

4. Page LK, Anchor C, Luy E, Kron S, Larson G, Madsen L, et al. Backyard raccoon latrines and risk for *Baylisascaris procyonis* transmission to humans. *Emerg Infect Dis.* 2009;15:1530–1. <http://dx.doi.org/10.3201/eid1509.090128>
5. Kelly TG, Madhavan VL, Peters JM, Kazacos KR, Silvera VM. Spinal cord involvement in a child with raccoon roundworm (*Baylisascaris procyonis*) meningoencephalitis. *Pediatr Radiol.* 2012;42:369–73. <http://dx.doi.org/10.1007/s00247-011-2151-y>
6. Wise ME, Sorvillo FJ, Shafir SC, Ash LR, Berlin OG. Severe and fatal central nervous system disease in humans caused by *Baylisascaris procyonis*, the common roundworm of raccoons: a review of current literature. *Microbes Infect.* 2005;7:317–23. <http://dx.doi.org/10.1016/j.micinf.2004.12.005>
7. Blizzard EL, Yabsley MJ, Beck MF, Harsch S. Geographic expansion of *Baylisascaris procyonis* roundworms, Florida, USA. *Emerg Infect Dis.* 2010;16:1803–4. <http://dx.doi.org/10.3201/eid1611.100549>
8. Miyashita M. Prevalence of *Baylisascaris procyonis* in raccoons in Japan and experimental infections of the worm in laboratory animals. *Journal of Urban Living and Health Association.* 1993;37:137–51.
9. Xie Y, Zhang Z, Niu L, Wang Q, Wang C, Lan J, et al. The mitochondrial genome of *Baylisascaris procyonis*. *PLoS ONE.* 2011;6:e27066. <http://dx.doi.org/10.1371/journal.pone.0027066>
10. Sorvillo F, Ash LR, Berlin OG, Tatabe J, Degiorgio C, Morse SA. *Baylisascaris procyonis*: an emerging helminthic zoonosis. *Emerg Infect Dis.* 2002;8:355–9. <http://dx.doi.org/10.3201/eid0804.010273>

Address for correspondence: Guangyong Yang, Department of Parasitology, College of Veterinary Medicine, Sichuan Agricultural University, 46 Xinkang Rd, Ya'an, Sichuan, 625014, People's Republic of China; email: guangyoyang@hotmail.com

EMERGING INFECTIOUS DISEASES

Free Online RSS Feed

in PubMed Central

Ahead of print

CME Peer-Review

podcasts wed

GovDelivery



Novel Divergent Rhabdovirus in Feces of Red Fox, Spain

To the Editor: Rhabdoviruses (family *Rhabdoviridae*) are enveloped single-stranded negative-sense RNA viruses belonging to the Mononegavirales order. The International Committee on Taxonomy of Viruses recognizes 11 genera (*Cytorhabdovirus*, *Ephemerovirus*, *Lyssavirus*, *Novirhabdovirus*, *Nucleorhabdovirus*, *Perhabdovirus*, *Sigmavirus*, *Sprivivirus*, *Tibrovirus*, *Tupavirus*, *Vesiculovirus*) (1). In addition, many recently described rhabdoviruses remain unassigned. Rhabdoviruses contain 5 major genes, encoding for nucleoprotein (N), phosphoprotein (P), matrix (M), glycoprotein (G), and RNA-dependent RNA polymerase (L). The *Rhabdoviridae* family includes pathogens of various animal species, humans, and plants. Viruses of the genus *Lyssavirus* are the most relevant to public health because they can cause rabies. Bats are the driving force within this genus; foxes and various other species of wild carnivores also can be infected with lyssaviruses and transmit them to humans and dogs (2).

During a viral metagenomic survey, conducted as described previously (3), of fecal samples collected from 4 red foxes (*Vulpes vulpes*) that were found dead in Álava, Basque Country, Spain, we identified the complete coding sequence and the partial leader and trailer sequence of a novel rhabdovirus, tentatively called red fox fecal rhabdovirus (RFFRV; 15,541 nt, GenBank accession no. KF823814; online Technical Appendix, <http://wwwnc.cdc.gov/EID/article/20/12/14-0236-Techapp1.pdf>) by mapping 8,287 of the 56,519 sequence reads in the sample of a red fox. A proportion of obtained reads contained sequences that were $\geq 99\%$ identical to mitochondrial DNA of *V. vulpes*, which confirmed

that the sample was collected from a red fox.

The obtained sequence of RFFRV was partially confirmed by specific primers and Sanger sequencing of PCR amplicons. Five major and 3 minor open reading frames (ORFs) were identified that had a genome organization similar to that of other rhabdoviruses (Figure, panel A). No significant hits were obtained by BLAST analysis (<http://blast.ncbi.nlm.gov/Blast.cgi>) of N, P, M, and G nucleotide and amino acid sequences, which was reported previously for novel divergent rhabdoviruses (4).

Predicted N, P, and M genes of RFFRV consist of 1,629, 2,490, and 813 nt, respectively, encoding for 543, 830, and 271 aa (online Technical Appendix Table 1). In addition to the absence of significant hits observed by BLAST analysis, no significant sequence homology was observed with known rhabdovirus proteins in pairwise alignments. Furthermore, no conserved motifs were detected in N, P, and M genes of RFFRV that are commonly observed in rhabdoviruses. However, intergenic regions between all major ORFs contained relatively conserved motifs that could be transcription termination/polyadenylation sequences (A/U) CU₇, similar to other rhabdoviruses (5). Adjacent to this termination signal was a stretch of conserved nucleotides that might function as a transcription initiation signal (online Technical Appendix Table 1).

The amino acid sequence of the G protein consisted of 669 aa and contained an N terminal signal peptide (1-MYHLIVLLVMLGQRAVA-17), a noncytoplasmic domain (aa 18–646), a transmembrane domain (647-ITAILMPLLSLAVVVGIMCC-667), and a cytoplasmic tail of 2 aa, similar to other rhabdovirus G proteins as predicted by using Phobius and TMHMM (<http://www.cbs.dtu.dk/services/TMHMM>) (6,7). We predicted 3 potential glycosylation sites in the ectodomain at positions 38–40

(NKT), 554–556 (NAS), and 592–594 (NIS) using NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc>).

Between the G and L genes, a complex intergenic region was present that contained 3 ORFs of 246 nt (7,413–7,658 aa), 231 nt (7,716–7,946 aa), and 459 nt (7,893–8,355 aa), of which 2 were overlapping frames (U1–3). Additional ORFs between G and L genes were detected previously in other rhabdoviruses (8,9). We detected transmembrane domains in the amino acid sequences of all 3 additional ORFs, suggesting they might act as viroporin (8,9).

The L gene of RFFRV contained 6,591 nt (2,197 aa). We detected several conserved domains and motifs, including RNA-dependent RNA polymerase, mRNA-capping region, mRNA capping enzyme, and virus-capping methyltransferase. Alignment of the deduced amino acid sequence of the L gene with the L gene of various other viruses belonging to the Mononegavirales order by using MAFFT version 7 (<http://mafft.cbrc.jp/alignment/software/>) and subsequent phylogenetic reconstruction by using a maximum-likelihood tree (WAG+F+I+G model with 100 bootstrap replicates in MEGA5 [<http://www.megasoftware.net>]) suggested that this virus belongs to a novel genus of the *Rhabdoviridae* family. In addition, pairwise identities of the deduced amino acid sequence of the L gene of RFFRV with that of other rhabdoviruses of the *Rhabdoviridae* family were only $\leq 35\%$ (online Technical Appendix Table 2).

Because the fox was found dead and no tissue samples were collected, whether RFFRV played a role in the animal's death is unknown. In addition, multiple attempts to isolate this virus on various cell lines of eukaryotes (Vero E6, MDCK, CRFK, N2a, and BHK cells, primary fox kidney cells) failed because of the absence of cytopathic effects and viral replication by quantitative reverse transcription PCR,

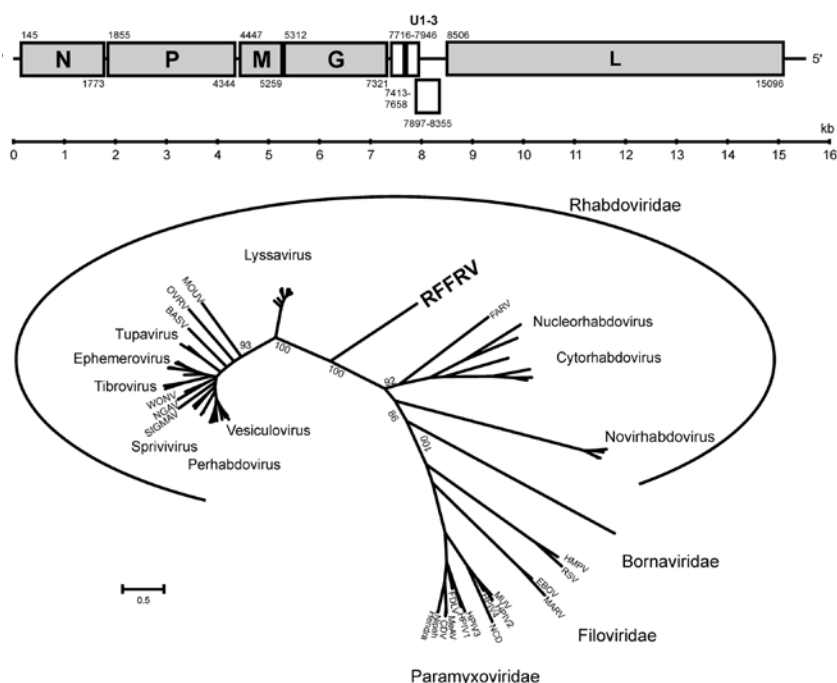


Figure. Genome organization and phylogenetic analysis of RFFRV. A) Genome organization of RFFRV. Indicated are the locations of the major ORFs (including the positions of the first and last nucleotide) and 3 theoretical minor ORFs between the G and L genes. B) Phylogenetic maximum-likelihood tree using the WAG+F+I+G model and 100 bootstrap replicates in MEGA5 (<http://www.megasoftware.net>) of the deduced amino acid sequence of the L genes of various viruses of the order Mononegavirales. G, glycoprotein; L, RNA-dependent RNA polymerase; M, matrix; N, nucleoprotein; ORF, open reading frame; P, phosphoprotein; RFFRV, red fox fecal rhabdovirus. Only bootstrap values in the close proximity of the branch of the RFFRV are indicated. Scale bar indicates nucleotide substitutions per site. Viruses and GenBank accession numbers are shown in the expanded figure legend online (<http://wwwnc.cdc.gov/EID/article/20/12/14-0236-F1.htm>).

despite a high number of reads in the original sample. The fox might have acquired the virus through spillover from a small prey, such as a bat, and additional studies are required to elucidate the prevalence, original host, and pathogenic potential of this novel virus.

Acknowledgments

We thank all researchers and institutions for their invaluable help during sampling and for providing the specimens used in this study, especially Patricia Lizarraga, Laura Elorza, Ricardo Gutierrez, and Luis Javier Chueca.

This work was funded by the European Community's Seventh Framework Program (FP7/2007–2013) under the project "European Management Platform for Emerging and Re-emerging Infectious

Disease Entities" European Community grant agreement no. 223498; the Virgo Consortium; and the Niedersachsen-Research Network on Neuroinfectiology of the Ministry of Science and Culture of Lower Saxony, Germany. In addition, this research was funded partially by the Basque Government through the research group on "Systematics, Biogeography and Population Dynamics" (ref. IT317-10; GIC10/76).

A.R.-G. holds a postdoctoral fellowship awarded by the Department of Education, Universities and Research of the Basque Government (ref. DKR-2012-64) and was awarded a short-visit research grant from the ConGenOmics Research networking program of the European Science Foundation to visit the Department of Viroscience, Erasmus Medical Centre and develop the current research project.

**Rogier Bodewes,¹
Aritz Ruiz-Gonzalez,¹
Anita C. Schürch,
Albert D.M.E. Osterhaus,
and Saskia L. Smits**

Author affiliations: Erasmus Medical Centre, Rotterdam, the Netherlands (R. Bodewes, A.C. Schürch, A.D.M.E. Osterhaus, S.L. Smits); University of the Basque Country, Vitoria-Gasteiz, Spain (A. Ruiz-Gonzalez); National Institute for Environmental Protection and Research, Ozzano dell'Emilia, Italy (A. Ruiz-Gonzalez); and Viroclinics Biosciences, Rotterdam (A.D.M.E. Osterhaus, S.L. Smits)

DOI: <http://dx.doi.org/10.3201/eid2012.140236>

References

1. International Committee on Taxonomy of Viruses. Virus taxonomy: 2013 release [cited 2014 Feb 9]. <http://www.ictvonline.org/virusTaxonomy.asp>
2. Matha IS, Salunke SR. Immunogenicity of purified Vero cell rabies vaccine used in the treatment of fox-bite victims in India. *Clin Infect Dis*. 2005;40:611–3. <http://dx.doi.org/10.1086/427700>
3. Bodewes R, Ruiz-Gonzalez A, Schapendonk CM, van den Brand JM, Osterhaus AD, Smits SL. Viral metagenomic analysis of feces of wild small carnivores. *Virology*. 2014;11:89. <http://dx.doi.org/10.1186/1743-422X-11-89>
4. Palacios G, Forrester NL, Savji N, Travassos da Rosa AP, Guzman H, Detoy K, et al. Characterization of Farmington virus, a novel virus from birds that is distantly related to members of the family *Rhabdoviridae*. *Virology*. 2013;10:219. <http://dx.doi.org/10.1186/1743-422X-10-219>
5. Albertini AA, Ruigrok RW, Blondel D. Rabies virus transcription and replication. *Adv Virus Res*. 2011;79:1–22. <http://dx.doi.org/10.1016/B978-0-12-387040-7.00001-9>
6. Coll JM. The glycoprotein G of rhabdoviruses. *Arch Virol*. 1995;140:827–51. <http://dx.doi.org/10.1007/BF01314961>
7. Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol*. 2004;338:1027–36. <http://dx.doi.org/10.1016/j.jmb.2004.03.016>
8. Gubala AJ, Proll DF, Barnard RT, Cowled CJ, Crameri SG, Hyatt AD, et al. Genomic characterisation of Wongabel virus reveals novel genes within the *Rhabdoviridae*. *Virology*. 2008;376:13–23. <http://dx.doi.org/10.1016/j.virol.2008.03.004>
9. McWilliam SM, Kongsuwan K, Cowley JA, Byrne KA, Walker PJ. Genome organization and transcription strategy in the complex GNS-L intergenic region of bovine ephemeral fever rhabdovirus. *J Gen Virol*. 1997;78:1309–17.

Address for correspondence: Rogier Bodewes, Erasmus Medical Centre, Department of Viroscience Dr. Molewaterplein 50, 3015GE Rotterdam, the Netherlands; email: r.bodewes@erasmusmc.nl

Ngari Virus in Goats during Rift Valley Fever Outbreak, Mauritania, 2010

To the Editor: Ngari virus (NRIV) is a single-stranded RNA virus belonging to the family *Bunyaviridae*, genus *Orthobunyavirus*. The genome comprises 3 segments, the small (S), medium (M), and large (L) segments, which encode the nucleocapsid (N) protein, the 2 glycoproteins Gn and Gc, and the RNA-dependent RNA-polymerase, respectively. Sequence analysis showed that NRIV is a reassortant between Bunyamwera virus (BUNV) and Batai virus (BATV), both from the genus *Orthobunyavirus*. S and L segments derived from BUNV, and the M segment derived from BATV (1,2). NRIV is more virulent than BUNV and BATV and is associated with hemorrhagic fever. NRIV was first isolated from *Aedes simpsoni* mosquitoes in 1979 and from humans in 1993, both in Senegal (3). During 1997 and 1998, humans were affected with hemorrhagic fever diseases in Kenya and Somalia that were caused by Rift Valley fever virus (RVFV) and by NRIV (2,4).

In 2010, during an ongoing RVFV outbreak in Mauritania, we collected 163 serum samples (62 from camels, 8 from cattle, and 93 from small ruminants) (5). RVFV RNA was isolated from serum samples as described previously (5). Further molecular testing of the samples was conducted by a SYBRGreen-based real-time reverse transcription PCR (RT-PCR) adapted from a conventional RT-PCR and based on generic primers (bun_group_forw 5'-CTGCTAACACCAGCAGTACTTTTGAC-3' and bun_group_rev 5'-TGGAGGGTAAGACCATCGTCAGGAAGT-3') that target a 250-nt sequence of the S segment of Bunyamwera serogroup members (6). Real-time RT-PCR was performed in a CFX 96 real-time PCR system (Bio-Rad, Hercules, CA, USA) by using 5 µL RNA with a QuantiTect SYBR Green RT-PCR Kit (QIAGEN, Hilden Germany) in a final volume of 25 µL. Cycling conditions included RT at 50°C for 30 min and 95°C for 15 min, followed by amplification with 44 cycles of 95°C for 15 s, 55°C for 25 s, 72°C for 30 s, and 77°C for 5 s. A melting curve analysis was then performed starting with 95°C for 60 s, and a temperature gradient was conducted from 68°C to 94°C in increments of 0.2°C.

Of the 163 serum samples tested, 2 samples from goats resulted in a positive signal with cycle thresholds of 23 (sample 51) and 28 (sample 65), respectively. Both samples showed similar melting peaks at ≈78.2°C and shared the identical partial nucleotide sequence of the S segment. The sequence belongs to the Bunyamwera serogroup, but the short partial sequence was not sufficient for accurate virus determination and identification. For this reason, both serum samples were used to inoculate cell monolayers of Vero E6 cells that were assayed for virus replication. Only sample 51 displayed a cytopathic effect after 72 h and was further analyzed. We isolated the viral RNA from cell culture with TRIzol reagent

¹These authors contributed equally to this article.

Novel Divergent Rhabdovirus in Feces of Red Fox, Spain

Technical Appendix

Technical Appendix Table 1. Characteristics of RFFRV genes and intergenic sequences*

Gene	Length, nt	Length, aa	Conserved intergenic sequence between 2 ORFs†
N	1,629	542	<u>UAG-60nt-ACAAAAAAACUAACCUCAGCUAUG</u>
P	2,490	829	<u>UAA-75nt-UCAAAAAAACUAACACAAGCUCGAAUAUG</u>
M	813	270	<u>UAA-23nt-ACAAAAAAACUAACCUAAAACUAGGUUAUAUG</u>
G	2,010	669	<u>UAA-70nt-ACAAAAAAACUAACACAAGCUAUG</u>
U1	246	81	<u>UAA-29nt-ACAAAAAAACUAACACUACCUCGUGUUUAUG</u>
U2	231	76	NA
U3	459	152	<u>UAA-116nt-ACAAAAAAACUAACACUAUCUGUGUUUAUACAAACAUG</u>
L	6,591	2,196	NA

*NA, not applicable; ORF, open reading frame; RFFRV, red fox fecal rhabdovirus.

†Start and stopcodon of the flanking genes are underlined.

Technical Appendix Table 2. Pairwise amino acid identities between the L protein of RFFRV and other rhabdoviruses

	RFFRV	LNIV	BEFV	RABV	IHNV	RYSV	PRV	SIGMAV	VSIV	TBIV	BASV	FARV	LBV	SHIBV	OZEV
RFFRV		27	31	34	23	27	33	30	32	31	30	29	35	35	35
LNIV NC_007642	27		25	30	23	35	25	25	25	25	25	33	30	30	30
BEFV NC_002526	31	25		45	24	24	60	57	60	57	52	25	46	46	45
RABV SRV9 AAT48626	34	30	45		25	28	46	44	46	44	40	29	86	87	88
IHNV L40883	23	23	24	25		22	24	24	23	23	20	27	25	26	25
RYSV NC_003746	27	35	24	28	22		25	24	25	23	24	30	28	29	28
PRV HM566195	33	25	60	46	24	25		60	74	54	50	26	47	47	47
SIGMAV Q410979	30	25	57	44	24	24	60		59	52	48	23	46	45	45
VSIV AAA48441	32	25	60	46	23	25	74	59		53	51	27	47	46	46
TBIV GQ294472	31	25	57	44	23	23	54	52	53		52	24	43	43	42
BASV JX297815	30	25	52	40	20	24	50	48	51	52		23	41	41	41
FARV HM627182	29	33	25	29	27	30	26	23	27	24	23		28	28	28
LBV JX901139	35	30	46	86	25	28	47	46	47	43	41	28		92	86
SHIBV ADD84511	35	30	46	87	26	29	47	45	46	43	41	28	92		87
OZEV FJ905105	35	30	45	88	25	28	47	45	46	42	41	28	86	87	

Red Fox Fecal Rhabdovirus, Partial Genome (GenBank Accession No. KF823814)*

GGATATCAAGTCCACCAATACCTTATATGTGCATGATCATGCACATCACATTCTACG
 CATACGACTCCCAGGGAGTACACTAAAAAAGACGTAACACAGACTTTGAATTACGT
 CAAGTCTAAGTTTTAAATTCGGTTAATTTCAATGGATCACGATAACGAAAAGCCAAT
 CTCGTACACTTCAATAGCGGAAGTTCCTGATAATGTTGCCATTGGGAGCACCATTTA
 CATTCAAGGTGAGCCCATCATCTATTTTGGAAAATCCGCTGCAACAGGAATTACGCG
 GAAAGGGGGGGCACAGAAAGACTGGACCAAAGACATGATCCGTGGAGTGAGAGTG
 TTCCTGCCCCAGACTGATGCTAATCTGCTCAATCTCATAGCCGGGGAAACCGAAGCC
 CCTGAGCTGGAGAAGTACACCATCCAAGATCCGGAAAAGAAGGGCATCTTGAAGAA
 ATTTGAAAGCAAGTGGGAGTTTGCGAATTGGGCAAACCTGCTGGTCGACTTGCAA

GCAACACAGGGAACATCCCCAAAGGGAGATTCCCATACTACTCTGCTCTATTCTCGA
TAACTGCCATTAAGGGAGCCCCTGTCCTGGCCCCTGCCATGAAAGACCTCGGGGACC
CTGTTTATGTGAAAGCTCCCGATGACCTCCACCCACCTACAGGAGACATAGAATGGC
ATGGTGATAAAATTAGTGTCGACGAGGCAGCCTACATAGGATATGGAGCATGGCTG
ATCATGCCTAGATTCACTATCAAAGCTGAATCCAAGAAAGATGAAATTGCAGCCAG
CAGCAAGGCATTTGACACTCTTAGGCGGTTGCTACCTGAGATCACCAAGCCACAAGT
GTTGGTATCTGTGGTGACGCAGCTCAGATTGGCATATCACGGAACGCTGGTTCCCGG
GTCTGCGTACCTTGCTGCAGAAGTAGCAATGAGAAGGGCAATGAATATAGAGTATG
ACCTAAAGGCTGACAGAACGGAGTGCAAAGCCGGGGAACACTTTCCAGGTTGTCAA
TTAAGAGTCCTGCAGGATATCCCCCAATACGACTCAGGCTTTTGGGGCTTTGGGCAA
GTTGGGCTAGAAATGGCCGGATACTCTGCTCTTAATATGCTGCATGCCGGCCTGGAC
ATCTACGGGAAAACCATTGCTGACCTAAGGATGCTAATCAACTGGAGGTGTTACGA
CAACTACATCGCAGATGAGATTAAAGAAGGCCCTTTGTTAGCAGATGACCCCTGGA
GAGCAGCGTCTTACTTACTAGCCCCTAATATAAGAACGCCACTAAGCATGGGGAAG
CACTCCATTGTAGCGTATTTGGGGCTATCTATCCAGTCAGCTGCAGCTAACATCAGC
ACAGGGGCTCCATCCCCACCAGAGGGAGTGAAAATGAACGAGCTGATCAGAAAGA
CAGTGTACGATCATGCTGTTGCAATTGTTAGCGAGTGGGACAATGACAGGCTTCAAC
CAAGCACCGTCACCACTGTAATGATTGGAGGACAAGTTATCCCCTTCAAGGGAGTTG
ACCCTAAGCGAGTGAATGACTTGTCAAGAATGTTCACTCAGAGACAGACCCCTTTGT
ATGAAGTACCACCCACAATCAGAGGAGGGAGCGATCACCCCTCTGTTTCCAGCGTC
CATACTAGTAGTAGAaGAGATGaTGAAGGGAGCTGGGAAGGAGGCAACGAAGAAGA
GTTGCTGAGAAAACCTCCATGAAAGGCGAGGCCAGTACGAGGAAGATACTAATTTAG
GGGGATTTTATTCAGCGACATAGACATCATCTTGAAATATTATAAGCTTGCTTTTTGA
TTCTTTAGCTTAGTATTAGTGCCATACAAAAAACTAACCTCAGCTATGGCTAATCG
TGTAAGGAAAATTCATTCTGTTGTGTCCGTGAATGATCCGGCCGAAGCTTGGCAAGT
TGCTCTTCAATATTGGGCTAAGCGGATCTTGTCAACAGATAAGGAGCTGCTCTGCCG
AATTACGAAGTCAATGGTTGCCACATGTCTGTGGTGTACGGGAACGTTAAACTCGA
CTGCGACGCCGTTTGGAGAGTCTGGGCCCTCCTTGCATTAGCATGGATGAAGAGCCC
TGTAGACGGAATTGCCATGTTGGCCTTACTAGAATTTGGCGCCAAACACGTACAGAG
GCTGCGAGCCGACCTGAGCAAGATCGGAGACCCTAGAATCGATGATTTTCGAGCGGT
TAACAAATACCAGCATTGCCCGTGTTCCAAAGAAATTCTTCAAAGAAAACAGCATC

ACTGACGACACCAAGAAGAAGCTAGAAAAGGCTGCAGCAATCTTGTCGGATCTAGG
GAAGAAGTGCAGCTATACTGGATTTTGTCTGCTATGCGGATGATAGCAACGTAGGAG
ATATGATGTCGCGGGTGTAAACGGCTTCTGGGGCATTGGGATCTGAGGATCCTATTG
GTGTGTGGAAGCCGTTGGAGTTAGATGAGAAAGCACCGTATCTTTTTGATGAAGTGA
AGCGCATTGGAGGAATCTCTGCAGAAAAACCAAGCCCACCTGAAAAGAAGATGTGC
AGGTTATTCGGGAAGAAAGGGAAATTTGCAGCGCCCGAGACAAGCTATAAGAGTGC
GCACCCGTTCTCTACTCAAGTATCTTCCAAAGAACCACCAAAGTTGAAGGTCCGAGG
CGTGGAACTCCCATCTCTTGATGAAGTGATCGGTCTTCTGCCCAAGGGAGCAGAGGC
TGCAGAACTTCTGGACGACACCAAGGAGTTGGAAGGCGCATCAGCTGAACCAGAGC
TAGCCGATCATGCCGAGCCGCAGCTGACTCAGTCACGCTCTGATTATCAGGGAGATC
AGAATTTCACTTCTAGCGTGCTCAAGGCTAGTCCCCACATGGACACAGTATAGACT
GCGACTCTAGCCAAATCGGGGCATCAAACCTTAGCGCTAGGAGCGCAAGCCTCTCC
CCCGAATGGAAGGAACCGGATGGAGCCACCAAGGGCGATGATCCTGAAGAGGTAG
AGTGCGCCAACGAGATAGTTGGAGGAAATAGAAGGGTTAGCGAAGAAGGCTGCTGC
TTGAATGAAGACACCAGAGAAGAAGGCAGCCTTCAATCTGATCAACACCCTGTGAA
AAATTACAGTAGCTGGGCAGAGTTCCGAGCGTCTTGGAAATCAGAAGTCTCTCCCAA
GCACGCGGAGGCTGCTGGATCGGCAATAGGCCACTCCCATAGCCCACAAGCAGAGG
CAGTTATGGATTCTGTAGGAGTTGCTTCTTCTACCATCAAACCTCTAAAGTCCGGAG
TAACGAGGCCTATGCAGACAGATGAGCCGCATGCCATCCTGGAAGCAGGATCGCAG
GGTATTACCAGCCTGTACCCTCCCCTCCCAGTTGTACCCGGAGTATTACTGCCGAGT
TCTCTCAAAGCTCCCTCCCCCGAAGGCGAAAGGGAAATTCACCGAAGAGTTCGG
ACTGATGTTACGCGGGATACGAAGAGGGCTAGAAGAAAGAGGCCTCAATTGGGAGG
CCAAATGGTTTGACACGCATCTGGAAGGCATTGCAAAGTTTGGCGACCCCGATACTA
TTGGGTTTCAGCTAGGTAGCTACTTTGCTCAACTCATTACAAATCCTTACCTTAAGAG
GGAGGAAACAAAACCGTCACCAAAGAGGATCACGCGCCTCCACCTCCATCGGCAC
CTCGAAAAGAAAGCGTGAATAAATTCAGTCAAGCTAGGCCCCATAGCCAAGTGCCC
ACCAAAGAAGATCCTGCACCGCACCCCTCCTAGAGCAGAAGAACCCGACCCAGAAG
CTGGATGACTGGGAAACCATCAAGGGTATATCTGCCAGGATCGGATGACGTTATGG
AATTCGCACCAATGAGGCGGAGCTTGACAACCTTATTACATCATGTGTGCTGGAGG
ACAAGTTCACTGAGCCATATATCCTTAAGCCTAAGGAGTTGTCCAAAGAACAGCTAA
AGAATCTCTTAGAGGTTGTGTCACAACATGGACAGAAGGCTTCCCAGCTGTTATGTG

AACATTTGACATTAAGAACTACAAGAGCATCTCAGCTTTGACAGCTAATTGGGCAA
TGAAGATAAAAGCAGAGATGCCAAGGTCCATGGGAATATACTTCACAAGGCTAGGA
GAAGCATATGGAAATGCCTTCTGGGAGGCGCTGCAAGTTGCTACTGCTGGCAGAGA
AGGGATTGACAAGTGGATCAGGGCTGCACTAGTCGTGAAGGCTAACACTGCGAGAG
GGAAAGTTCCTCCCCCTGGCTGGTTTGTGACAACAAGAGCAAAAGACAGCGACAAC
CCCAGCGTCCGTGCAAGAGCAAGTGCCTGTTTAAATTCTTAAATGAATAATATTAG
TCTAACCCCTGTAGTTGTAGTCTTGCCTCTATATTAGCTTACAAGTAGTTCTTCATC
TCATCTTCATTTCAAAAAAACTAACACAAGCTGCGAATATGAGCAATTTTAGGTATA
CTCTGATGAAATTTGCATCAAAGATGTCTTTGACATCTAAATCTAAGTATAAGGTAC
TTGGAATTGGTGACGAATTGGGTCAATCGAATGTCAACATTATCCATGAGGGCGAG
GAAGACCATACGTCGATTTATTCAGAGAGCCCTTCATCTTCTAAAAAGAAAACAAGT
CGCACAATGAGGCCTTGGGTTTATCCTACATATCCTGAGCAAAAGAACATAGTTAGA
GGAACACTGCGCGGAGCCAGGCCTAAGAAAGTTGTTCTGTCTATTA ACTTAACTGGA
GTCTCTGAGCCTCTTGATTTTCCGGAAGTGGCCACAGTCATATCAGACATTCTTGAC
GGAATGGAAATGTCTGCAATGAGAAAGCTGGTCTTGCAAATGATGATCTTGTCAACT
AGACCAATGGGGAGGTTTTCTGATGGGTGTTACCGATACATCTTCAGTAATTGTTTT
AGCACCTCATGCTTCCGTCTAGTTGCCTAGGAGAGTTCATGCATGATTCAGGGGAT
TTCTGTTGACCACGTTTGATGACAAGACATACTTTGGATCATAACAGGCTGACATTC
GAAGACGCGATAGGCTCCAATCATCAATACCCGCTCCCCTTGTGGTTTCACTTGTCTT
ACACCAACGCGCTCAAGCCAGACTTCACCCCGGACTAGCAGTCTGCAGAGCTAAA
GGGAAGTTCTTGATCTACCTTGATGACAAATACTTCAGCTTCAAAACGTATAGCGAC
AACATCGTGCTATTGCTGATGGGCCGTGCTCCTAAGGCAATTAAGCAGAAGTAATTA
CCGCTTGAGTTTGGATTTTACAAAAAACTAACCTAAACTAGGTTATAATGTATCA
TCTAATTGTTCTCCTTGTTATGCTAGGACAAAGAGCAGTAGCAGAGCCAGCTGACAT
TGCACGAGCCGCCAAGATCATAACAGTATGCAGAGGAAAGGCTGAGAAATAAGACG
GGTCTGAGGATTGAGCTGTCATCGAGAGATCCGGCTCAAAGCGCAGCATCCCCACC
TGATGTGCCTGCGACTAAGCAGCCTGCTACCAAGCCCGCTGCTACTACAGCTGCCGG
AACTTCAGCCCCGATTAAGGTGGGGTTGCCTGAGTCTGAATACCTGCGTCAGGTTGT
GGAGCCTTGGTTCACCGTCTGTGAGAGCAAGGTCCAAGGACACTGACCGATGTAG
TGCATCGAATCCTCAGGAAAGGGGATGGCAAAGCTCATAGAATGCTTCGAGGAACA
CCCAATGGAGATGATCACTCCGTTTATACAAAATTCGGATCCGGGATCTCCCCTGCA

CCAAGCCGACTGGTAGTCCTTAGTGAATCATTGCAAGAAGTGCAAAGGCCGCAACT
ATATCGAGTCTCGAGGCTAAGGTCAAGCTCTACGCCCCACCGGGTAGTACTATGCGA
AGTTATGTCTGCATTCTCACCTCTGAATGTGGAAGAGATGGAGTGCACGGGAGTACT
GGGTAAGCTGTCTGACACGTCCTCGGTTACGCTCAATGTGGTGGACCCGGGGTATGT
ACTGACAATGATCCGCGTCTGCTCTTGCTCAGCAATGAGGGTCAAACGATGGACCTG
GACTAGCTTCTTCAACGTGAAAGAGAGGCTAGAAGACCGCTGGAGGATACCCCCTG
ACCCTAAAGAATGTGAGTCTCATTGCGAACCAATTTTCGCAGGAGAAGTGACTTCGG
CTGTAAGAGGGAGTGTTCCCTGATTATTGCTATGCTTGGATGTCTACGTGTGAAGTTC
AGGGAGATGTGTACCAAGTCTCTTTAGGAACAGCTAAATTCCATAGGTTCTTGAACC
AAATAAGAGCCCCCTTTATTGTGGACGACCCTTGCACCCCATCGGCGCCATGTAAAG
GATCAGGGGATAGTCTCGTCTTAGCAAAAATAGAAGAGGATAACCCTCGATTCACC
ACCATGAATGGGGAATTGACGCCAAGGTACAATTGGGATACGCACAAAGTTTATCA
GGTGCACCTGCCAGGAGTCACAACACTAGCATCTTGGACGCGTCATGCGGTTTTCTACA
TGGGGGTTATGTGTATTATCAGCTAATGAGCGGCCGCATCGTTAGCGTGAGCGTAGG
CACGCTACAATTTGGTGATAAGGTCAAGCCTCCGATCTGCACAGAATGGAAGGGGC
CGTATATGCCGTTTGTCTTACCTGATTCTCAGGTTGCCTCGACTTCTGAACAGTTGCG
GCAGGACCTGCTCCACTGCCAGACTAGAAAAGAAGTCGTCCTTAATGCACTGGCAA
CTAAGAGGCTCCCATCTATAACGCTGTTTGAAGGCTTGGGTTACAAAGGATCGGAGT
CATACGGTCTGGTGTCAAGGAAGGGGCTCCTGTATGCAGCACCTTGCCCATCAGTAG
AGTACACAGATCTTGAGCATGTAGAAGGAAACATATGGATGGTCGTAAATAATGGG
AGAGAAGTTGGGTGCCTGGATGGAGGGTTAAACTTTGCTGTCAAGAGCGGTTGCGT
GGTTAATCCGAACGCATCCGTCTCTATCCTTCTAGGCGAGTGGAAGGTAATATCTGA
TAGGGACGGCAAACCTTTAGCAGAACCAATCCCAAAGGCAGGTTGGGGGAGCATTC
CAGCTCTCGAAAACATATCGGCTGCATTCGGGGGATTACTTAGCCAGTTTAGAACAGC
CGCCTCTGTGGGATGATGGGAACGGGCCATTATCATTCTACTAGCAACTCAACAG
GAGACCCGGTCATTCATTCCGGGGCATCATCACTGTGGAGCTCTATGTCCCTAGCCA
GTAAAATCACTGCTATTCTGATGCCTCTTTTATCCTTAGCAGTAGTGGTAGGGATAAT
AATGTGCTGCCGCAGATAACTTGCATATCATGTACAAGAAAGATATTGTTACCTTCT
CTCTGATAGCTTTTAGTTAATCAAATTAACAAAAACAAAAAACTAACACAAGCTATGA
GTTGTTTAACTCAAGATTTGAGGAAGGAGATACGAGAGCGGGTGTCTAATAGCATA
AGCCAGAATCCAAGTGTAACAACACTTATAATTATCGAAGTTATACTATGGGTCGGC

ACTTTAGCCTATATTTTCGATAGCGTTGGGGTGCCATCGGTATCTTCAGGCTCGCATA
AAGAATTCTGTAGAACGGAGACAACGATTTAGGGGAGATGGGGTCTTGCTGCCCGA
TACTGGACCTGCGTAATAACTAAATAGGTCGATCCTTTAGGTGGTACAAAAAACTA
ACACTACCTCGTGTTTATGGAAGAAAAGTGCTCAGATTCCGATTACTTTAAAGAGTT
GACTGACGCAGCTATAGAGGGAGTGTGGGCATCTCCTTTATACCCAATTACATTGGC
TACGGTCATAATATTCCTTATACTGCTAATATTTGTAGTAGCGTGGAGAGCAGCTGT
GATTGCAAAGATACGACATCGTATAGATGAGTCCACAGCACGGCGCCTCCGCCCTGAT
TTCGGGGGTTCGACCCTGATCACTTCTCGCTTATCATTGAGTCATACCCGATTCTTG TG
GGCAATTATGCAGCCACACTCATCTTAACTGTCACCCTAGTCTTGTCTTTTAGTCTTG
CTACAATTATTTCTTTTTTTCATCTGTAAACA ACTTAACGGATATCCGAGGGCGTTCT
CATCTATCACTTAGGTAATGAGGTCTCCGAATTTGCGTCCCATGCTTTAGCAGCAGTT
TGTAGCGCACAAATCGACATCTGCCAGACTAACAAGGGAGCTTGAACAGTTTCGCCTA
AGTGGCCAAGTGACCGAAGTAACACCGAGCTCGGGGGCAACGGGGCAGCTGAGCA
TACCACAGCGCTACAGAATGCTGTTAGTAGAAAAGGACATGTTGGAAAATGAAATG
TTCATTGCAGAGCATTTCGCTAGTCTAATCTCTAACAACTACTTTGATTTCTTCATCC
TCTCTTGTCTTATGTTTGTGAACTGCTTTGACTCAGTGATTTCTCACATAACAGCCTA
AATATAACCGCAAATTATCTCTCAAAAACAAAAAACTAACACTATCTGTGTTATAC
AAACATGGAAAATCGTTCCTTTTTGATCGCATTCCGGACTCTGTGCTGCACAGCCC
CCTTATCCTTAGAGAGGAGTTGTGCTTTTGCCGTCCAGACAAGGAAACCAGTGACTA
TCCAGCCATTCCGCAGCCAGACGAATTCACAATCTCTCGCCGCCCTGAGTTGATAGA
GCTCAGGTCACTGTCACTCACTGGGGCAATCTCAGGAGTAGGGTGTATCAAATGGCT
CGCTTATATGAGATTGAATTGGATCGAAGCCAAGGATATTATGAGGCTCATGGCTGA
AGAATACCAATCATCCACAATCCACCCAGAGATAGCATCCGACTTCAATAACGTGT
GTGAAGGTCTCTTCTCAAGGCTGCATGACCAGTGGGGAGAAAATCAGCATCCCAGT
GCGCTTCAGGACATTTGGAACGTTTTTGTGGCGCAGTACCAATCAGACGCCATTATT
GACAACTGGACAAAATACCGAATACTCTCAGCAGCTCTCTTATGGTCTAATCATAAA
GAGAGAGGCAACCAATCAGAATTTCAAACAGCATCATTGCTAAAGCAGGATGGGG
AGTCCGCGCTATAAATGGAGTGGAAGTCATAACTCACAAGTTGTTAACTAAAAGAG
AACGCTTGGGAAAGCTAACTGTAGGCGCATACTGGTGTCTGCTGGAGCTTGCCAACC
AGCCTCCAGATTGTTTCAGAAAATCACTATTGCTGTGTGAAAAGGACCTGTATTTGC
AGCGCTGGATTTCACTCTTCTTGTGTA ACTGTCAAACGCCTTCGGAGGAATCCCTTT

GCCTCAGCCAAAGCACTTGATCAAAAATGTACCGAATGGGTGACATCATCCTCGCCAC
AGAAGGAAACGAAGGGTACAATGCAATCAAGACAATAGAAAGCATCATGGTTCGGTT
CTGCACATAAGCAACTAGAAGACGAGTCAATTTCCCCTACAGGTGACAAATTCTACG
AGTTTATCACGAAAGAAGGGCTTTCAGCTTGCCATCTAGAATCATCGTCATCTGCTA
TGAGAAAAATCATCCATATAGGGGAGATAGTAACCCCTGAAGTTGCCATAGAGATG
GCAGGATGCTACAGACACTGGGGCCACCCAATCATCAATCCTGTTGGAGGTCTCACT
GCAGTCCGAGAGAACGCAACTGCCAGCTTCTACGAATGAAAGGCTCATGATATC
ACTAGCTGCTGATCTCAACTATTTGTTGTTGAGGAGTTATTTTGAAGAGCATGGCAG
GTGGCCCCCGGGAGTGAGATATGAAGCGGCGCCAGAGATAGCCAAAGACCTGTTCA
ACAAGTGGGTATCAACTAATCAATTCCCTGGCCCTACCTCTGCTAGTCAAGTCAGGA
ACAGCTGGTTTTTCGTAACATATGACAGCTTATTCGACAAAAATCAAGAAATTCCAA
TCCTCAGCCTAATTTTCAGACAAGTCTCACTCAGTTGGACGATCCGCCCTGTCAACTA
TGTGCCTCAAAAAGAATCTGCTGCTGAGTCCTTCACGGAGGGTGCTTCAATCAACTC
TCAGCTATGCAGAGATCGATGTGAACAAATTTTTGGACTCTATAGACAGCACAGAA
AATGGTTTAAGCAACGACGACTTAGTCATCCTATTGCGAGAAAAaGAGCGAGAGCTC
AAGGTAAAGGGGCGGTTCTTTTCCCTAATGACGTACAAATTGAGGACTTACTTTACA
GCAACCGAATACCTAATTGCAAAGCACATTCTGCCTGTACTCCCTGAAATTACAATG
ATGCAGGGGCAAATTGACCTGTGGAAGACTTTCAAAGGTGCGGTACGCACTGTGAG
TCAAGAGAAGTCTACACATCACATGATTCATGTAGATTTTCGAAAAATGGAACAACCTT
CCAGAGAGAAGAGTCTACAGCTCCGGTTTTCCAGATTATGGACAGAGCCTTCGGGTG
GTCAAATGTCATATCAAGGACTCATAACTTCTTCTCTAGGTGCTTCGTGCGGCTATGCT
GGGAGAATAGATATGTTCCCTATCGGTCTGACAGACAACCTGGCCGTGGTGTGGACC
GGACACAAAGGAGGCTTTGAAGGGCTAAGACAGAAGGGATGGAGTGTCGTAGGGG
CTTTGCTAATACGGCACGTGATGAGACTCACTGGTCTGCATGGGAAAGTGCTAATTC
AAGGAGATAACCAGGTAATCATACTTGAGTACCCTCTTGCCAGCTCATCGAATACTT
CATCGCTTGCTCTCGAGAGGCATAGGCATTCGAAAATGACAACCTCACTTCTTGTCTG
TGTTCTCTGAATTGTCTAAGAGCATCGGATTAAGAATTAAGCCAGAAGAGACATGG
ATATCCTCGAGGATGGTTTACTACGGGAAGTTCCTGTGATCGACGGATCGGCAAGA
GGAATGGTGCTTAAGAAACTCTGCAGGATCTTTGCAATATCAAATGACCTCACTCCT
AGCATCTCAAATAGCATCTCTTCTCTGCACACGTCATGCATTGCCGGCTGTATCCAAT
CATCATAACGCTGTCCCAATTTACATGTTTGCTATCTTTTATGGCTTCTTCCTGCTCCAC

GACTTCTTCGAATACAACCCCTTGTCCATGGAGCCTATGATCTTCACATTCAGACGG
GCTTTATTCGGAACAACCTGCTGTTTCGAGACATTACCAAGGTACGAGCGCCACAACC
AAGGACTCTATCATCATTTCTGTAGCAAGAGACTGGGCACCAGCAAGCAATAAGCA
GCAATCGGCTCTCGGACTAGACCTCCTGACGCGAGATTCATCTCTTGGAGGTCTTGG
AGGAGCATCCATTTTAAAGTATATGATCAGGCAGTTTCCTGATCCTGTACAGAGGG
CCTATCATTCGCAAAAGTGGGGTTAAGCAAGGCAGAAACACCTGAGCTAAAGACCA
TTTTCTTCTTATGGAGGCCTTCCCATCAAGTCAGGCTCAGTAGAACTCCTTCTAGA
AGCACCCATGTTCGATCAATCTGCCATCCTCTACTCGTCCTGTAAATGTACTGCGGCG
GTTTGTGAAGAGAAGATCACACAGAACATGGACACGCTGGTCAACAATCAGCTTG
TGTATCAGGCGCTGTTCGATGAAGCAAGATATGCGACCGCCATTCTTGAAGATGTGTG
CATCCCTACAGCCATTCTTCCCCGATTTGCTTCTTCTCTGTACAGTGCCACCCCAT
AGGAGTGGCAGAAACAATTGTTGAAAATTCCTCGGGACTAAAAGTGTGCCAGAA
TTGCTCTCAGAGAGGGTGCAGGGCAGCTGGGAAAGAACTAGTAGAAAGCGAAGTT
GAGTTGATAACAACGTATCTAACAGTTCTCAAAGTGTCTGCTCTGGAGCATTGGGGT
TGTTCTAGCAAATTGGCTGACAGCTATAGGTCGAAAAGTTGGGGAGGCCAAATTTG
GGGAGTCACAGTGCCTCACCTGCAGAGCAGTTTTACTTAGAGTGGCCAATGACAA
GCGGGTGCTTCAAGGAAAGCCTTGCTTGGATGAATGATTTTGCCGAAGaAACTGTGA
AAGTTAGCTTGAATGAAaGAATGATCTATGGAGACACCTCTCCGCGATCAGCCGGAC
CTATGAATCCTTACTTAGGTGGTTCGAACTAGTGAGCGGTCTCTGTCCAATCATGGC
TAGAGCTCGAAACATCATCTTCTATTGCAAAGAGACTTGTAGAGCTAATAAATGCAT
GTGGTTGGGCATATCAAAAGGGTTCGCACTTGCATCACGCAATAAAAGAAATGCTC
ACACAATTGACTGGATTAGTGATTCCGAGCGATGCAGTTTGTAGAATCACAGAGTCC
GGGAGTTTGGGACATAGGTTCTCAGATCCTAGAGTGTCTTCTGGAGCAATGTCAGCC
ACCAACTATAACACTGCCACTCATATTGCGTTCAGTACAAACGACCTAGTGAGGTTG
AAGCGCGGGGAGGACAACCTATATGGTTTTGTCCAAGGGCTATTCGTGTATTTACA
AGCATTATTGGGGAGCTAGTAAGGAGAGGAGATAATGGGTCTCCGACTGCCACCT
TCACCCTAAATGCGAAGAATGCATGATACCATGTCCAGATGTTTCAATGGAGCTTTC
AAACCCGGTAGTCCCGCTTTATTCAATGTTTTTATCACCTCGCCTCCCGAATGATTT
TTAACGTTGCCGGAGAATCTCTATGTTGACATTGATCCGTTCAAAGTAAAGCCTGAT
CAAATCAAGACTATTGTCAACATCGTTCCTGCATACAGTACAGTCTCTGTTGACACC
CTCTCGGGTGAGGAGATATCAGATGCTATATCCAATGTTGCAGCAATCATAGCAGTA

ACAAAAACAAATCCTCAAGGAAGATGGACTCTGGGGATACGAACTGCCATCCCGGT
GTATGCAGGTTACAAGTGCTCGTGGAAGGTACTTAAGAGCAAGATTGTATTCTGGAC
TTTTGCTATGCTTGCATTCCGGCACATCAATGAGTCAGAGGAGCATGACTGCTGTGA
GGCTTGTCAGCACATACTTAGAAAGCAAGCAGCAGATCACTGGAACCTGCATAACC
CAACAATGCTGGGAGGGATGCTGAGCAACACGGACATCATTGTCGAGTTGGCTAAA
GAGCAGATTGCGCTCAGCGCAGGATTTCCGATATCGCCTATGGAACTTTCAAGCGCA
GCAGGACAGTCATTATTCGAGCAGTGTGTCTCCCATGGAAAAGAATTCTCACTACCA
GAATTCCCGAAGGGAGTGATAGCTCTCCCGAGAGACATGACAGGAGGTGGAACTCC
GGCATATTTAATCGCCTGCATGATGTGCTATGCCGGCTTCTTCAAAGGAGCCAACAC
GGGACCACTAAAAGCACTTGCAGGGCTGACCACTGAGATGAAAACATCTGTTATAG
CAAGCCAAGCATTCGAAGAAAGCGAAGGAGTCTGGAGCAGGCTTGTAGATATGCTT
CACGAAAACACATCAACCTTGTGCTGCGTGCTTGACATTGACTTCAAAAACACTAACA
AAAAGGGGTCAAGACGTAGATCATATTCTCCCGGGAAGCATACCTGCTATTTCACTT
GATACCGTCAGAGGGGCATACTGTATATGGAACACAGCATCAACAACAAAAGGGCC
TGAATCTTCTTACTTGGAAAGAGGAGGGAGAGCCTGAGGTGACACAAACCCCTATCA
TGCATCTATTCCGCCCCATCTTAATGTCTGCTAACGGTTCAATCAAGATTCTGTCTGT
CTTGAAAACTATCCAGCTTCAACCAAAAAATGTTGCAGTGGTCGGAGATGGTAATG
GGGGATTTTTGAGGGCACTATTGGGTCTCTACCCTACAGCTGAGTTCTTTTTCAATTC
CCTATCTGAAATGTCCAACTACGGCGATCAAGGGGAGTTCACAAATTACCCCGCTTG
TGTGCTAGACTCTCCAAGTGATTGGGGAAGGGTGATCAATCTCGGAGAAGTCTTAAC
ACTGCCTTCAGACTTGCCTCAGAGAGGTGGCCTTCCGAGGTAACGAAATACTGGAT
GAAGAAATGCTTTGACCCAGACCTGCTGGTGTGCGATGCAGAGAGCTATGATTCTAT
CAAGAGGTCAGAAATCTTGAGCAACCTCCTCAAGGTATGCGGCCAGCGGTCTCACG
TCATAATCAAAGCCCACTTGTCAAGGATGCAGGACGTCCAGAGCTTGGTCAATGCGC
ACTTGCATTCACACAATGTGATGTACTTGAGAGCCCCTTACAGCAACTGGGGGAAGT
TTGAGGTGTACTTGTGGTACACTCGAATTGTAGAAAGAAACGATATGAGGTATGCAC
CAAAGATACATATCTCATGTGGACTAGAGCTGAATTCAACGCTCACCTCAGCTCC
AATTAACAGCGTTCCACAGCTGGGTTTTCCCTCATCTTGAAAAAGACGATAGACTCC
GAGCTGTCCGGCGGATGCTAAGCCTACGACGGCACATTAGCAGGTACACTAAGATA
AGACCGCAAGAGCTGTCACGAGTAGAGAGGAAAGCAAAAGAGCACGGAGACATGG
GGTTCTGTGAATTGCTGTTGCACTACTTAGGATCCAAGTTCATATCATCTATGAACA

ATTCTGTTATCAGGGCAACTTCACATGCTGGGAGATTGAGAGCCAGCTTAAAGCGAT
CGGAGCTGATCGGGCTAATAGGGACTGCTATCGGTGCCTCATGCTTGGTGTGAGCCA
TTTCTCATCCGCATGATTACTGCTTGTCTTGGAACCTTACAAGAATGGGGTGGATGT
TGTTTTGGGCAATGATAAAGGGAATATCAAGCCAATTGTGGTCAGAGCCTGGTCAG
GTGAGTCCACTCCTGTAACTATAACCCCAAAGGAACTCAAAAAGCCTGTTCATTCTA
TGGTGAGATACTTGGGGGGCTTATTGTGGAGCTCTGGTATTGATAGAATCACTCAA
GGTTAGAGTGGCTGGTAAGAGAAAAATATGAACTCGGCTTCTCGGAAGGGTACAAT
CACTTGGCAGTAGCCACTTGGGATAACGATACCTTAGACGTAGATCCATTGGAGTGT
ATAAGATGGGAGATGGAACAAGCTGCCGTCGAAGAAATCGATTTGGGGGGATAAT
TGATGAATAGCACTTGATTTCTAAAATAGTATCAGAGAAAAACAAAAACTAAAA
AACAAAAAACATATATACATATAAAATTAGGAGCAATAACTGAAGATCAAAGAG
CATTAAGAGATCTAAGAGAAAAAGAGAGAAGGACCTGGCAAAAAGCTAAAAGA
TAGAAATAAGGAAACAAAAGAATTTTCTGAGAGGGCAGCACATAAGGATGATAGG
GATCACATGAAGGCTTCGGGAACCTTGAAGCTTGTAaaaaaaAGCTCAAAGACAAA
ATTGCAGAGAAGAGGGGAACCCAAGTTGCTCTGATTGCTGACTGGCGACATCTGAC
CATTTGCTCTTTTACTGAAAGACCAATCCTGATATAGATGACTTCTAGTTGTCTGGCT
ATGAGTGTGAGTATCGTACAAGGCATGGGTATGGGGGTATGGAGGGCAGCACCATT
GTGGGT

*Coding sequences of identified open reading frames are underlined; deduced amino acid sequences below, overlapping sequence of open reading frames U2 and U3 gene are in italics.

Deduced Amino Acid Sequences RFFRV Genes

N Gene RFFRV

MDHDNEKPISYTSIAEVPDNVAIGSTIYIQGEPIIYFGKSAATGITRKGGAQKDWTKDMIR
GVRVFLPQTDANLLNLIAGETEPELEKYTIQDPEKKGILKKFESKWEFANWANLLVDL
QSNTGNIPKGRFPYYSALFSITAIGAPVLAPAMKDLGDPVYVKAPDDLHPPTGDIEWH
GDKISVDEAAAYIGYGAWLIMPRFTIKAESKKDEIAASSKAFTLRLPEITKPQVLVSVV
TQLRLAYHGTLVPGSAYLAAEVAMRRAMNIEYDLKADRTECKAGEHFPGCQLRVLQDI
PQYDSGFWGFQVGLEMAGYSALNMLHAGLDIYGKTIADLRMLINWRCYDNYIADEIK
EGPLLADDPWRAASYLLAPNIRTPLSMGKHSIVAYLGLSIQSAAANISTGAPSPPEGVKM
NELIRKTVYDHAVAIVSEWDNDRLQPSTVTTVMIGGQVIPFKGVDPKRVNDLSRMFTQR

QTPLYEVPPHNQRRERSPSVSSVHTSSRRDDEGSWEGGNEEELLRKLHERRGQYEEDTN
LGGFYSAT.

P Gene RFFRV

MANRVRKIHSVVS VNDPAEAWQVALQYWAKRILSTDKELLCRITKSMVAHMSV VYGN
VKLDCDAVWRVWALLALAWMKSPVDGIAMLALLEFGAKHVQRLRADLSKIGDPRIDD
FERLTNTSIARV PPKFFKENSITDDTKKKLEKAAAILSDLGKKCSYTG FCCYADDSNVGD
MMSRVLTASGALGSEDPIGVWKP LLELDEKAPYLFDEVKRIGGISAEKPSPEKKMCRLF
GKKKGKFAAPETSYKSAHPFSTQVSSKEPPKLKVRGVELPSLDEVIGLLPKGAEAAELDD
TKELEGASAEPELADHAEPQLTQSRSDYQGDQNFTSSVLKASPPHGH SIDCDSSQIGASN
LSARSASLSPEWKEPDGATKGDDPEEVECAN EIVGGNRRVSEEGCCLNEDTREEGSLQS
DQHPVKNYSSWAEFRASLESEVSPKHAEAAGSAIGHSHSPQAEAVMDSVGVASSTIKPL
KSGVTRPMQTDEPHAILEAGSQGITSLYPPLPVVPGVLLPSSLKSSLPPKAKGKFTEEFGL
MLRGIRRGLEERGLNWEAKWFDTHLEGIAKFGDPDTIGFQLGSYFAQLITNPYLKREET
KPSPKEDHAPPPPSAPRKESVNKFSQARPHSQVPTKEDPAPHPRAE EDPDRSWMTGKPS
RVYLPGSDDVMEFRTNEAELDNFITSCVLEDKFTEPYILKPKELSKEQLKNLLEVVSQHG
QKASQLLCEHLTLRNYK SISALTANWAMKIKAE MPRSMGIYFTRLGEAYGNAFW EALQ
VATAGREGIDKWIRAALVVKANTARGKVPPPGWFWVTTRAKDS DNPSVRARASALFKFL
NE.

M Gene RFFRV

MSNFRYTLMKFASKMSLT SKSKYKVLGIGDELGQSNVNIHEGEEDHTSIYSESPSSKK
KTSRTMRPWVYPTYPEQKNIVRGTLRGARPKKV VLSINLTGVSEPLDFPEVATVISDILD
GMEMSAMRKLVLQMMILSTRPMGRFSDGCYRYIFSNCFSTSCFPSSCLGEFMHDSGDFC
STTFDDKTYFGSYRLTFEDAIGSNHQYPLPLWFHLSYTNALKPDFTPGLAVCR AKGKFLI
YLDDKYFSFKTYS DNIVLLLMGRAPKAIKQK.

G Gene RFFRV

MYHLIVLLVMLGQRAVAEPADIARA AKIIQYAEERLRNKTGLRIELSSRDPAQSAASPPD
VPATKQPATKPAATTAAGTSAPIKVGLPESEYLRQVVEPWFTVCE SKVPRTLTDVVHRIL
RKGDGKAHRMLRGTPNGDDHSVYTKFGSGISPAPSRLVVLSESLQEVQRPQLYRVSRLR
SSSTPHRVVLCEVMSAFSPLNVEEMECTGVLGKLSDTSSVTLNVVDPGYVLT MIRVCSC
SAMRVKRWTWTSFFNVKERLEDRWRIPDPKECESHCEPIFAGEVTSAVRGSVPDYCYA

WMSTCEVQGDVYQVSLGTAKFHRFLNQIRAPFIVDDPCTPSAPCKGSGDSLVLAKIEED
NPRFTTMNGELTPRYNWDTHKVYQVHLPGVTTSSILDASCGFLHGGYVYYQLMSGRIVS
VSVGTLQFGDKVKPPICTEWKGPYMPFVLPDSQVASTSEQLRQDLLHCQTRKEVVLNA
LATKRLPSITLFEGLGYKGSESYGLVSRKGLLYAAPCPSVEYTDLEHVEGNIWMVVNNG
REVGCLDGGLNFAVKSGCVVNPNASVSILLGEWKVISDRDGKLLAEPKAGWGSIPAL
ENISAAF GDY LASLEQPPLWDDGNGPIIPTS NSTGDPVIHSGASSLWSSMSLASKITAILM
PLLSLAVVVGIIMCCRR.

U1 Gene RFFRV

MSCLTQDLRKEIRERSNSISQNPSVTTLIIIEVILWVGTLAYISIALGCHRYLQARIKNSVE
RRQRFRGDGVLLPDTGPA.

U2 Gene RFFRV

MEEKCSDSDYFKELTDAAIEGVWASPLYPITLATVIIFLILLIFVVAWRAAVIAKIRHRIDE
STARLRPDFGGRP.

U3 Gene RFFRV

MSPQHGASALISGVDPDHFSLIIESYPILVGNYAATLILTVTLVLSFSLATIIFFFSSVNNLT
DIRGVLIYHLGNEVSEFASHALAAVCSAQSTSARLTRELEQFRLSGQVTEVTPSSGATGE
LSIPQRYRMLLVEKDMLLENEMFIAEHSLV.

L Gene RFFRV

MEKSFLDRIPDSVLHSPILREELCFCRPKETS DYP AIPQPDEFTISRRPELIELRSLSLTG
AISGVGCIKWLAYMRLNWIEAKDIMRLMAEEYQSSTIHPEIASDFNNVCEGLFSRLHDQ
WGENQHPSALQDIWNVFAQYQSDAIDNWTKYRILSAALLWSNHKERGNQSEFQNSII
AKAGWGVRAINGVEVITHKLLTKRERLGKLTVGAYWCLLELANQPPRLFRKSLLLCEK
DLYLQRWISLFLCNCQNAFGGIPLPQPKHLIKMYRMGDILATEGNEGYNAIKTIESIMVG
SAHKQLEDESIPTGDKFYEFITKXGLSACHLESSSAMRKIIHIGEIVTPEVAIEMAGCYR
HWGHPIINPVGGLTAVRENATAQLPTNERLMISLAADLNYLLRSYFEEHGRWPPGVRY
EAAPEIAKDLFNKVVSTNQFPPTSASQVRNSWFFVTYDSLFDKNQEIPILSLISDKSHSV
GRSALSTMCLKKNLLSPSRRVLQSTLSYAEIDVNKFLDSIDSTENGLSNDDL VILLREKE
RELKVKGRFFSLMTYKLRTYFTATEYLI AKHILPVLPEITMMQGQIDLWKTFCGAVRTV
SQEKSTHHMIHVDFEKWNNFQREESTAPVFQIMDRAFGWSNVISRTHNFFSRCFVGYAG
RIDMFPIGLTDNWPWCWTGHKGGFEGLRQKGWSVVGALLIRHVMRLTGLHGKVLIQG

DNQVIILEYPLASSSNTSSLALERHRHSMKMTTHFLSVFSELSKSIGLRIKPEETWISSRMVY
YGKFPVIDGSARGMVLKKLKCRIFAISNDLTPSISNSISSLHTSCIAGCIQSSYAVPIYMFAIF
YGFLLHDFFEYNPLSMEPMIFTFRRALFGTTAVRDITKVRAPTTKDSIIISVARDWAPAS
NKQQSALGLDLLTRDSSLGGLGGASILKYMIRQFPDPVTEGLSFAKVGLSKAETPELKTI
FSSYGGLPIKSGSVELLEAPMSINLPSSSTRPVNVLRRFVEEKITQNMDTLVNNQLVYQA
LSMKQDMRPPFLKMCASLQPPFPFASLYSATPIGVAETIVGKFLGKTKTVARIALREGA
GQLGKKLVESEVELITTYLTVLKLKSALEHWGCSSKLADSYRSKSWGGQIWGVTVPHPAE
QFYLEWPMTSGCFKESLAWMNDFAEETVKVSLNERMIYGDTSPRSAGPMNPYLGGRS
ERSSVQSWLELETSSSIKRLVELINACGWAYQKGSHLHHAIKEMLTQLTGLVIPSDAVC
RITESGSLGHRFSDPRVSSGAMSATNYNTATHIAFSTNDLVRLKRGEDNYMVLFFQGLFV
YFTSIIGELVRRGDNGSPTAHLHPKCEECMPCPDVSMELSNPVVPLYSMFLSPRLPNDFL
TLPENLYVDIDPFKVKPDQIKTIVNIVPAYSTVSVDTLSGEEISDAISNVAIIAVTKTNPQ
GRWTLGIRTAIPVYAGYKCSWKVLKSKIVFWTFAMLAFRHINESEEHDCCEACQHILRK
QAADHWNLHNPTMLGGMLSNTDIIVELAKEQIALSAGFPISPMELSSAAGQSLFEQCVS
HGKEFSLPEFPKGVIALPRDMTGGGTPAYLIACMMCAYAGFFKGANTGPLKALAGLTTE
MKTSVIASQAFESEGVWSRLVDMLHENTSTLCCVLDIDFKTLTKRGQDVDHILPGSIPA
ISLDTVRGAYCIWNTASTTKGPESYLEEEGEPEVTQTPIMHLFRPILMSANGSIKILSVLK
TIQLQPKNVAVVGDGNGGFLRALLGLYPTAEFFFNLSSEMSNYGDQGEFTNYPACVLDS
PSDWGRVINLGEVLTLPDDLSSERWPSEVTKYWMKKCFDPDLLVCDAESYDSIKRSEILS
NLLKVCGRSHVIAKAHLSRMQDVQSLVNAHLHSHNVMYLRAPYSNWGKFEVYLWYT
RIVERNDMRYAPKDTYLMWTRAEFNAHPQLQLTAFHSWVFPHLEKDDRLRAVRRMLS
LRRHISR YTKIRPQELSRVERKAKEHGDMGFCELLLHYLGSKFISSMNNSVIRATSHAGR
LRASLKRSELIGLIGTAIGASCLVSAISHPHDYCLFLELYKNGVDVVLGNDKGNIKPIVVR
AWSGESTPVTITPKELKKPVHSMVRYLGGLLWSSGIDRITQRLEWLVRKYELGFSEGY
NHLAVATWDNDTLDVDPLECIWEMEQAAVEEIDFGGIIDE.