL-
PSW-----KGQRFVL-TGVDTYSGYGFAFPARNA-SAKTTIHGLTE-LI--YRHGIPHSIASDQGTHFTARE-----------------
VRQWAHDH-GIHWSYHVPHHPEAAGLIERWNGLLKMQLQHQL------GGN--SLEGWGRVLQKAVYALNQRSIYG---------
TVSPIARIHGS--------
>HIV-1
```

```
HEK----YHSNW----RAMASDFNLPPV--VAKEIVASCDKCQLKGEQV----------------DCSPGIWQLDCT----------
HLEGKVIL-VAVHVASGYIEAEVIPAE-TGQETAYFLLK-LA--GRWPV-KTIHTDNGSNFTGAT----------------
VRAACWWA-GIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRD----------QAEHLKTAVQMAVFIHNFKRKGGIG-------
GYSAGERIVDIIATD----
>FIV
HEK----FHSDP----QYLRTEFNLPKM--VAEEIRRKCPVCRIRGEQL----------------KIGPGIWQMDCT----------HFDGKIIL-
VGIHVESGYIWAQIISQE-TADCTVKAVLQ-LL--SAHNV-TELQTDNGPNFKNQK----------------MEGVLNYM-
GVKHKFGIPGNPQSQALVENVNHTLKVWIRKFLP----------ETTSLDNALSLAVHSLNFKRRGRIG-------
GMAPYELLAQQESLR----
>EIAV
HEN----WHTSP----KILARNYKIPLT--VAKQITQECPHCTKQGSCV----------------MRSPNHWQADCT----------HLDNKIIL-
TFVESNSGYIHATLLSKE-NALCTSLAILERLF------SPKSLHTDNGTNFVAEP----------------VVNLLKFL-
KIAHTTGIPYHPESQGIVERANRTLKEKIQSHRD----------NTQTLEAALQLALITCNKGRESM-G-------
GQTPWEVFITNQAQV----
>Gypsy
HNR-----HRAAQENIKQVLRDYYFPKMGSLAKEVVANCRVCTQAKYDR-HPKKQELGET-P--IPSYTGEMVHIDIF----------
-STDRKLFL-TCIDKFSKYAIVQPVVSR-TIVDITAPLLQ-II--NLFPNIKTVYCDNEPAFNSET----------------VTSMLKNS-
GIDIVNAPPLHSSSNGQVERFHSTLAEIARCLK------LDK--KTNDTVELILRATIEYNKTVHSV-T-------RERPIEVVHP-------
--
>Zam
NSN-----HRGIDETFLHLKRETYFPNMKNKISELIRNCETCLKLKYDR-QPQNIVFETP-E--TPSKPLDIIHIDIY----------
TINNNFNL-TIIDKFSKFAAVYPIPNR-NGINCIKAIKN-FF--SQFGLPKKLIHDQGVEFCNDI----------------FRKFCSQY-
NILLHVTSFQQSSSNSPVERLHSSLTEIYRILL------DTRK-KPTDHEEIMSETVITYNNAIHST-T-------
KHTPFELFNGRTHLF----
>Tom
HEK-----HPGIQKMTKLFKENHYFPNSQLLIQNIINECRVCNLAKTEH-RNTKMPFKVT-P--SPGHCRDKFVIDIY----------
SSEGKHYL-SCIDIYSKFATLEQIKTK-DWIECKNALMR-IF--NQLGKPTLLKADRDGAFSSLA----------------LKQWLESE-
GVELQLNTAKTGVAD--VERLHKTINEKIRIIN------SSK--NDEIKLGKMENILYIYNHTRHDT-T-------
GQTPAHIFLYAGQPT----
>Tor4a
--------HTGIDQTFSKIAKKYVWPRMNMAIREFIRTCDFCQKDKPNF-HPNRAPVLSF-R--TPNGPYEIYGFDLITL--PPT----
DFGNKYVM-VMIDFFSKFAYCEPLKSR-NSSYLLSKFRN-IF--KNPFFPKMVVLDNAREHSE------------------LAKFMTNN-
NIEPHFTPPRHPSSNGQVENFNRTLKSRLRAKC-----------KYENWDLVLQEILHDINASEHSV-T-------
KKSPFFIQSGAECPHNIFD
>DpCIN1
F-------------------------------------------------------------VQKGINDTWQIDLVEMI-PFYK---EKGYRYLL-
TVIDTFSKFAYAVPVKSK-TALDVLSALKK-IL--KSKHKPKNIQSDQGKEFFNVA----------------FGELMKKI-NINHYHT--
YTHLKASICERFNRTLKNRMWKIF------SMQ--GNHTWIYILHDLLKDYNNSYHRT-I-------KMRP--------------
>DrCIN1
V-------------------------------------------------------------VVYGMDTQFQADLVDMT-AYSTE--NDGNKFML-T-
IDVFSKYAWSRVLKNK-SGVEVTKAFES-IL--EEGRVPRKLQTDQGKEFFNKH----------------FQDLMKKH-DINHFAT--
ATDLKASVVERFNRTLKSRMWRFL------TAT--NSRRYINVLQDIMTGYNNSYHKT-I-------KMRP--------------
>SmCIN1
Y-------------------------------------------------------------KTMGVDDLWQMDLMEMI-PFSKI--NKGYKYIL-
TCVDIFSRFARAIPTKTK-SANEMAEAIKE-LF---IDGKPRNLQTDLGKEFYNSK----------------VREIIK---GINHYSV--
FSQFKAAHVERFNRTLRDRLKKYF------VYK--GNKTWINVLQKAIYSYNYSPHRGLN-------GMRP--------------
>ApGINGER2a
YDTHSAIGHGGRNRMVSELKNNYC-NITNETIMVFLKLCAECHKKTVCS----KVGLAPK-PF-LQPTFNSRAQVDLIDMQ-
SQR----HNDYRFIM-NYQEHLTKYIVLKPLKSK-CAEEIAYNLID-IY--TLFGAPTILQSCSGKKFVTSI----------------
INKLNVIWDEIKIDYGWPKHSESQESVERTNRDVQEMLSAWM------TKK--DSLDWPSALKFIQFKKNRTFHSG-I-------
GTSPYEAMFGCTARIGYES
>BmGINGER2a
HDAHLATGHGGRNRLVKGINNKYK-NITTEAVMLYLSLCEACQKKGK------KKGIAIK-P--TTAEMNSRCQMDLIDLQ-
TSP----DNEFKFIS-NYQDNL-NFVVLRPLKTK-TASEVVKQLMD-IF--TLLGAPSS-NVINDREFANKI----------------
IQELANMRDGVKIIHGKPRHSQSQGSIERANQDVQNMITTWL------ITN--QTTKWS--------------------------------------------
>FcSCAND3
HNTHLSIGHGGRTRMEKELQAKYK-NITKEVIMLYLTLCGPCQQKHSKL----KKVLTSK----PIKEVSSRCQVDLIDMQ-
LNP----DGEYKFIM-HYQDLRTKLSFLRSLKSK-RPKEVAHALLD-IF--TIIGAPSVLQSDNGREFSSQI----------------
VYELSNIWPELKIVHGKPQPCQSQSSMNQTNDDIQKRIISWM------QTN--NSSHWAEFLWFIQMIQNQPYHRG-M-------
QPNPFEGTFSSEAKPGLCH
>HsSCAND3
```

```
HNTHLSIGHGGRTRMEKELQAKYK-NITKEVIMLYLTLCKPCQQKNSKL----KKVLTSK----SIKEVSSRCQVDLIDMQ-
LNP----DGEYRFIL-HYQDLCTKLTFLRSLKSK-RPTEVAHALLD-IF--TIIGAPSVLQSDNGREFSSQV----------------
VSELSNIWPELKIVHGKSQTCQSQSSAEQT-EDIRKRIFSWM------QTN--NSSHWTEFLWFIQMSQNQPYHRS-M-------
QQTPCESAFSSEAKLGLSH
>MeSCAND3
HDTHLNIRHGGRTRMEKELQPKYH-KRGYNVI--FNSLQTLSAKEFKVT----KKCIVSK----PIKEVNSRCQVELIDMQ-LNP-
---DGEYKFIM-SYQDHQTKLSFLRPLKSR-RCEEVAHALLD-IF--TIIGAPSILQSEHGRDILSQL----------------
VNELNMMWSKLKIVHGKPRPCQGHGSVVQPKEDIQNMIISWM------QKN--SSLHWSEGLWFIQMMKNRPFHRD-T-------
QHSPREIVFDSEKST----
>HsKRBA2
HDTHLSIGHGGRTRMLKELQGKYG-NVTKEVIVLYLTLCKQCHQKNPVP----KRGLAPK-PM-TFKDIDSTCQVEILDMQ-
SSA----DGEFKFIL-YYQDHSTKFIILRPLRTK-QAHEVVSVLLD-IF--TILGTPSVLDSDSGVEFTNQV----------------
VHELNELWPDLKIVSGKYHPGQSQGSLEGASRDVKNMISTWM------QSN--HSCHWAKGLRFMQMVRNQAFDVS-L-----
--QQSPFEAMFGYKAKFGLYS
>BtKRBA2
HDTHLNIGHGGRTRMLKELQGKYG-NVTKEVIVLYLTLCKQCHQKNPVP----KRGLAPK-PM-
PCKDNDSRCQVEILDMQ-SNA----DGEFKFIL-YYQDHLTKFIILRPLKAK-QAHEVVSVLLD-IF--
TILGPPSVLESDSGMEFTNQV----------------VNELNEVWPDLKIVLGKSHPGQGQGFLERVSRDVKTMLSAWM------QSN-
-HSHHWTEGLRFMQMVRNQAFDGS-L-------QQSPYEAMFGCKAKFGLYS
>LaKRBA2
HETHLNIGHGGWTCMLKELQGKYG-NVTKEVIVLYLTLYKQCHQKNPVP----KKGLVPN-ST-
TLKDVNSRCQVEILDVQ-SNA----NAEFKFIL-CYQDHLTTFIILXPLKAR-QVHEVVSVLLD-IF--
IILGTPRVLESNSGIEFTNQ-----------------VNEVSEVWPDLNIVSSQYHTGXGQGSKEQXNEDVQNTMSARM------QSN--
HSHHWTEVLXFMPMVGNQVFNVS-L-------HQS-----LQCKVTLGLYS
>HmGINGER2-1
CQN---GLHLGALKTYKKIIEGYA-NIARKAVEIFISFCPTCNLNKRQL---KKAPLQPI----ISTGFLQRLQIDLIAME-SKP----
DKEFRYIG-HVVDHFSKFHILFPMRNK-TALETANNLKR-VL--SYFGLPAILHSDNGSEFVNDV----------------
IKSLLLTWPKTQIVNGSPRHSQSQGLVEQGNFTIERLISSRE------YDS--GQSNWSEWLVEIQCKFNV--------------------------
------
>Zeco1
HKQ---FGHASADRLQRLMHSSCF-----AILQQIVDDCEICKKYKRTK-PKPAVGL----P--MASEYNETMAVDLHELE-PSI----
-----WYL-HIIDHFTRFSAGSIVKTK-KSSEIVNSFIH-WI--SVHGAPQKLYSDNGGEFNNQE----------------IRDMAENF-
NIEIRTTAGYSPWSNGLLERHNQTLTEIILKV-------KE---NGCDWHTALDWALMAKNSMLNVH---------GYSPYQLVFG---
------
>1731
HKR---NGHLNTSSLQEMVRKKMVY-GVEKVVFKPDAVCKTCMLAKIHV-QPFPKTTRS-----
RAEELLDMIHSDLCGPFSTPS----LAGSKYFL-TFIDDKSRRIFVYFLRKKDEVFTKFVEFKK-
LVERQTGRKIKCIRSDNGGEFVNNV----------------FDDYLKAH-GIARQLTIPHTPQQNGVAERANRTLVEMARCML------
LQSELGEALWAEAINTAVYLRNRSTSRALQ-------SKTPMEEWTG---------
>Tnt-1
HKR---MGHMSEKGLQILAKKSISY-----AKGTTVKPCDYCLFGKQHR-VSFQTSSE------RKLNILDLVYSDVCGPMEIES---
-MGGNKYFV-TFIDDASRKLWVYILKTKDQVFQVFQKFHA-LVERETGRKLKRLRSDNGGEYTSRE----------------
FEEYCSSH-GIRHEKTVPGTPQHNGVAERMNRTIVEKVRSML------RMAKLPKSFWGEAVQTACYLINRSPSVPLA-------
FEIPERVWTN---------
>Vitico1-1
HLR---YGHLNVKGLKLLSKKEMVFGL---PKIDSVNVCEGCIYGKQSK-KPFPKGRSR-----RASSCLEIIHADLCGPMQIAS--
--FGGSRYFL-LFTDDHSRMSWVYFLQSK-----------------------VLRTDRGGEFLSND----------------FKVFCEEE-
GLHRELTTPYSPEQNGVVERKNRTVVEMARSMM------KAKNLSNHFWAEGVATAVYLLNISPTKAVL-------
NRTPYEAWYG---------
>SIRE1-4
HQR---FGHLHLRGMKKIIDKGAV-RGIPNLKIEEGRICGECQIGKQVK-MSHQKLQHQ-----
TTSRVLELLHMDLMGPMQVES----LGGKRYAY-VVVDDFSRFTWVKFIREKSETFEVFKELSL-
RLQREKDCVIKRIRSDHGREFENSR----------------LTEFCTSE-GITHEFSAAITPQQNGIVERKNRTLQEAARVML------
HAKELPYNLWAEAMNTACYIHNRVTLRRGT-------PTTLYEIWKG---------
>Copia
HER---FGHISDGKLLEIKRKNMFSDQLLNNLELSCEICEPCLNGKQAR-LPFKQLKDKT----
HIKRPLFVVHSDVCGPITPVT----LDDKNYFV-IFVDQFTHYCVTYLIKYKSDVFSMFQDFVA-
KSEAHFNLKVVYLYIDNGREYLSNE----------------MRQFCVKK-GISYHLTVPHTPQLNGVSERMIRTITEKARTMV------
SGAKLDKSFWGEAVLTATYLINRIPSRALS-------SKTPYEMWHN---------
>pCretro6
HRI---LGHAHLASIKRLCLGLA--EGVDVDMTSDDQFCESCVRAKQHV-QPFPDQATRT-P--
DKLKVGEIVASDTWGPAQVRS----VHSFYYYM-SMTDLKSRFSGVYFSAQKSLAKMALTEFRG-
```

```
LIRRRTGSSIMCLRIDNGTEFINNE----------------FLQYCKSQ-GIVVETTAPYSPAQNGVAERLNRTLRESARAML------
LAAELPRKFWPEAVSYACLIKNRLPHAALK-------GMTPYQALTG---------
>Hydra1-2
HEI---LGHCNYEKLQNVVK-----GMITGKIDKSNLNCKICTQGKFVQ-SRNREPDT------RAKAALELVHTDLAGPIDPEA---
-KDGFKYAL-AFTDDYSGAVFVYFLKAKSDTARATEKFIA-DT--APYGRIKCVRSDNGTEFTAKE----------------
FQSLLSKN-GIRHETSAPYSPHQNGTAERNWRTLFEMARCMI------LESGLPKMLWTYAVMTAAVIRNRCYNSRIH-------
QTPYYALTG---------
>Koala
HSR---LGHPAFPIVKRIVQSHKL--SCLDVESNNISVCDACQKAKSHQ-LPFGLSTS------EVHSPFELVYSDVWGPA-PTS----
VGGKKYYV-SFIDAYSRFTWIYLLKFKSEVFEKFHEFQN-HVERFFDRKIKTIQTDWGGEY--QK----------------LNSFFNKI-
GILHHVSCPHTHQQNGSAERKHRHIVEVGLSLL------AHAFIPLKFWDEAFSSAVYLINRIPTKVLQ-------YRSPLEQLYN--
-------
>CoDi6.3
HCA---LGHPCEATTRATAKAFV-------RLIGQMKPCKDCALSKAKA-KKISKVPVK-----RASKPGGRLCIDISSPS-TKS----
VGGKCHWL-LVVDDCTDYAWSFFLNKKSETNDIMIALIK-LK--QYDIDVKTIRCDNSGE--NNA----------------
LQRSCKQE-GITFEYTAPNTPQQNGRVERRFPTLYGRVRAML------RDVSVSKRLWAEAANTATDLDNMLLKQG-------
ETTNSFHKFFG---------
>Ty1B
HRM---LAHANAQTIRYSLKNNTYFNESDVDWSSAIDYCPDCLIGKSTK-HRHIKGSRLKYQ--
NSYEPFQYLHTDIFGPVHNLP----KSAPSYFI-SFTDETTKFRWVYPLHDR-REDSILDVFTT-IL--
AFQASVLVIQMDRGSEYTNRT----------------LHKFLEKN-GITPCYTTTADSRAHGVAERLNRTLLDDCRTQL------
QCSGLPNHLWFSAIEFSTIVRNSLASPK-S-------KKSARQ------------
>Ty4
HKR---MGHTGIQQIENSIKHNHY--EESDLIKEPNEFCQTCKISKATK-RNHYTGSMNNHS--TDHEPGSSWCMDIFGPV-
SSS----ADTKRYML-IMVDNNTRYCMTSTHFNK-NAETILAQVRK-NI--QFDRKVREINSDRGTEFTNDQ----------------
IEEYFISK-GIHHILTSTQDHAANGRAERYIRTIITDATTLL------RQSNLRVKFWEYAVTSATNIRNCLEHKS-T-------
GKLPLKA------------
>ISSc
W----------------------------------------------------------------QVERPGALWHGDVCHVT-GCT----VGGPLRIH-
GLLDDASRYVVALEAHTTEKEIDMLAMTVD-AL--RRHGKPDALYLDNGSTYRGDV----------------LKTACARL-
GITLLHAKPYDPEARGKMERFWRTLREGCLTYL------GAVAINTTLRA----FLDRRYHPAPHAGLL-------GQTP-------------
-
>ISAf
F---------------------------------------------------------------AYDQINELWQGDLSHGP-TIR----VNGKTFLI-
AYIDDCSRFVPYAQFFPS-EKFDGLRIVTK-AV--LRCGKPKRIYSDNGKIYRSEV----------------LQYACAEM-
GITLIHTQPYDPQSKGKIERFFRTVQTRFYPLLELDPPKSLEELNERFWR----WLEEEYHRKPHASLD-------GKTP-----------
---
>ISLa
FER---------------------------------------------------------------SRPNQMWQTDIFSFR--------LGGAAYLL-
AFIDDYSRYMVGLGLYRRQTAENLLEVYRR-AT--GEYNCPAEMLTDNGRQYGTTR----------------FEKELKKD-
RIKHIRSQPHHPMTLGKIERFWKTIWTEFLDRCQFD--------CMETAQQRITLWIKYYNHRPHQGIG-------GLCP-------------
-
>ISTv1
Y----------------------------------------------------------------WALSPNLIWHADVHYFR--------GVRGQYIY-
GIIDDFSRKVLACIQIPDMLAATTANVAAS-AF--IEFGAPYCFWTDNGSE---NA----------------FQDLLDLW-
GVQWRHTDSHSPYQNGKIERFWPTLERCQSIQ------------------AIPAFIWEYNNTPHEDLT-------GTTP--------------
>ISEc
F---------------------------------------------------------------YASGPNQKWAGDITYLR--------DEGWLYLA-
VVIDLWSRAVIGWSMSPRMTAQLACDALQM-AL--WRRKRPRNVHTDRGGQYCSAD----------------YQAQLKRH-
NLRGSMSAKGCCYDNACVESFFHSLKVECHFIS-----------REIMRATVFNYIECDYNRRRHSWCG-------GLSP--------------
>CoDi6.7
HHR---LGHASMTKIRMLAKVGLL-PASLKECQ--IPLCTSCLYGKATR-RPWRTKGSSTAS--
KVNGPGQCISVDQLVSTTPGY----LRGYHAAT-VFVDNFSRLSYVHVQKGT-SAEETIQAKHA-YA--
RSHGVVTKYHADNGI-FADNK----------------FREAVKED-RQTLSFCGVNAHFQNGIAERRIRELQDHARTML------IHA--
TSHLWPYALRMANELHNHLPSLR-G-------SVPPIEIFA----------
>ApBIN1j
HIR---LLYIGPQGLLAHIHRKYWPIRGRSLARQTVHKCMICFRFKPRM-LQPVMAPLPSVRD-TRCRAFERAGVDLCGPI-
NIRSGL-RKVKHYIA-VFVCMVTRAVIWSLFEIC-------LRMRS-FL--CYSGA-KKILSSWASDLKESK----------------
LNDQLSEL-GIEWQFILPSAPHSGGLWESAVKSAKQHLVKSS------
NGSLLTYEETSTLLCRNEAVLNSRPLTATSSDPSDFNVLTP--------------
>Tas
```

```
HEE---LLHAGISSILAKMREAYWIPRGRQAIKRALNNCFHCRRWKSKPFQMPPMPPYPPERV-SKHPPFENTGVDYIGPF-
TIRSMK-RITQNDGS-VYSLVSQLELYISKSPLTYPDQVSFSVYVR-FV--SRRGLPKRMLSDNGTQFVWAR---
SVLTSVSKHQTDNAILDYCAAH-NIQWSFITPLVPWQGGIYERMVGPVKGSMKKTI------
GWKRLTQEELQTLTTDIEAVVNCRPIIPLTSE--NTTVLRPVDFLLPHGNSSP---
>Cer7
HGK----YHTNEQQTMEATREKAWIPCLRRQVKKIIGKCVKCQRYNRAPMKYPNMADMPSFRV-
RRSRPFENTGLDYFGPM-TFRKEDGSTESCWGC-VLSCATTRLTHIELVQQC-STKAFINAIRR-FV--
SERGIPDRIVSDNAPQFLGQQILEVSARASKENALDKDILEYLGNT-SIEWSFITPYAPWQGGMYERMMRSIKQAIYKSI---
---GKNILTLDDLETVMKEVSGIINSRPLTYVTEGAG--TTICPKDFINPEMRMT----
>Kobel
HST---VGHLGRNSILAKVREKYWIYGASQLAKRVARSCVTCQRYHAAP-
CEQVMAELPADRVVAEVSPFTFCGMDYFAPL-TVRRGR-SEVKRYGV-IFTCFSSRAVHLEIAHSL-ETDACINAIRR-FM--
ARRGPVRSIRSDNGTNLVGSE---KELRHALEGLDQSRMNDVLCAE-GIEWHFNPPAASHFGGVWERMIRSIRKILYSLL---
---REQPVDDETLSTLFCEAENIVNNRPLTTTSSDPNDLLPLTPNM------------
>Tamy
HRR---FCHANHATVVNELRQTYWILSLRDAVKKVLHQCQWCRTRKMKP-
QMPPTGDLPVERLRYGSHPFTCTAVDYFGPM-FVTIGR-RKEKRWGA-LFTCLTTRAVHLELVPSL-STSSMIMALRR-MS-
-ARRGTPRVIYSDNGTNFIGAN---HELREEIGKLKKNELIDAANQE-GIRWKFIPPGAPNMGGAWERMVRTVKTALSAIL--
----NERSPPEEVLHTLLTEVEHTVNSRPLTHLSVNPEDEESLTPNHFLIGRSCGS----
>Zebel
QR----LLHPGSERVLAELRRQYWVLRGREAIRKHQHTCRDCQFWRAKP-
QTPQMADLPSSRLQLYKPPFYSTGVDCFGPF-TVKVGR-RQEKRWGV-LYKCMTTRCVHLDLLEQL-DTDAFLLSLRR-
FI--ARRGKPMELLCDNGTNFVGGD---RELRETFNAMAPK-LQEQLAEQ-
RIRFRFNPPSAPHFGGTWEREVKSVKSALRVIL------
REQSVPEAVLQTLLVEVEGILNSKPLGYISSDVADLDPVTPNLLLMGRRDAS----
>Purbel
HE----LGHHGVAVTTAKTRRKYWILQGYRLAKTVKHRCVTCRAAECRR-
ETQIMANLPSCRLAPFTPPFHYTSCDYFGPY-LVKVGRNKKAKHYGI-IFTCLNTRAVHLEMATDC-STMEFLQALRR-FI--
AVRGQPAQFLSDNGTQFVGAE---RELREMVRGWSERELKDFCAEK-
RVVWKFVTPGAPHQNGCAEAMVKSCKFALKRAI------
GEQVLTPLELYTCFLEVANLVNERPIGRVGNDPDDGGYLCPNDLLLGRSSSK----
>ApBIN1h
------LGCIAPVPLTQDINKLYYLRQVRRATGAAVNGSAPCVSASATP--------------------LRNRVDYAGPL-
QMRELRLRVFKIYIA-VFVCFTTKAVHLEVVSDL-STDAFFAAFDR-FV--ARRGLPNEVFSDCGTNFVGAD---
RQMRSLINSAEGQAIGNARA---TCDWHFNPPSAPHFGGLWEVAVRSTKRLLVRVI------
GNHIFTYEEFSTILARVEAVLNSRPLTPASTDPHDLDCLT----------------
>Bel
HVS---YLHTGVDATFTNLRQQYWILGARNLVRKAVFQCKSCFLQRKGT-SNQIMGELPIPRV-QASRCFQHTGLDYAGPI-
AIKESKGRIGKAWFS-IFVCLTTKALHIEVVSEL-TTQAFIAAFQR-FI--
ARRAKPTDLYSDNGTTFHGGKKTDDMRRLAIQQAKDEELAGFFANE-
GISWHFIPPSAPHFGGMWEAGVRSIKLHMKRIL------GSKALTFEELSTVLTQIEAILNSRPLCPTGDN--
SLDPLTPAHFLTGSPYTA----
>Cer10-1
HEI---NGHLPEQYTLSALKTMYWIPLCKALVRSVISNCIKCKKVFGLPYPYPNTKQLPGCRT-EPSKPFAKVGLDYLGPV-
EYLRDDQSIGKAYVL-LYTCLVTRAAVLRVVPDA-TTESYLMALRT-IF--HQVGVPSEVHSDNGAKMTNDD---IREGSEI--
--DEIFTCFLASQ-EIKFIYITPWSPWQGGVYERIIGLLKHQIHKIC------
GDQKLDFFSLQYVVSSAQAMINNRPLVAHARSPNDMITLRPMDFMIPG--------
>Pyggy
HES-----HTRGEPIKDNLPRRFFWAEHDRSIRRYVRNCDVCEKTNYRE-KVLRTITNT-----RPEDIEKEISIDFIKGL-PTS-----
EGITYLI-VVTNRLSKGSIFILLPNI-KTETVVRAFLR-VV---AYYLLEAITSDRGS-FVS---V---------------LERLCEIL-
KIRRRLLISFY-LTNGSTERIN-SVEAYTRAFI-----------SAQIDASLCSIAQIAINSRDATL-TGV-----------------------
```

Dataset S8. IN alignment generated by PROMALS3D and used to reconstruct the phylogenetic tree in Fig. 4.

```
>ApGIN7a
HDSLL-GGHAGVSR--TVKRLKLNYNWRHMKKDIKHYIKHCEICQKNKSHLKTKKPML----ITSTI--
VKPFEKNCLDIVGPLPK-TLTGNMYILTLQDELSRFSLAIVLSTT---DAQT-VAQAFV-EVSYGIPESILTDRGTNFL----------
---FDVFKSMCKLL-DIKKSKTTPWHPQGNCFLERSHKTLKTYLRSFVDKD----SNWDKLLCYATFCYNTTVHTSTN-------
FTP--------------
>Cer3
HESLG-GGHFGYRK--TLHKIKRNYYWPNMRSDVLKWTLQCKICQQKRNPHPSTRELQ----KIVIT--
TKVFEKVGVDLTGPLRM----TASGNKYMVCWFTKFVISVPLPNASTETVAD-AMMKEL-LLKFGTPSQLVSDRASTFT----
----------SEAFRAFCRKL-EIQQHLAIPYHSKGNGATERTFRTFHNMVSKHVNKTH---TDWDQILPYMTFVYNTTVHDTTG-
-------ETPFFLIFGRDPVFAIDK
>Mdg1
DDPSE-GGHSGISR--TLRKMKNCCCWPRMTKAISEYVETCLKCQQAKTTKHTKTPLT----ITETP--
ATAFDKVLIDTIGPLPR-SENGNEYAVTIICDLTKYLVTVPIPNKSAKSVAK-AIFENF-ILKYGPMKTITTDMGTEYK-------
------NQIIDDLCKYM-KIKNITSTAHHHQTLGTVERSHRTFNEYVRSYISVDK---TDWDIWIQYFTYCFNTTPSVVHE-------
YCPYELVFGRLPRQFIDF
>DrGIN2-2
HFNDI-GHHLGQKK--TVHRIQSKYYWLGIVKDVVDWIKVCETCQHAERSKNLARTV-----RPIKV--
DGPWEILAIEIIGPFPGTVHGGNTHIVIITDYYSKWVEAFPVQKKDGLCVA--RCISSS-VCRFGPSKTIFCSQNADFC--------
-----EEVMKQLCERW-NITVRVLSVDQPQRNALYDRSCNLLRDTIKQMVVEKQ---VEWDDFLDPILAIFRTSVNPTTK------
--FTP---------------
>Tat4-1
HEGAG-GNHSGGRS--LAAKIKKHQYWPTMITDCIRFVAKCEPCQRHASIIHSPTEAL----TTSTA--
PYPFMRWAMDIVGPLPS--SNGKRFMLIMTDYFTKWVEAEAYNKIQAGEVQR-FVWKNI-ICRHGLPYEIITDNGSQFI-----
---------SRQFEDFCAKW-RIRLNKSTPRYPQGNGQAEATNKIILDGLKKRLEAKK---GAWADELD-GVLSHRTTPRRSTG---
-----ATPFSLTYGIEAMAPAEV
>Gmr1
RDALL-GAHLGMDK--TRERVVARFYWPGVRRDVARYCQECPDCQRVRPRAVERSPLI----PMPII--
ETPFERIALDIVGPLPR-TSRGHRYLLVILDYATRYPEALPLRAA-TSKAY--ARISAT-LSRVGLPKEILTDQGSCFM---------
----SRVVKELLKLL-QVSQLRTSVYHPQTDGLVERFNQTIKRMLKKSIEADG---KNWDQLLPHVLFAIREVPQASTG-------
FSPFELLYGRRPRGILDL
>SURL
-------GHLGIEK--CKKRAREVIYWPRINADIAEMVQSCTSCLMYKPKQQAESLH-----PHAVP--SRPWEKVAVDLFTL-------
NKREYMVIVDYYSQFIEVCTMTSTSSKAVI--NHMKAI-FARHGTPCELMSDNGPQFA-------------SQEFKSFAKEW-
DFHHTTSSPHYPQSNGLAENAVKIVKNLLLKS--------QHCGQDIHRALQVYRSSPL-ACG--------
KSPAELLYNRRLRSNLPM
>HmGINO1
HNDAG-GAHQGIVR--TQNKIKNLYYWMSITSVVEAWIKTCKQCQSFEKIKTIAPQL-----HPIQV--
TERWAVLGVDLIGPLPE-TVNGNKYVLTITDLFTKWVVARVLPDKSAAAVA--TALVNN-FHTYGPPKKVITDQGREFV--
------------NEVKTKIFQLM-GIKHSVTSAYHPQSNGQDERTNQTFKCALSKYTNNEM---
NDWDRFSSXIAYGLNISCQKSTK--------VTP--------------
>HmGINA1
YDSNH-AAHVGLNN--TRAKLKGSFYWLGMVNDITKCVKECDKCQRMEKIRTIAPEL-----KSIKV--
NGLWDFLGIDLIRPLPI-TKLGNKYILTITDLWSKYIEAFPIPDKPAFYVS--KCLTTL-FYRFGPPKKILSDQGREFV----------
---NSLNEQLFFLF-QIKHLITSAYYPQTIGQDERTNQTIKKSLSKLSNDTQ---DNWDELLEAVLFGLRTCVQKSTK--------
FTP--------------
>HmGINNY1
------GGHFGRDK--TLAKISERYYWKGMVNDVRSFCKYCDKCQRANRFEKHSAEL-----HPIKV--
DEVWSTVGIDLIGPLPL-TEKGNKYIITATCLFSKWPEAASLSDKTATSAA--EFLYTC-FTRHGCCEVQISDQGREFV-------
------NEVNHELNKMM-GTKCNVTSAYHPQSNGEDERFNQTLQRQLLKYVDEKQ---NTWDLYIESILFSYRVSVQDSTK---
-----QTP---------------
>HmGINER1
------ASHLGINA--VRKQISSRFFWHSIVEDITKHVAQCERCQKTSNRNLKVSP--ALK-PVKVE--QQVMKQVGVDLIKLPE-
--SNGFNYAIVLIDYFSKWTEAEPLIDKTAVSVA--AFLYRV-ICRHGCFQIQINDQGREFV-------------NSVSIALHDMS-
GVQQRVTSAYHPQANGLVERQNQTIKKAIVKVLNENV---KSRASVLDGILFALRVKVQDSMG--------YSP--------------
>HsGIN1
HENDS-GAHHGISR--TLTLVESNYYWTSVTNDAKQWVYACQHCQVAKNTVIVAPKQ-----HLLKV--
ENPWSLVTVDLMGPFHT-SNRSHVYAIIMTDLFTKWIVILPLCDVSASEVS--KAIINI-FFLYGPPQKIIMDQRDEFI----------
---QQINIELYRLF-GIKQIVIS-HTSGTVNPMESTPNTIKAFLSKHCADHP---NNWDDHLSAVSFAFNVTHLEPTK--------
NTP--------------
```

```
>RetroSor1
HSGMC-GSHIGPRA--LSAKALRQFYWPTHIRDAEEIVKTCKACQTFSPIQSGPSALT----QLIPA--
SWPLQRWGMDLVGPMPT-AQGGNKFAVVAIEYFTRWIEAKPLTTITSETI-RKFFWQNI-VCRFGVPRLLTVDNGKQFD---
-----------SDNFKEFCHLI-GTKIAFASVYHPESNGAVERANRTIFSAISKTLLNLR---GKWVEELPRVVWSHNTTVSRATG---
-----FTPFKLLYGEEAMLPEEI
>Kabuki
HN----LSHPGIRT--TRKMITNTFFWPNMNKDIGMWAKPCIRCQESKIYRHTISDL-----GNFVD--AGRFEHIHVDIVGPLPT-
SPQGFRYCITIIDRCTRWPEAFPVHDITADKV-AKTIYDGW-ITRFGCPVRITSDQGRQFE-------------SNLFLKLTKLL-
GINKICTTPYHPQSNGAVERWHRSLKVALMTRLND-----TSWVDELSTAMFGLRASVRTDSG--------
VSAAQLTYGKNIRLPGDF
>Tor1
HL----PGHARTTN--MLGSMRDTVIFPKMAKKIKQKIMTCPTCLATGSTERPRFTPI----TPPKE--SHPYMTVTVDLLGPLPS-
TPTRKKYILAAIDNLTRWIELRCIPDKTAAHVA--KALMDI-FFLRGPPRAVSCDNGREFS-------------NTLLRELLASF-
GTCINYGTPYRPQGQGLIERANREIVKHLKGLNIEE----HKWDTFIPSIQLSMNLTYHSALG--------
SSPFQAMHGWTLAEPLFS
>Tor2
HDL---KGHLGVDR--TAQFLSW-CWWPHKMDDIRTYVASCANCLKQKGYDMQPSRP-----HLYRA--
TRPHEIIYCDFITLPTS-SRSGKRHALTVMCGFSRWLQVYAVHRCRSIDAA--RGLMNY-FLQFDFPRTLSSDRGRHFE------
--------NELLADLCKLL-QIRQNLHCAYRPESSGVIERCHKTLKSSLWAMVRDNP---LDWELALPSVVSAMNRSTNAATK--
------VSPYKCIFGRDPSFNGIA
>Mdg3
HNE---LGHVGRDK--MIEAIMKNYWFPNLKQKCSTHISNCLKCISFSPKTGKTEG--FLH-NIPKG--
NKPFEIIHIDHYGPVDL--ARPKKHILVIVDAFTKFVRLYATKTTNTK-EVI-QSLNDY-FRAYSRPKCIISDRGACFT-----------
---SGDFDSFLKEC-NVKHIKIATGSPQANGQVERINRSLGPMISKLIEPDQG--LHWDLVLEKVEYTLNNTLHRSIK--------
QYPSIMLFGLQQKGQIMD
>Woot
HQL---LGHFGATK--VYNSMKREYYWPNMYRTIKKRLRSCDLCQKTKSSNRPHQGPL----TPILY--
DHIGDLVCVDFYGPLPT-GRLGASYVFVVIDVFSKFLKLYPLRKATAK-IAA-KRLIED-FSGYIKPKCVLSDHGTQFI-------
------SNTWQNSLRAA-DIQPTLSSIRHPESNP-SERVMRELGRIFRAYCRENH---ASWVNHLSNIEDCLNYVPHISTG--------
FSPYEILYGRTPPNPLDA
>ApGIN4c
HDY---YAHTGPLK--TAHALKNICYFPSFNKTIRCIVQACELCQKCKPKTTRIAGPL----QPILS--NQPLDKLLVDFYGPLPT-
GIFQFSYIFVVVVDNFTRFVKLYPLQSANAK-ICI-KKLTTDYFPNYGIPKNIVSDHGRQFI-------------SNYWQTSLKKY-
KVQVSHTSVYHPQSNP-AERVMRELGRMFRTYCHEQH---SLWPQYVPYIEWTLNNVRHESTH--------NTP--------------
>CRM
HGGGL-MGHFGAKK--TEDILAGHFFWPKMRRDVVRLVARCTTCQKAKSRLNPHGLYL----PLPVP--
SAPWEDISMDFVLGLPR-TRKGRDSVFVVVDRFSKMAHFIPCHKTDDATHIADLFFREI-VRLHGVPNTIVSDRDAKFL----
----------SHFWRTLWAKL-GTKLLFSTTCHPQTDGQTEVVNRTLSTMLRAVLKKNI---
KMWEDCLPHIEFAYNRSLHSTTK--------MCPFQIVYGLLPRAPIDL
>Del
HRSKF-TIHPGSTK--MYRNLKINFWWSGIKREVVEYVSRCLICQQVKADHHHHSGLL----QPLPV--
SEKWEHILMDFIIGFPS--KRCHDSIWVIVDRFTKSAHFIPIHTTISGKDLA-LYIKEI-IRLHGIPTTIVTDRDTKFT-------------
SRFWGSLKSL--GTELFFSTAFHPQTD-GSERTIQILEDMLRSCSLDFK---GNWEEHLPLVEFAYNNSYQSSIG--------
MAPFEALYGRPCRSPTCW
>Galadriel
YDSAW-AGHPGVER--MLALLSRVYFWPKMEDDIEAYVKTCHVCQVDKTERKKEAG--LLQ-PLPIP--
ERPWLSVSMDFISGFPK--VDGKASIMVVVDRFSKYSVFIAAPELCSSEVAAELFYKHV-IKYFGVPADIVSDRDTRFT------
--------GRFWTALFNMM-GTELKFSTANHPQTDGQTERINHLLEEYLRHYVTASQ---RNWVELLDTAQFCYNLHKSSATE-
-------MSPFEIVLGKQPMTPLDV
>Amn-san
HSSKQ-AGHPGSEK--TLELLRRLVWWPTIRKDVRDFVAACTVCATTKASHSRPCG--LLH-PLPIP--
SRPWTHLGMDFIVELPP--SCGNTVIWVVIDRFSKMAHFIPLRKLPSAVELAHLFIQHI-FRLHGFPVEIVSDRGSQFV--------
------SRFWRSLCKSL-GVSLQFSSAYHPQTNGAAERVNQALEQFLRNHVSLCQ---DDWSDLLPWAEFAHNNASHSSTG---
-----RSPFLSVYGQHPLAFPQD
>Sushi-ichi
HSSCF-ACHPGVRR--TAEFVQRRFWWPNLQEDVREFVGACTVCARSKASHRSPAG--LLH-PLPVP--
SRPWSHVALDFVTGLPV--SQGNDTILTIVDRFSKGVHFVALPKLPSAAETAELLVSHV-VRLHGIPLDVVSDRGPQFT-----
---------SRVWQAFCKGI-GATVSLSSGYHPQSNGQAERANQAMEAALRCVTTSNP---
ASWSKFLPWVEYSLNAMESSATG--------MSPFQCFLGYQPPLFPQQ
>Reina
HDSPI-GGHSGFPV--TYHRIKKLFYWAGMKGQIKEFVQSCEICTKAKADRNRYPG--LLL-PLPIP--
DQAWQVISLDFISGLPT--SRRFNCILVVVDKFSKYAHFLAMSHPFTALSVAKLFLSQV-YKLHGLPLSIISDRDPIFT---------
```

```
-----SNLWQELFKLV-GTKLCLSSAYHPQSDGQTEWVNQCVEAYLRCFVHGCP---KQWSNWLSLAEFWYNTCFHTALG--
------QSPFEVLYGHTPSQLGLS
>Maggy
HDSPV-AGHPGKAK--TYDLLSREYYWPGMLHYVSLWVKKCQTCRRINPSREGHQG--LLR-PLPTP--
ERSWQHLSMDFITHLPQ--SNGHDAILVVVDRLTKMRHFVPCKGTCNAEDTANLYLHHV-WKLHGLPLTIVSDRGTQFV-
-------------SKFWKHLTTRL-KIDSLLSTAHHPETDGQTERFNASLEQYLRAYVAYLQ---
DDWESWLPLAEFTANSHKSETTG--------TSPFYATYGFHPRMGFE-
>Skippy
HAHPL-HGHQGVTK--TMKRLQELGYRHFKKGQVEKVIKQCDLCAKTKAQRHKPYG--QLQ-PLPVA--
QRPWDSITMDFITKLPT--GIFYDSIMVIVDRLTKFSYYLPYREATDAEELSYVFYRHI-VSIHGLPTEILSDRGPTFA---------
----ATFWQSLMARL-GLNHRLTTAFRPQVDGQTERMNQVLEQYLRCYINYEQ---NDWVEKLPIAQLAYNTAYNESTK----
----LTPAYANFGFTPNAYHNA
>TF1
HEEGK-LIHPGIEL--LTNTILRRFTWKGLRKQIQEYVQNCHTCQINKSRNHKPYG--PLH-PIPPP--
ERPWESLSMDFITALPE--SSGYNALLVVVDRFSKMAILLPCTKSITAEQTARMFDQRV-IAYFGNPKEIIADNDHIFT--------
------SQTWKDFAHKY-NFVMKFSLPYRPQTDGQTERTNQTVEKLLRCVCSTHP---NTWVDHISLVQQSYNNAIHSATQ----
----MTPFEIVHRYSPALSPLE
>Diaspora
HSSAP-GGHLGVQR--TARKVLDCFYWPTIFKDAWKICSTCEQCQRAGNTLTWRQQ--MPQ-QPMLF--
CEVFDVWGIDFMGPFPV--SFGYVYILLAVDYVSKWVEAKPTRTNDAKVVAD-FVRSNL-FCRFGVPKAIVSDQGTHFC--
------------NKTMHALLKNY-GVVHRVSTPYHPQTNGQAEISNREIKRILEKIVQPSR---
KDWSTRLDDALWAHRTAYKAPIG--------MSPYRVVFGKACHLPVEI
>Athila4-1
HGSAY-GGHFATFK--TVSKILQAFWWPTMFKDAQEFVSKCDSCQRKGN-INRRNE--MPQ-NPILE--
VEIFDVWGIDFMGPFPS--SYGNKYILVAVDYVSKWVEAIASPTNDAKVVLK-LFKTII-FPRFGVPRVVISDGGKHFI--------
------NKVFENLLKKH-GV------------KQVEISNREIKTILEKTVGITR---KDWSAKLDDALWAYRTAFKTPIG--------
TTPFNLLYGKLCHLPVEL
>ApGIN3c
-----------MVE--YESIARSYFWWPSIDKEIEDMARNCNECINARPNPPKSVLT-----PWKWP--QRQWTRVHCDFLGPY------
KNKTFLIIVDATTKWLEVFQVNSMTAQTVI--TKLSEL-IARFGIPRTITTDNAKCFN-------------GDEFQSYCKTL-
GIRHLTGAPFHPASNGFAESGVKIVKQFFNKC--------STSLHHLNKFLLMYRNTVHSTTN--------ETP--------------
>Mag
HD-----PHMGIVK--TKSLARSYVWWPGI-DEAETECRACTVCAAVADAPSTHAPR-----SWPWP--SRPWSRLHLDFLGPI---
---GGVTYLVVVDSCSKWIEAIKMQRTTAQAVI--SVLRDL-WSKFGLPKQTVSDNGPPFS-------------SSDFQKFLIHN-
GIKHIYSAPYHPASNGAAENAVKICKRAIKKALKQN----LNVDTALCRFLLAYRNTEHATTG--------
DSPANILQGRSLRMRLDN
>SPM
HQ-----EHTGIVR--MKAVARSYFWYPKLDADIEKLSQSCEACLRMRNDPPKVPFV-----PWSNA--RKPWERVHVDFFELE-
------GKDYLVLVDTYSKWVNVELMTSTTSAKTI--ETLRSW-FAIFGLPETLVSDNGPQFT-------------SEEMEVFLSKN-
GVKHVLVPPYHPASNGAAERTVQIVKRTLQKYFLSDK---VSIRHRLDNFLFSYRTTPQTVTG--------
VSPYELVFKFAPRTRFSL
>Cer4
HE-----GHPGVVR--MKQKACSFVFWTGIDKDVEKLVRGCENCQESAKMPRVAPLR-----PWPEP--
QKAWSRVHIDFAGPV------NGHWFLVIVDAKSKYAEVKMTKTISASATV--SLLEEV-FATHGYPELLVSDNGTQFT--------
------SNQFKLMCQEY-GMEHKTSAVYYPRSNGAAERFVDSLKRGLAKITRSG----VVTQQALNKFLICYRNTPHSALA-----
---ATPAECHFGRKIRTKMSL
>CoDi6.7
HHR---LGHASMTK--IRMLAKVGLLPASLKECQI---PLCTSCLYGKATRRPWRT-ASTR-SIHKV--
NGPGQCISVDQLVSTTPPTLKRYHAATVFVDNFSRLSYVHVQKGTSAEETIQAKHAFER-YARSHGVVTHYHADNGIFA-
-------------DNKFREAVKED-
RQTLSFCGVNAHFQNGIAERRIRELQDHARTMLIHATKDSHLWPYALRMANELHNHLPSLRGS--------VPPIEIFA----------
>pCretro6
HRI---LGHAHLAS--IKRLCLGLAEGVD-VDMTSDDQFQCESCVRAKQHVQPFPD--QAT-RTPDK--
LKVGEIVASDTWGPAQVRSVHSFYYYMSMTDLKSRFSGVYFSAQKSGLAKMALRGLIRR-MSRTGSSIMLRIDNGTEFI-
-------------NNEFLQYCKSQ-
GIVVETTAPYSPAQNGVAERLNRTLRESARAMLLAAELPRKFWPEAVSYACLIKNRLPHAALK--------MTPYQALTG-----
----
>BFV
HDGPG-SAHLGRSA--LLLKLQQKYWWPRMHIDASRIVLNCTVCAQTNSTNQKPRPPL----VIPHD--
TKPFQVWYMDYIGPLPP--SNGYQHALVIVDAGTGFTWIYPTKAQTANAT---VKALTH-LTGTAVPKVLHSDQGPAFT----
----------SSILADWAKDR-GIQLEHSAPYHPQSSGKVERKNSEIKRLLTKLLAGRP---TKWYPLIPIVQLALNNTPNTRQK----
----YTPHQLMYGADCNLPFEN
```

```
>EFV
HKGPG-SGHLGKNT--MYIKILNKYWWPNLIKDISKYIRTCTNCIITNTDNVPNKSYI----VQEKT--
GLPFQKYYMDYIGPLPP--SDGYYHVLVIVDEGTGYTWLYPTKAQTANAT---VKALNH-LTGTAIPKVLHSDQGSAFT-----
---------SATLVAWAKDK-GIQMEYSSPYHPQSSGKVERKNSEIKRLLTKLLVGRP---TKWYPLIPTVQLALNNTPNAKIG---
-----KTPHQLMYGVDCNLPFQD
>SFV
HE----LAHTGREA--TLLKLQTTHWWPNMRKDIITVLRQCKPCLQTDSTNLTPIPPV----SQPRP--VKPFDKFFIDYIGPLPP--
SHGFSYVLVIVDARTGFTWLYPTKAPSTNAT---INSLNL-LLGTAIPKVLHSDQGSAFT-------------SATFAEWAKDR-
SIQLEFSTPYHPQSSGMVERKNKEIKRLITKLLVGRP---TKWYPLLPTVQLALNNTYSPRTK--------
LTPHKLLFGVDGNVPFAN
>FFV
HN----ISHAGREA--VLLKIQENYWWPKMKKDISSFLSTCNVCKMVNPLNLKPISPQ----AIVHP--
TKPFDKFYMDYIGPLPP--SEGYVHVLVVVDAATGFTWLYPTKAQTSKAT---IKVLNH-LTGLAIPKVLHSDQGSAFT------
--------SEEFAQWAKER-NIQLEFSTPYHPQSSGKVERKNSEIKKLLTKLLVGRP---LKWYNLISSVQLALNNTHVVSTK-----
---YTPHQLMFGIDCNLPFAN
>MuERV-L
HEQ---SGHGGRDG--GYPWAQQ-HGLPLTKADLATAAADCQICQQQKPTLSPRYG------TIPRG--
DQPATWWQVDYVGPLPS--WKGQRFVLTGVDTYSGYGFAFPARNASAKTT-IHGLTECL-IYRHGIPHSIASDQGTHFT----
----------AREVRQWAHDH-GIHWSYHVPHHPEAAGLIERWNGLLKMQLQHQLGGNS---
EGWGRVLQKAVYALNQRSIYGT---------VSPIARIHGS--------
>HERV-L
HEQ---SGHGGRDA-GGYHWAQNQHGLPLTKADLATATAECPICQQQRPTLSPRYG------TIPQG--
DQPATWWQVDYIGPLPS--WKGQRFVLTGIDTYSGYGFAYPACNASAKTT-IHGLTECLLIHHHGIPHSIASDQGTHFT-----
---------AKEVWQWAHGM-GIHWSYHVPHHPEAAGLIERWNGLLKSQLQCQLGDNT---
QGWGKVLQKAVYALNQHPIYGT---------VSPIARIHGSQT------
>CoEFV
HCG---IGHGGNNT--GKTLRRL-YWWPGLYTECKRYVANCPECVNVTPTTRPPNL------RPPRG--ALFDKIF-LDFVGPLPR-
-SNGYTAILILVESLSSFTWLLPCRDQSASTT-VTALSARF-FPASIEPKCFHSDQGGDFT-------------SQLFKKMCSER-
NIRVEYSTPHHPQSAGVVERKNRGLKAALT-KLVRNR---RKWFQVLDIVQTGLNNTPIARNE--------
ATPFFLMYGQHMNTPLTN
>ApGIN5c
------------------------YSTSDLKKIGKIMNNCVTCILANRKQGKKEG------SIDKS--DTPLAMWHIDFLGPLTP-
TPKGYRHILSVIDGFTKFCWLFPTKSVTAQEVI--EKLTVM-EATFGNPGKIVCDRGTAFT-------------SKDFGDYCSGR-
EIRHILITIGVPRANGQVERLNTIIINVLTKLSLNKT---DQWYRQYQSIKEAVELEYVRCHE-------------------------
>Xen-1
HG----PTHVSKNL--MNNLISKLYFAPGITTLTRNYTAACTICAQCNPGRMEKPP------NLAKP--LYPFQRIQIDHIQMPR---
CGRFEYVLVVVDMFSGWPEAFPVANMTAKTT-AKKLLSEI-VCRYGVPEVIESDQGPVFT-------------ASVTKDIWTAL-
GVTLHFHTPYHPQSSGKVERMNGTLKTKMLKMSQDSG---MLWPDSLPIALFSVRYTPRGVNN--------
LSPFEILFGCAPRL----
>SnRV
HL----YGHPSEES--LRKVLTKRFVWEDMGQHCKEITNTCLTCAKYKVLRAGPPMG-----VGRSA--
EGPCQKLQVDHVGPLGP-GTHGYRYLTTMVDVYTGWFWAKPCRGPTTGATI--AALEEH-ISIWGVPYSIQSDNGTAFT---
-----------SKAMQEWANTY-GIEWKVGAIYHPQSQGKVERKHRLLKDRLKRATHEG----KNWVQALPSILLFINSM-
HPRDQ--------FSAYELMTGRVPHL----
>MuLV
HQ----LTHLSFSK--TKALLERNYYMLNRDRTLKDITETCQACAQVNASKSAVKQG-----TRVRG--
HRPGTHWEIDFTEVKPG--LYGYKYLLVFIDTFSGWVEAFPTKKETAKVV-TKKLLEEI-FPRFGMPQVLGTDNGPAFV----
----------SKVSQTVADLL-GVDWKLHCAYRPQSSGQVERMNRTIKETLTKLTLATGS--RDWVLLLPLALYRARNTPG-
PHG--------LTPYEILYGAPPPLVNFP
>KoRV
HQ----LTHLGPDK--LLQLVGRTFHIPNLQSVVREITSKCQVCAVTNAVTTYREPG-----RRQRG--
DRPGVYWEVDFTEVKPG--RYGNRYLLVFIDTFSGWVEAFPTKTETALTV-CKKILEEI-LPRFGIPKVLGSDNGPAFV------
--------AQVSQGLATQL-GIDWKLHCAYRPQSSGQVERMNRTIKETLTKLALETGG--KDWVTLLPLALLRARNTPG-QFG--
------LTPYEILHGGPPPVLASG
>FLV
HK----LTHLSARK--MKTLLEREFYLPNRDLHLRQVTESCRACAQINAGKIKFGPD-----VRARG--
RRPGTHWEVDFTEIKPG--MYGYKYLLVFIDTFSGWAEAYPAKHETAKVV-AKKLLEEI-FPRYGIPQVLGSDNGPAFI-----
---------SQVSQSVATLL-GINWKLHCAYRPQSSGQVERMNRSIKETLTKLTLETGS--KDWVLLLPLVLYRVRNTPG-PHG--
------LTPFEILYGAPPPMAHFF
>REV
HR----ATHLGESK--LTELVRKHYLICGIYRAARDITTRCVACAQVNPGAAPVEKGL----SRIRG--
AAPGEHWEADFTEMITA--KGGYKYLLVLVDTFSGWVEAYPAKRETSQVV-IKHLIHDI-IPRFGLPVQIGSDNGPAFV-----
```

```
---------AKVTQQLCEAL-NVSWKLHCAYRPQSSGQVERMNRTLKETIAKLRKETGG---DWVSLLPQALLRARCTPG-REG-
-------LSPFEILYGLKPPVVP--
>Loki-Ofa
HDSPV-GGHRGYRA--TLQTLQQVAYWPLMARDTKSYVQGCLVCCQFQPSRPLGRAPL----QKRGV--
TFPWSHLQVDWVGPVPK-SSRGNKYLLTVTCAFTKWVECLPAPNDTAMTT-ATLLLNHV-
FSRWGLPLSVDSDRGTHFT--------------SSVMTTLYDIL-GVEVKFHISHHPQSSGQVERANRTIVSMLKKYVSSNG---
KDWDVKLPLVLMAIRSTPHRSTG--------VTPFEMMTGREMTLPLHL
>Loki-Aro
HDEPC-GGHRGEQP--TIHMLWQVAYWFKIWTDVKDYVNGCTICHKFQPVLPRHRAPL----QSAPS--
AQPWSHIQIDWIGPVTR-STWGNXYALTITDRFTKWTKCLTALADTAETT-AVLLMKHV-FTHWGLPQSVDSDGGTHFT-
------------ATIMEIEWKTL-GVKRQLHIAYQPQSSGQVERANQTVINMLKKQLDANQ---
KDWDIKLPLVLMTIXAMPSKATQ--------VSLFEMMTDRQIVLPLHL
>Loki-Xla
HDAPT-AGHRGEKI--TYELLRDYAYWPHMLQDVRTYCQGCLVCPQFQPQGPNHRAPL----MKRGI--
SMPWSDIQIDFIGPVTT-SSKGNRYMLTVTCLFTKWVECLPCKTCSSSVC-ASLLINHI-FSRFGLPQRIESDRGSHFT---------
-----SEVMTKMWEIL-GVKRKLHIAYRPASSGGVERYNQSIVNILKKFVKESG---KDWDVKLPLVLMAIRATPSAATK------
--LSPFEMITGRKMVLPQHL
>Loki-Xtr
HDAPT-AGHRGEKI--TYELLRDYAYWPHMLQDVRTYCQGCLICPQFQPQGPNHRAPL----MKRGI--
SMPWSDLQIDFIGPVTT-SSKGNRYMLTVTCLFTKWVECLPCKTCSSSVC-ASLLINHI-FSRFGLPQRIESDRGSHFT--------
------SEVMTKMWEIL-GVKRKLHIAYRPASSGGVERYNQSIINILKKFVKESG---KDWDVKLPLVLMAIRATPSAATK------
--ISPFEMITGRKMVLPQHL
>Loki-Npa
HDAPT-AGHRGEKI--TYELLRDYAYWPHMLKDVQTYCQGCLVCAQFQPQGPTHRAPL----QKKGF--
SLPWSDIQVDFIGPVTR-SSRGNKYMLTVTCVFTKWVECLPAHNNTAETC-ATLLINHV-FSRYGLPLRIDTDRGSHFI------
--------SEVMSKMWKIL-GVKRKLHIAYRPASSGSVERYNQSIVNILKKFVKESG---KDWDMKLPLVLMAIRATPSAATK---
-----MSPFELMTGRKMVLPQHL
>Loki-Rca
HDAPT-SGHRGEKI--TYEILRDYAYWPHMLRDVKAYCQGCLICPQFQPASPTHRAPL----QKRGI--
CMPWSDLQMDFIGPVTK-SSRGNRYMLTVTCLFTKYV-CIPAPNNSSDTC-AALLINHI-FSRFGLPQRIESDRGTHFT-------
-------SEVMTKLWKIL-GVKRKLHIAYRPASSGGVERYNQTIVNILKKFVKETG---KDWDVKLPLVLMALRATPNTVTK---
-----MSPFELMTGRRMVLPQHL
>Loki-Pma
HDEPC-SGHRCEMN--TGQMLRQVAYWPKMGEDIHKHVSNCLVCCKYQPTTSQHRAPL----QKRGT--
DYPWSNLQIDWIGPVNR-STKGNKYMLTVTCAFSKWIECLPAPNNTAETT-AILLINQI-FSRWGLPSTIDSDKGSHFT------
--------AEVITEVWKML-GVKRQLHVAYHPQSSGQVERANRTIVSLLKKYVSTTG---
RDWDTKLPLVLMAMRATPNRATG--------RTPFEIMTGRLMTLPVHL
>Loki-Lca
HNEPC-SGHRCEMN--TGQMLRQVAYWPKMGEDVHKHVSNCLVCCKYQPTNSQHRAPL----QKRGT--
DFPWSNLQIDWIGPVNR-STKGNKYMLTVTCAFSKWVECLPAPNNTAETT-ATLLVNQI-FSRWGLPSTIDSDKGSHFT----
----------ADVIAEVWKTL-GVKRQLHVAYHPQSSGQVERANRTIVSLLKKYVSTTG---
CDWDTKLPLVLMAIRATPNRATG--------RTPFEIMTGRQMTLPVHL
>Loki-Dcl
HDTPS-ARHHGTST--TYETLEQAAYWPSMQQDLAEYVKEFFPCYQFQPTKTNHRASL----QRRGV--
TSPWSDLQSDWVGPLPR-STRGYKYFLTVVFQFTKWVECLPAPNDTAQTT-AYLLMNHI-FSRFGLPLRVNSDRGTHFT--
------------AEIMQQTWKLL-GIQARLHISQHPIASRQVERMNRTVVSMLRKFVSPNQ---
KDWDVKLPLVLMAIRATPNEATG--------VSPFELMTGRQMTLPLHL
>Loki-Lch
HDEEC-GGHRADRA--TYETLKQVAYWPNMYQDVQDYTKGCLVCCRFQSSSPKHRAPL----QTKGI--
TMPWSDLQIDWIGPVTR-SSRGNKYLLTVTCLFTKWIECLPAPNDTAETT-AMLLINHV-FSCWDLPMRIESDRGSHFT----
----------GKVIENVWNTL-GVQRKLHVSHHPQSSGQVERANKTIVNILKKYISTSG---KNWDVKLPLVLMAIKATPNKSTG-
-------VTPFEMMTGREMTLPLHL
>Loki-Str
HDEPC-GGHRGVKA--TCETLRQVAHWPHMEQDVARYVRGCLVCCQFQPTKPLHRAPL----QRKGV--
SYPWSSIQIDWVGPVAR-SARGNKYLLTVTCAFTKWVECLPATNDTAETT-AVLLLNHV-FSRFGLPGEVDSDRGTHFS---
-----------AAVMTELWKLL-GVKAKLHIAYHPRSSGGVERSNQSIVRILRKYVAANH---
KDWDLKLPLVLMAIRATRNRSTG--------MTPFEMMTGRQMTLPLHL
>Loki-Lro
HDAAC-AGHHGTRA--TYETLKQVAYWPGMQQDVAEYVKGCLVCCQFQPANPNHRAPL----QRRGV--
TFPWSDLQIDWVGPLPR-STRGNKYFLTVVCQFTKWVECLPAPNDTAQTT-AYLLMNHI-FSRFGLPLKVNSDRGTHFT---
-----------AEIMQQIWKLL-GIHAQLHISHHPISSGQVERSNQTVVSMLRKYVSANQ---KDWDVKLPLVLMAIRATPNESTG-
-------VSPFELMTGRLMTLPLHL
```

```
>Loki-Tfu
HDAPC-AGHHGTKA--TYETLKQVVYWPGMQQDVTEYVKGCLRCCQFQPANPNYRAPL----QRRGV--
TFPWSDLQIDCVGPLPR-STRGNKYFITVVCQFTKWVECLPAPNDTAHTT-AYLLMNHI-FNRFGLPLRVNSDRGTHFT----
----------AEVMQQIWKLL-GVKAKLHISHHPISSGQVERTNRTVVSMLRKYVAANQ---
XDWDVKLPLVLMAIKATPQESTG--------ISPFELMTGRQMTLPLHL
>Loki-Oja
HDLPS-SGHKGINA--TYSALKQIAFWPHMKKDVTDYVNGCLVCCQFQPTNPTHRAPL----QKKGI--
SFPWSDIQIDWVGPLTK-SSRGNKYFLTVTCAFTKWVECLPAPNDTAQMT-ALLLMNHV-FSRFGLPMRVNSDRGTHFT-
-------------AEVMQQVWNTL-GVKANFHVSHHPQSSGQVERANRTVVNILKKYVAASH---
KDWDIKLPLVMMAIRATPHESTG--------TSPFEMMTGRQMTLPLHL
>Loki-Ame
HDSPS-AGHKGIKE--TYRTLKQVAFWPRMREHVASYIKGCLVCCQFQPANPLHRAPL----QRKGL--
TFPWSDLQMDWVGPLTK-SARGNKYFLTVTCAFTKWVECLPAANDTAITT-ACLLVNHI-FTRFGLPCRVDSDRGTHFT--
------------AEVMQQMWHVL-GVKAKLHVSYHPQSSGQVERANRTVVSILKKFVATHH---
RDWDVKLPLVLMAIRATPHDSTG--------VSPFEVMTGRQMTLPLHL
>Loki-Sme
HDSPC-SGHRGIKA--TLKSLQEVAYWPHMQKDVLDYIKGCLVCCQFQPASTLHRAPL----QRRGV--
SFPWSDLQIDWVGPLTK-STRGNKYFLTVVCAFTKWVECLPAPNDTAQTT-AYLLMNHI-FSRFGLPGRVNSDRGTHFT--
------------AEMMQQLWQLL-GVKANFHISYRPQSSGQVERANRTVVNILRKYVSSNH---
KDWDVKLPLVLMAIRATPHESTG--------LSPFELMTGRQMTLPLHL
>Loki-Sri
HDSPC-AGHKGSKA--TYNALRQVAYWPQMRKDVADYIKGCLVCCQFQPANPVHRAPL----QRRGI--
SFPWSDLQIDWVGPLTK-STRGNKYFLTVTCAFTKWLECLPAPNDTAQTT-AYLLMNHI-FSRFGLPSRVNSDRGTHFT---
-----------ADVMQELWQIL-GVKANFHISHHPQASGQVERANRTVVSVLKKYVTSNH---
RDWDVKLPLVLMAIRATPHEATG--------FSPFEMMTGRQMTLPLHL
>Loki-Cne
HDSPV-GGHRGYKA--TLHTLQQVAYWPLMAND-----------------------------XRRGV--TFPWSHIQIDWVGPVPR-
SARGNKYFLTVTCSFTKWVECLPAPNDTATTT-AVLLLNHV-FSRWGLPLSVDSDRGTHFT-------------
SNVMTALFEML-GVEVRFHISYHPQSSGQVERMNRTVVNMLRKYVDSHA---KDWDVKLPLVLMAIRFTPQTSTG-------
VTPFEMMTGRQMTLPLHL
>Loki-Sau
HDTPI-GGHRSYKA--TLKTLQQVAYWPFMARDTKLYVQGCLICCQFQPSRPLNRAPL----QERGI--
TFPWSHLQVDWIGPVPK-SSRGNKYLLTVTCSFTKWVECLPAPNDTAVTT-AVLLMNHV-FSRWGLPLSIDSDRGTHFT--
------------SSVTTALYDIL-GVEAKFHIAYHPQSSGQVERANRTIVNMLKKYVGSNG---
KDWDMKLPLVLMAIRSTPHRSTG--------VTPFEMMTGRQMTLPIHL
>Loki-Bsp
HDSPV-GGHRGYKA--TLQTLQQVAYWPSMSKDVKIHVQSCLVCCQFNPSRPLNRAPL----QKRGM--
TFPWSHIQVDWIGPVPK-SSRGNKYLLTVTCAFTKWVECLPAPNDTAVTT-AVLLLNHV-FSRWGLPLSVDTDRGTHFT--
------------SHVMTELYHIL-GVNVRFHISHHPQSSGQVERANRTIVQMLKKYVSSHG---
KDWDMKLPLVLMAIRATPHRATG--------VTPFEMMTGRQMTLPLHL
>Loki-Ape
HDSPV-GGHRGPQP--TLETLRQVAYWPSMSRDVKAYVKGCLVCCQFQPSRPLGRAPL----QKRGV--
TFPWSHLQVDWVGPVPR-SSRGNKYLLTVTCAFTKWVECLPAPNDTATTT-AVLLLNHV-FSRWGLPLSVNSDQGTHFT-
-------------SGVMTAMYDIL-GVKARFHIPYHPQSSGQVERANRTIVNMLKKYVNGSA---
RDWDVKLPLVLMAIRSTPHRTTG--------ITPFEMMTGREMTLPLHL
>Loki-Pfl
HDSPV-GGHRGPQP---XETLRQVAYWPSTNRDVKAYVRGCLVCCQFQPSQPLGRAPL----QKRGV--
TFPWSHLQVDWVGPVPR-SSRGNKYLLTVTCAFTKWVECLPAPNDTATTT-AVLLLNHV-FSRWGLPLSVDSDRGTHFT-
-------------SGVMTAMYDIL-GVKTRFHIPYHPQSSGQVERANRTIVNMLKKYVNGSA---
RDWDVKLPLVLMAIRSTPHRTTG--------ITPFEMMTGREMTLPLHL
>Loki-Pyo
HDSPV-GGHRGTRL--LXSTLQQVGYWPSMSNDTMSYVRGCLTCCQFQPSQPLNRAPL----QGRGV--
SFPWSHLQIDWIGPVPK-SSRGNKYLLTVTCAFTKWVECLPASNGTAVTI-AVLLLNHV-FSRWGLSISIDSDRGTHFT------
--------SHVMTSLCKIL-GVEVRFHISHHPQSSGQVERANRTIIQMLKKYVGSNG---KDWDIKLPLVLMAIRLTPHRSTG-----
---VTPFEMMTGRQMTLPLHL
>Loki-Bpe
HDTPV-AGHRGVQA--TYQALQQVAYWPHMRKDVQEYTKGCLVCCQFQPSRPLHRAPL----HKRGI--
TFPWSDLQIDWVGLLTK-SARGNKYILMVTCAFTKWVEGLPAPNDTAQTT-AFLLMNHV-FSRFGLPCKIDSDQGTHFT--
------------AAVMSELWQTL-GVKATFHISHHPMASGQVERTNRTVINILKKYVSANQ---
RDWDVKLPLVLMAIRATPNKSTG--------ISPFEMMTGRQMTLPLHL
>MmCGIN1
```

```
HDSPI-GEHQGLED--TYKTVRLLGWWPGMQDHVRDYCRSCLFCIPRNLIGGELKV-IESP-WPLRS--
TAPWSSLQIEVVGPVTV-SEEGHKHVLIVADANTRWVEAFPLKPYTHVAV-AQVLLQHV-FARWGVPIRLEAAQGPQFA-
-------------RHVLVSCGLAL-GAQVTTLSRALQFPCLMSSEAYWEFKRALKEFIFLYG---
KKWAASLPLLHLAFRASTTEATP-------------------------
>HsCGIN1
HDIPL-GAHQRPEE--TYKKLRLLGWWPGMQEHVKDYCRSCLFCIPRNLIGSELKV-IESP-WPLRS--
TAPWSNLQIEVVGPVTI-SEEGHKHVLIVADPNTRWVEAFPLKPYTHTAV-AQVLLQHV-FARWGVPVRLEAAQGPQFA-
-------------RHVLVSCGLAL-GAQVASLSRDLQFPCLTSSGAYWEFKRALKEFIFLHG---
KKWAASLPLLHLAFRASSTDATP-------------------------
>ISSc
---------------------------------------------------------------LRWQV--
ERPGALWHGDVCHVTGCGGKAMPLRIHGLLDDASRYVVALEAHTTKEIDML--AMTVDA-
LRRHGKPDALYLDNGSTYR-------------GDVLKTACARL-GITLLHAKPYDPEARGKMERFWRTL--REGCLTYLGA---
VASLVDINTTLRAYHPAPHAGLL--------QTP--------------
>AciFLERV
HEA---LGHAGADP--TRRELKKQQLWIPV-REIRRVLKDCVECGRYNAGRRGQRMGG----LTIKS--
TVPWGSICMDVAGPLGITGMKGEKYLLVLVDSMSGYVAVKPVRQANGSSVV--SMLDQA-
CCYMGIPKELRTDNGTHFR-------------NAKVDQWCQEN-
GVMRVYSPPYTPQANGVVERTIGLVKSWIGKNANGKEWKVLELVKTLNDRHRQGRPSPAEELN--------
QRPFTSIQGVEQLKPVPD
>AliFLERV
HEA---LGHAGLLP--TRRGLVKQNLWIPM-KQIQQVLKRCTVCGKYNAGRRGKRTEG----LTIKS--
TIPWGSICMDVAGPMGVTGTKGEKYLIVLVDSMSGFVVVRPVRKANGNSVV--SMLDYT-CACLGIPRELRTDNGTHFK-
-------------NVQVNRWCEQH-
GVMRVYSPPYTPQANGVAECTIGLVKNWLAKNANSKEWKSVDIGRALNDRDRAGRPSPSDELN--------
QRPFVTIQGVEQLKPALG
>JSRV
HD----LHHQNSHS------LRLQFKISR--EAARQIVKSCSTCPQFFVLPQYG--------VNPRG--LRPNHLWQTDVTHIP---
QFGRLKYVHVSIDTFSNFLMASLHTGESTRHCI--QHLLFC-FSTSGIPQTLKTDNGPGYT-------------SRSFQRFCLSF-
QIHHKTGIPYNPQGQGIVERAHQRIKHQLLKQKKGNELYSPSPHNALNHALYVLNFLTLDTEG--------
NSAAQRFWGERSSC----
>SRV-1
HT----LHHLNAQT------LKLMFNIPR--EQARQIVRQCPICATYLPVPHLG--------VNPRG--LLPNMIWQMDVTHYS---
EFGNLKYIHVSIDTFSGFLLATLQTGETTKHVI--THLLHC-FSIIGLPKQIKTDNGPGYT-------------SKNFQEFCSTL-
QIKHVTGIPYNPQGQGIVERAHLSLKTTIEKIKKGEWYTKGTPRNILNHALFILNFLNLDDQN--------
HSAADRFWHSNPRK----
>MMTV
HA----LHHQNAAA------LRFQFHITR--EQAREIVKLCPNCPDWGHAPQLG--------VNPRG--LKPRVLWQMDVTHVS---
EFGKLKYVHVTVDTYSHFTFATARTGEATKDVL--QHLAQS-FAYMGIPQKIKTDNAPAYV-------------SRSIQEFLARW-
KISHVTGIPYNPQGQAIVERTHQNIKAQLNKLQKAGKYY--TPHHLLAHALFVLNHVNMDNQG--------
HTAAERHWGPISAD----
>RSV
HT----ALHIGPRA------LSKACNISM--QQAREVVQTCPHCNSAPALEAG---------VNPRG--LGPLQIWQTDFTLEP---
RMAPRSWLAVTVDTASSAIVVTQHGRVTSVAVQ--HHWATA-IAVLGRPKAIKTDNGSCFT-------------
SKSTREWLARW-GIAHTTGIPGNSQGQAMVERANRLLKDRIRVLAEGDGFPTSKQGELLAKAMYALNHFERGENT-------
-KTPIQKHWRPTVLT----
>LDV
HA----FLHLGAKA------LAQECNIPL--TEVERLVKSCSHCQKTPLWMAG---------VNPRG--LEPNQIWQTDFTMYP---
PFRPNHWLAVTVDTYSGVIVATAHRGTKSRHAI--HHWTAA-IAYLGRPSVIKTDNGSAFT-------------
SKAATAWAQRW-HITLKHGIAYNSQGQAIVERANRTLKARLEALGEGEGWPPEHQPELLYRALYSMNFFAR-STE-------
QTSVARHWSHQRLP----
>EIAV
HE----NWHTSPKI------LARNYKIPL--TVAKQITQECPHCTKQGSGPAGCV-------------MRSPNHWQADCTH-------
LDNKIILTFVESNSGYIHATLLSKENALCTS--LAILEW-ARL-FSPKSLHTDNGTNFV-------------AEPVVNLLKFL-
KIAHTTGIPYHPESQGIVERANRTLKEKIQSHRDNT----QTLEAALQLALITCNKGRESMGG--------QTPWEVFITNQAQV--
--
>HIV
HE----KYHSNWRA------MASDFNLPP--VVAKEIVASCDKCQLKGEAMHGQV-------------DCSPGIWQLDCTHL-------
EGKVILVAVHVASGYIEAEVIPAETGQETA--YFLLKL-AGR-WPVKTIHTDNGSNFT-------------GATVRAACWWA-
GIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQA----EHLKTAVQMAVFIHNFKRKGGIG--------YSAGERIVDIIATD--
--
>FIV
```

```
HE----KFHSDPQY------LRTEFNLPK--MVAEEIRRKCPVCRIRGEQVGGQL--------------KIGPGIWQMDCTHF-------
DGKIILVGIHVESGYIWAQIISQETADCTV--KAVLQL-LSA-HNVTELQTDNGPNFK--------------NQKMEGVLNYM-
GVKHKFGIPGNPQSQALVENVNHTLKVWIRKFLPET----TSLDNALSLAVHSLNFKRRGRIG--------
MAPYELLAQQESLR----
>HmGINGER2-1
-----NGLHLGALK--TYKKIIE-GYANIARKAVEIFISFCPTCNLNKRQLKKAPL------QPIIS--TGFLQRLQIDLIAMESK-
PDKEFRYIGHVVDHFSKFHILFPMRNKTALETA--NNLKRFVLSYFGLPAILHSDNGSEFV--------------NDVIKSLLLTW-
KTQIVNGSPRHSQSQGLVEQGNFTIERLISSREYDSGQ--SNWSEWLVEIQCKFNV---------------------------------
>1731
HKR---NGHLNTSS--LQEMVRK-KMVYGVEKVVFKPDAVCKTCMLAKIHVQPFPKT-----TRSRA--
EELLDMIHSDLCGPFSTPSLAGSKYFLTFIDDKSRRIFVYFLRKKDEVFTKF-VEFKKLVERQTRKIKCIRSDNGGEFV------
--------NNVFDDYLKAH-
GIARQLTIPHTPQQNGVAERANRTLVEMARCMLLQSELGEALWAEAINTAVYLRNRSTSRALQ--------KTPMEEWTG----
-----
>Tnt-1
HKR---MGHMSEKG--LQILAKK-SLISYA---KGTTVKPCDYCLFGKQHRVSFQT------SSERK--
LNILDLVYSDVCGPMEIESMGGNKYFVTFIDDASRKLWVYILKTKDQVFQVF-
QKFHALVERETRKLKRLRSDNGGEYT--------------SREFEEYCSSH-
GIRHEKTVPGTPQHNGVAERMNRTIVEKVRSMLRMAKLPKSFWGEAVQTACYLINRSPSVPLA--------EIPERVWTN-----
----
>Hydra1-2
HEI---LGHCNYED--ISKLQNV-VKGMKITGKIDKSNLNCKICTQGKFVQSRNRE------PDTRA--
KAALELVHTDLAGPIDPEAKDGFKYALAFTDDYSGAVFVYFLKAKSDTARAT-EKFIAD-TAPYGRIKCVRSDNGTEFT--
------------AKEFQSLLSKN-
GIRHETSAPYSPHQNGTAERNWRTLFEMARCMILESGLPKMLWTYAVMTAAVIRNRCYNSRIH--------QTPYYALTG----
-----
>ISTv1
--------------------------------------------------------------VRYWA--LSPNLIWHADVHYFRG----
VRGQYIYGIIDDFSRKVLACIQIPDLAATTA--NVAASA-FIEFGAPYCFWTDNGSENR---------------GAFQDLLDLW-
GVQWRHTDSHSPYQNGKIERFWPTLERCQ-------------SIQAIPAFIWEYNNNTPHEDL--------------------------
>CoDi6.3
HCA---LGHPCEAT--TRATAKAFGVRL------IGQMKPCKDCALSKAKAKKISKV-----PVKRA--SKPGGRLCIDISSPSTK-
SVGGKCHWLLVVDDCTDYAWSFFLNKKTNDIMI--ALIKEL-KQAYDIDVKIRCDNSGENN--------------ALQRSCKQEGL-
GITFEYTAPNTPQQNGRVERRFPTLYGRVRAMLRDVSVNKRLWAEAANTATDLDNMLLKQGET--------TNSFHKFFG---
------
>Cer7
HG----KYHTNEQQ--TMEATREKAWIPCLRRQVKKIIGKCVKCQRYNRAPMKYPNMADMPSFRVRR--
SRPFENTGLDYFGPMTFDGSTESCWGCVLSCATTRLTHIELVQQCSTKAFI--NAIRRF-
VSERGIPDRIVSDNAPQFLGQQILNRASKENALDKDILEYLGNT-
SIEWSFITPYAPWQGGMYERMMRSIKQAIYKSIGKNILTLDDLETVMKEVSGIINSRPLTYVT--------EGAGTTICPKDFIN-
---
>Cer10-1
HEI---NGHLPEQY--TLSALKTMYWIPLCKALVRSVISNCIKCKKVFGLPYPYPNTKQLPGCRTEP--
SKPFAKVGLDYLGPVEYQESIGKAYVLLYTCLVTRAAVLRVVPDATTESYL--MALRTI-
FHQVGVPSEVHSDNAAIFKLGAKMTDIREGSEIDEIFTCFLASQ-
EIKFIYITPWSPWQGGVYERIIGLLKHQIHKICGDQKLDFFSLQYVVSSAQAMINNRPLVAHA--------RSPNDMITLRPM---
---
>Bel
HVS---YLHTGVDA--TFTNLRQQYWILGARNLVRKAVFQCKSCFLQRKGTSNQIMGELPI-PRVQA--
SRCFQHTGLDYAGPIAITPRIGKAWFSIFVCLTTKALHIEVVSELTTQAFI--AAFQRF-
IARRAKPTDLYSDNGTTFHGGKKTLRLAIQQAKDEELAGFFANE-
GISWHFIPPSAPHFGGMWEAGVRSIKLHMKRILGSKALTFEELSTVLTQIEAILNSRPLCP----------TGDNSLDPLTPAHF---
-
>ApBIN1h
--------------------------------------------RQVRRATGAAVN-GRSA-PCVSA--
SATPLRNRVDYAGPLQMKSRVFKIYIAVFVCFTTKAVHLEVVSDLSTDAFF--AAFDRF-
VARRGLPNEVFSDCGTNFVGADRQMRSLINSAEGQAAIG-NARA-
TCDWHFNPPSAPHFGGLWEVAVRSTKRLLVRVIGNHIFTYEEFSTILARVEAVLNSRPLTPAS--------TDPHDLDCLT-------
-
>Kobel
HST---VGHLGRNS--ILAKVREKYWIYGASQLAKRVARSCVTCQRYHAAPCEQVMAELPA-DRVVA--
VSPFTFCGMDYFAPLTVVKR----YGVIFTCFSSRAVHLEIAHSLETDACI--NAIRRF-
```

```
MARRGPVRSIRSDNGTNLVGSEKELRHALEGLDQSRMNDVLCAE-
GIEWHFNPPAASHFGGVWERMIRSIRKILYSLLREQPRDDETLSTLFCEAENIVNNRPLTTTS--------SDPNDLLPLTPNM---
--
>Tas
HEE---LLHAGISS--ILAKMREAYWIPRGRQAIKRALNNCFHCRRWKSKPFQMPPMPPYPPERVSK--
HPPFENTGVDYIGPFTIRITQNDGSVYSLVS--QLELYISKSPLTYPDQVSF-SVYVRF-
VSRRGLPKRMLSDNGTQFVWARSVLSVSKVHQTDNAILDYCAAH-
NIQWSFITPLVPWQGGIYERMVGPVKGSMKKTIGWKRLTQEELQTLTTDIEAVVNCRPIIPLTSEN--
TTVLRPVDFLLPHGN------
>Zebel
------LLHPGSER--VLAELRRQYWVLRGREAIRKHQHTCRDCQFWRAKPQTPQMADLPSSRLQLY--
KPPFYSTGVDCFGPFTVGRRQEKRWGVLYKCMTTRCVHLDLLEQLDTDAFL--LSLRRF-
IARRGKPMELLCDNGTNFVGGDREL-RETFNAMAPKLQEQLAEQ-
RIRFRFNPPSAPHFGGTWEREVKSVKSALRVILREQSVPEAVLQTLLVEVEGILNSKPLGYISSDVADLDPVTPNLLLMGR
RDAS----
>Tamy
HRR---FCHANHAT--VVNELRQTYWILSLRDAVKKVLHQCQWCRTRKMKPQMPPTGDLPVERLRYG--
SHPFTCTAVDYFGPMFVGRRKEKRWGALFTCLTTRAVHLELVPSLSTSSMI--MALRRM-
SARRGTPRVIYSDNGTNFIGANHELREEIGKLKKNELIDAANQE-
GIRWKFIPPGAPNMGGAWERMVRTVKTALSAILNERSPPEEVLHTLLTEVEHTVNSRPLTHLSVNPEDEESLTPNHFLIG
RSCGS----
>Purbel
HE----LGHHGVAV--TTAKTRRKYWILQGYRLAKTVKHRCVTCRAAECRRETQIMANLPSCRLAPF--
TPPFHYTSCDYFGPYLVRNKKAKHYGIIFTCLNTRAVHLEMATDCSTMEFL--QALRRF-
IAVRGQPAQFLSDNGTQFVGAERELREMVRGWSERELKDFCAEK-
RVVWKFVTPGAPHQNGCAEAMVKSCKFALKRAIGEQVLTPLELYTCFLEVANLVNERPIGRVGNDPDDGGYLCPNDLL
LGRSSSK----
>Cigr-1
HNN---LGHPGFSR--LYHFIKTRN-LPFSSEEAKIICRQCKTCAEVKPRFYKPASR-----NLIKA--MRPWERLSMDFKGPVK---
-GRNSYLLIIVDEYSRFPFVFPCKRMTSAVVI--QCLTNL-FCLFGFPAYVHSDRGLSFI--------------SHEVKVFLTER-
GIATSKSTPYHPQGNSQCERVNQTIWRTVKLLLHGKLMAEEQWEEVLPQALHAIRSLVCLATN--------
ETPHERFFNFSRRS----
>ApBIN1j
HIR---LLYIGPQG--LLAHIHRKYWPIRGRSLARQTVHKCMICFRFKPRMLQPVMAPLPS-VRDTR--
CRAFERAGVDLCGPINIKVTPLKHYIAVFVCMVTRAVIWSLFEIC--------LRMRSF-LCYSGAKKILSSWASDLSK-----------
---ESKLNDQLSEL-GIEWQFILPSAPHSGGLWESAVKSAKQHLVKSSNGSLLTYEETSTLLCRNEAVLNSRPLTATS--------
SDPSNVLTP---------
>Gypsy
HN----RAHRAAQE--NIKQVLRDYYFPKMGSLAKEVVANCRVCTQAKYDRHPKKQ--ELG-ETPIP--SYTGEMVHIDIFST--
-----DRKLFLTCIDKFSKYAIVQPVVSRTIVDIT--APLLQI-INLFPNIKTVYCDNEPAFN--------------
SETVTSMLKNSFGIDIVNAPPLHSSSNGQVERFHSTLAEIARCLKLDKK---NDTVELILRATIEYNKTVHSVTR--------
ERPIEVVHP---------
>Tom
HEK---LLHPGIQK--MTKLFKENHYFPNSQLLIQNIINECRVCNLAKTEHRNTKM--PFK-VTPSP--GHCRDKFVIDIYSS------
-EGKHYLSCIDIYSKFATLEQIKTKDWIECK--NALMRI-FNQLGKPTLLKADRDGAFS--------------SLALKQWLESE-
GVELQLNTAKTG--VADVERLHKTINEKIRIINSSKN---EIKLGKMENILYIYNHKTRHDTT--------QTPAIFLYAGQP------
>Zam
-----NSNHRGIDE--TFLHLKRETYFPNMKNKISELIRNCETCLKLKYDRQPQNI--VFE-TPETP--SKPLDIIHIDIYTI-------
NNNFNLTIIDKFSKFAAVYPIPNRNGINCI--KAIKNF-FSQFGLPKKLIHDQGVEFC--------------NDIFRKFCSQY-
NILLHVTSFQQSSSNSPVERLHSSLTEIYRILLDTRK---TDHEEIMSETVITYNNAIHSTTK--------HTPFELFNGRTH-----
>DpCIN1
--------------------------------------------------------------RRKFVQ--
KGINDTWQIDLVEMIPFYKEKGYRYLLTVIDTFSKFAYAVPVKSKTALDVL--SALKKI-LKISHKPKNIQSDQGKEFF------
--------NVAFGELMKKI-NIN--HYHTYTHLKASICERFNRTLKNRMWKIFSMQGN--HTWIYILHDLLKDYNNSYHRTIK----
----MRP---------------
>DrCIN1
--------------------------------------------------------------RNRVVV--
YGMDTQFQADLVDMTAYSTENDGNKFMLTIDVFSKYAWSRVLKNKSGVEVT--KAFESI-
LEEGRVPRKLQTDQGKEFF--------------NKHFQDLMKKH-DIN--HFATATDLKASVVERFNRTLKSRMWRFLTATNS--
RRYINVLQDIMTGYNNSYHKTIK--------MRP---------------
>SmCIN1
```

```
------------------------------------------------------------TRSYKT--
MGVDDLWQMDLMEMIPFKINKGYKYILTCVDIFSRFARAIPTKTKSANEMA--EAIKEL--FIDGKPRNLQTDLGKEFY-----
---------NSKVREIIKGI-----NHYSVFSQFKAAHVERFNRTLRDRLKKYFVYKGN--KTWINVLQKAIYSYNYSPHRGLN------
--MRP---------------
>HTLV-2
HG----LTHCN---------QRALVSFGATPREAKSLVQTCHTCQTINSQHHMPRG------YIRRG--LLPNHIWQGDVTHYKY---
KKYKYCLHVWVDTFSGAVSVSCKKKETSCETI--SAVLQA-ISLLGKPLHINTDNGPAFL-------------SQEFQEFCTSY-
RIKHSTHIPYNPTSSGLVERTNGVIKNLLNKYLLDCPN--LPLDNAIHKALWTLNQLNVMNPS--------KTRWQIHHSP-------
-
>BLV
HK----LTHCN---------SRALSRWPNSAWDPRSPATLCETCQRLNPTGGGKMR------TIQRG--WAPNHIWQADITHYKY---
KQFTYALHVFVDTYSGATHASAKRGLTTQMTI--EGLLEA-IVHLGRPKKLNTDQGANYT-------------SKTFVRFCQQF-
GISLSHHVPYNPTSSGLVERTNGLLKLLLSKYHLDEPH--LPMTQALSRALWTHNQINLLPIL--------KTRWELHHSP--------
>WDSV
HG----VSHTHKGG--IMSYFSKFWTHPKASQTIDLILGHCQICLKHNPKYKSRLQG-----HRPLP--SRPFAHLQIDFVQMCV--
--KKPMYALVIIDVFSKWPEIIPCNKEDAKTV---ILMKDI-IPRWGLPDQIDSDQGTHFT-------------AKISQELTHSI-
GVAWKLHCPGHPRSSGIVERTNRTLKSKIIKAQEQLQL--SKWTEVLPYVLLEMRATPK-KHG--------
LSPHEIVMGRPMKTTYLS
>ApGINGER2a
HSA---IGHGGRNR--MVSELKN-NYCNITNETIMVFLKLCAECHKKTVCSKVGLAP-----KPFLQ--
PTFNSRAQVDLIDMQSQ-RHNDYRFIMNYQEHLTKYIVLKPLKSKCAEEIA--YNLIDI-YTLFGAPTILQSCSGKKFV--------
------TSIINKLNVIWDEIKIDYGWPKHSESQESVERTNRDVQEMLSAWMTKKDS--LDWPSALKFIQFKKNRTFHSGIG-----
---TSPYEAMFGCTARIGYES
>BmGINGER2a
HDAHLATGHGGRNR--LVKGINN-KYKNITTEAVMLYLSLCEACQKKG--KKKGIAI------KPTT--
AEMNSRCQMDLIDLQTS-PDNEFKFISNYQDNL-NFVVLRPLKTKTASEVV--KQLMDI-FTLLGAPS-SNVINDREFA--------
------NKIIQELANMRDGVKIIHGKPRHSQSQGSIERANQDVQNMITTWLITNQT--TKWS------------------------------------------
---
>HsSCAND3
HNTHLSIGHGGRTR--MEKELQAKYKNI-TKEVIMLYLTLCKPCQQKNSKLKKVL-------TSKSI--
KEVSSRCQVDLIDMQLN-PDGEYRFILHYQDLCTKLTFLRSLKSKRPTEVA--HALLDI-FTIIGAPSVLQSDNGREFS---------
-----SQVVSELSNIWPELKIVHGKSQTCQSQSSAEQT-EDIRKRIFSWMQTNNS--SHWTEFLWFIQMSQNQPYHRSMQ------
--QTPCESAFSSEAKLGLSH
>FcSCAND3
HNTHLSIGHGGRTR--MEKELQAKYKNI-TKEVIMLYLTLCGPCQQKHSKLKKVL-------TSKPI--
KEVSSRCQVDLIDMQLN-PDGEYKFIMHYQDLRTKLSFLRSLKSKRPKEVA--HALLDI-FTIIGAPSVLQSDNGREFS-------
-------SQIVYELSNIWPELKIVHGKPQPCQSQSSMNQTNDDIQKRIISWMQTNNS--SHWAEFLWFIQMIQNQPYHRGMQ---
-----PNPFEGTFSSEAKPGLCH
>MeSCAND3
HDTHLNIRHGGRTR--MEKELQPKYHKR-GYNVI--FNSLQTLSAKEFKVTKKCI-------VSKPI--KEVNSRCQVELIDMQLN-
PDGEYKFIMSYQDHQTKLSFLRPLKSRRCEEVA--HALLDI-FTIIGAPSILQSEHGRDIL--------------
SQLVNELNMMWSKLKIVHGKPRPCQGHGSVVQPKEDIQNMIISWMQKNSS--LHWSEGLWFIQMMKNRPFHRDTQ------
--HSPREIVFDSEK------
>HsKRBA2
HDTHLSIGHGGRTR--MLKELQGKYGNV-TKEVIVLYLTLCKQCHQKNPVPKRGL-------
APKPMTFKDIDSTCQVEILDMQSS-ADGEFKFILYYQDHSTKFIILRPLRTKQAHEVV--SVLLDI-
FTILGTPSVLDSDSGVEFT--------------NQVVHELNELWPDLKIVSGKYHPGQSQGSLEGASRDVKNMISTWMQSNHS--
CHWAKGLRFMQMVRNQAFDVSLQ--------QSPFEAMFGYKAKFGLYS
>BtKRBA2
HDTHLNIGHGGRTR--MLKELQGKYGNV-TKEVIVLYLTLCKQCHQKNPVPKRGL-------
APKPMPCKDNDSRCQVEILDMQSN-ADGEFKFILYYQDHLTKFIILRPLKAKQAHEVV--SVLLDI-
FTILGPPSVLESDSGMEFT--------------NQVVNELNEVWPDLKIVLGKSHPGQGQGFLERVSRDVKTMLSAWMQSNHS--
HHWTEGLRFMQMVRNQAFDGSLQ--------QSPYEAMFGCKAKFGLYS
>LaKRBA2
HETHLNIGHGGWTC--MLKELQGKYGNV-TKEVIVLYLTLYKQCHQKNPVPKKGL-------
VPNSTTLKDVNSRCQVEILDVQSN-ANAEFKFILCYQDHLTTFIILXPLKARQVHEVV--SVLLDI-
FIILGTPRVLESNSGIEFT---------------NQVNEVSEVWPDLNIVSSQYHTGXGQGSKEQXNEDVQNTMSARMQSNHS--
HHWTEVLXFMPMVGNQVFNVSLH-------------QSLQCKVTLGLYS
>Tor4a
--------HTGIDQ--TFSKIAKKYVWPRMNMAIREFIRTCDFCQKDKPNFHPNRA--PVL-SFRTP--NGPYEIYGFDLITLPP--
TDFGNKYVMVMIDFFSKFAYCEPLKSRNSSYLL-SKFRNVI-FKNPFFPKMVVLDNAREHS---------------ELAKFMTNN-
```

```
NIEPHFTPPRHPSSNGQVENFNRTLKSRLRAKCKYE-----NWDLVLQEILHDINASEHSVTK--------
KSPFFIQSGAECPHNIFD
>Pyggy
HESLS-SGHTRGERLPIKDNLPRRFFWAEHDRSIRRYVRNCDVCEKTNYREKVLRTI-----TNTRP--EDIEKEISIDFIKGLPT-
-SEGITYLIVVTNRLSKGSIFILLPNIKTETVV--RAFLR--VVAYYLLEAITSDRGSFVS----------------VLERLCEIL-
KIRRRLLISFYL-TNGSTERINSV-EAYTRAFISAQI-----DASLCSIAQIAINSRDATLTG--------VALFFLQHRYNVD-PLQL
>Ty1B
HRM---LAHANAQT--IRYSLKNNTITYFDVDWSSAIDYQCPDCLIGKSTKHRHIK-GSRL-KYQNS--
YEPFQYLHTDIFGPVHNLPKSAPSYFISFTDETTKFRWVYPLHDRREDSIL--DVFTTI-LAFIASVLVIQMDRGSEYT--------
-----NRTLHKFLEKN-GITPCYTTTADSRAHGVAERLNRTLLDDCRTQLQCSGLPNHLWFSAIEFSTIVRNSLASPKSK-------
------------------
>Ty4
HKR---MGHTGIQQ--IENSIKHNHYEESLDLIKEPNEFWCQTCKISKATKRNHYT-GSMN-NHSTD--
HEPGSSWCMDIFGPVSSSNADTKRYMLIMVDNNTRYCMTSTHFNKNAETIL--AQVRKN-IQYVRKVREINSDRGTEFT---
-----------NDQIEEYFISK-GIHHILTSTQDHAANGRAERYIRTIITDATTLLRQSNLRVKFWEYAVTSATNIRNCLEHKSTG---
-----------------------
>ZFERV-2
----------------------------------KIKVNCPNCRKAIDSRFNDY-------GSIKT--
DHENDEISMDFAGPIRPQTKKKNCYFLLITDNASEYTQIFPCKQATEEVVI--ESLQQW-FKFRGVPNRVRADGAFNLS-----
---------GRKVEQFIKDS-GAKLIKSVKYQPSSNGIAERRIKEVKNYFCKNSD------LPWDEQLIDCLKFVNY--------------------
------------
>ISEc
--------------------------------------------------------------------SGPNQKWAGDITYLR---
TDEGWLYLAVVIDLWSRAVIGWSMSPRMTAQLAC-DALQMA-LWRRRPRNVVHTDRGGQYC-------------
SADYQAQLKRH-NLRGSMSAKGCCYDNACVESFFHSLKVECIHGEHFIS---EIMRATVNYIECDYNRWRRHSWC--------
LSP---------------
>ISLa
-------------------------------------------------------------RFFER--SRPNQMWQTDIFSFRLG---
GQAAYLLAFIDDYSRYMVGLGLYRRQTAENLL-EVYRRA-TGEYNCPAEMLTDNGRQYG-------------TTRFEKELKKD-
RIKHIRSQPHHPMTLGKIERFWKTIWTEFLDRCQFDC---ETAQQRITLWIKYYNHQRPHQGG--------LCP--------------
>ISAf
-------------------------------------------------------------KRFAY--
DQINELWQGDLSHGPTINGKAQKTFLIAYIDDCSRFVPYAQFFPSKFDGLR--IVTKEA-VLRCGKPKRIYSDNGKIYR------
--------SEVLQYACAEM-GITLIHTQPYDPQSKGKIERFFRTVQTRFYPLLELDP---PKSLEELNERFWRYHRKPHASLD-------
-KTP---------------
>Koala
HSR---LGHPAFPI--VKRIVQSHKL--SCLDVESNNISVCDACQKAKSHQLPFGL------STSEV--HSPFELVYSDVWGPAPT-
SVGGKKYYVSFIDAYSRFTWIYLLKFKSEVFEKFHEFQNHV-ERFFDRKIKIQTDWGGE----------------YQKLNSFFNKI-
GILHHVSCPHTHQQNGSAERKHRHIVEVGLSLLAHAFIPLKFWDEAFSSAVYLINRIPTKVLQ--------RSPLEQLYN---------
>Vitico1-1
HLR---YGHLNVKG--LKLLSKKEMVFGLPKIDS---VNVCEGCIYGKQSKKPFPKG-----RSRRA--
SSCLEIIHADLCGPMQIASFGGSRYFLLFTDDHSRMSWVYFLQSKV----------------------LRTDRGGEFL-------------
SNDFKVFCEEE-GLHRELTTPYSPEQNGVVERKNRTVVEMARSMMKAKNLSNHFWAEGVATAVYLLNISPTKAVL------
--RTPYEAWYG---------
>SIRE1-4
HQR---FGHLHLRG--MKKIIDKGAVR-GIPNLKIEEGRICGECQIGKQVKMSHQKL-----QHQTT--
SRVLELLHMDLMGPMQVESLGGKRYAYVVVDDFSRFTWVKFIREKSETFEVFKELSLRL-QREKCVIKRIRSDHGREFE-
-------------NSRLTEFCTSE-
GITHEFSAAITPQQNGIVERKNRTLQEAARVMLHAKELPYNLWAEAMNTACYIHNRVTLRRGT--------TTLYEIWKG-----
----
>Copia
HER---FGHISDGK--LLEIKRKNMFSDQLLNNLELSCEICEPCLNGKQARLPFKQLK----DKTHI--
KRPLFVVHSDVCGPITPVTLDDKNYFVIFVDQFTHYCVTYLIKYKSDVFSMFQDFVAKS-EAHFLKVVYLYIDNGREYL--
------------SNEMRQFCVKK-
GISYHLTVPHTPQLNGVSERMIRTITEKARTMVSGAKLDKSFWGEAVLTATYLINRIPSRALV--------KTPYEMWHN-----
---
>Zeco1
HKQ---FGHASADR--LQRLMHSSNKDKECFAILQQIVDDCEICKKYKRTKPKPAV------GLPMA--SEYNETMAVDLHELE-
-----PSIWYLHIIDHFTRFSAGSIVKTKKSSEI-VNSFIHTW-ISVHGAPQKLYSDNGGEFN-------------NQEIRDMAENF-
NIEIRTTAGYSPWSNGLLERHNQTLTEIILKVKENG----CDWHTALDWALMAKNSMLN-VHG--------YSPYQLVFG---------
```