

**A case for a negative-strand coding sequence in a group of positive-sense RNA viruses**

Journal:	<i>Virus Evolution</i>
Manuscript ID	VEVOLU-2019-067.R1
Manuscript Type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Dinan, Adam; University of Cambridge, Department of Pathology Lukhovitskaya, Nina; University of Cambridge, Department of Pathology Olendraite, Ingrida; University of Cambridge, Department of Pathology Firth, Andrew; University of Cambridge, Department of Pathology;
Keywords:	RNA virus, overlapping genes

SCHOLARONE™
Manuscripts

We thank the reviewers and editor for their careful assessment of our manuscript and for their kind and helpful comments and suggestions. Please find below our detailed responses.

Associate Editor: Darren Obbard

I agree with the reviewers, and I enjoyed reading the manuscript very much. However, in addition to those points raised by the reviewer, I would also like to ask for a couple of clarifications or small additions.

(1) Narnaviruses classically lack a virion. However, is there experimental evidence for this in any narna-like virus infecting non-fungi, or is this just a generalization from the *Saccharomyces* viruses?

To the best of our knowledge, this has not been experimentally investigated outside of the two *Saccharomyces cerevisiae* narnaviruses. The lack of an obvious capsid coding sequence (except in ourmia-like viruses, as discussed in the ms and in Dolja & Koonin, 2018, PMID 29103997) suggests that they do lack a virion, but one could envisage other scenarios, such as a helper virus providing capsid proteins, or unidentified capsid-encoding segments. We have added the text "Unlike classical narnaviruses and mitoviruses, ourmiaviruses encode a capsid protein on one of their genome segments, and some of the "ourmia-like" viruses encode a capsid protein in a separate ORF on the RdRp-encoding RNA.", and "In none of the sequences did we observe an additional ORF that might encode a capsid protein. Nonetheless, it is possible that some of these viruses might encode a capsid protein on an undetected genome segment (cf. Lye et al. 2016) or in an extra ORF 5' or 3' of the RdRp ORF that has been missed as a result of an incomplete assembly, or might exploit a helper virus for the provision of a capsid protein."

(2) I was confused by the note that there was 'Scant conservation' in the protein encoded by the rORF, as in the results this was described as "Overall, the amino acid sequences predicted to be encoded by rORFs were rather more divergent than the corresponding set of RdRps (mean pairwise identities 20.1% and 26.6%, respectively)." To me this implies that it is almost as highly conserved (20%) as the RdRP (27%). Based on closer pairs only (say RdRp of 50% protein identity), is there good evidence for some conservation in the rORF protein (i.e. more than expected based on the AA identity seen in the RdRp, the codon usage, and conditioning on the absence of rORF stops?)

Regarding the first point: These levels of global identity (20.1%, 26.6%) are very much in the "twilight zone", i.e. the region where protein homology inference based on sequence alignment identity becomes uncertain. For example Rost (1999, PMID 10195279) analyzed more than a million sequence alignments between protein pairs of known structures and found that, "above a cut-off roughly corresponding to 30% sequence identity, 90% of the pairs were homologous; below 25% less than 10% were". The point we wished to make at the referenced part of the ms was not so much a statement on global conservation (which is minimal) but rather whether or not more local regions of conservation could be detected. From Figure 7 we can see that the RdRp has many more, stronger regions of amino acid conservation than the rORF. Whereas the 20.1% and 26.6% global values don't distinguish whether or not conservation is localized to the same regions in each sequence pair, Figure 7 is based on multiple sequence alignments and thus highlights localized regions that are more conserved than average across many sequences.

To clarify this, we have now reworded "We found scant evidence for amino acid conservation in the rORF." to "In contrast to the RdRp, we found little evidence for conserved amino acid motifs in the rORF protein (**Figure 7**)."

The second point is non-trivial. One way would be to apply synplot2 to the RdRp ORF of more closely related sequences than currently shown in Figure 7. However, this does not directly address the question and does not distinguish between rORF amino acid conservation and, e.g., embedded functional RNA structures. Instead we performed some sequence randomization

1 experiments to attempt to address this point with, as suggested, pairs of more closely related
 2 sequences. "Conditioning on the absence of rORF stops" was achieved by restricting each
 3 analysis to the part of the RdRp ORFs that are overlapped by the rORFs. A simple way to control
 4 for the amino acid identity and codon usage in the RdRp ORF might be to shuffle all the RdRp
 5 codons used for each amino acid X between all sites in the RdRp ORF where amino acid X is
 6 present – i.e. shuffle the RdRp nucleotide sequence while preserving its amino acid sequence
 7 order and codon usage statistics. Then the % amino acid conservation between the two RdRps will
 8 be preserved and one can assess how the shuffled RdRp codons have affected the rORF % amino
 9 acid conservation. This approach is however flawed because the two sequences are not
 10 phylogenetically independent: e.g. if sequence 1 has a CUG codon at position *i*, then sequence 2
 11 is also more likely than random to have a CUG codon at position *i* (e.g. if the ancestral sequence
 12 had a CUG at that site). Thus it is essential to shuffle alignment columns not separate sequences.
 13 Of course, if alignment RdRp codon columns are shuffled, then the % amino acid identities for both
 14 the RdRp ORF and the rORF will be unchanged from the original sequence pair since the rORF is
 15 in phase with the RdRp ORF so has the same codon columns. To move forwards, we decided to
 16 look at conservation of amino acid *k*-mers, figuring that functionally important amino acid motifs
 17 often occur in clumps. We count the % of *k*-mers that are conserved between sequence 1 and 2,
 18 for *k* = 1, 2, 3, ..., for each of the RdRp and rORF, and for the WT pairwise alignment and shuffled
 19 alignments. We shuffle RdRp codon columns between sites that encode the same amino acid pair
 20 (i.e. if sequence 1 has amino acid X and sequence 2 has amino acid Y at positions *i* and *j*, then, in
 21 our shufflings, it is permissible to switch the sequence 1 and 2 codons at position *i* with the
 22 sequence 1 and 2 codons at position *j*). This will have no effect on the RdRp amino acid conserved
 23 *k*-mer counts nor the rORF amino acid conserved 1-mer counts, but it may change the rORF
 24 amino acid conserved *k*-mer counts for *k* > 1.
 25

26 When applied to, for example, the sequence pair KP642119 and KP642120 (61% amino acid
 27 identity in the RdRp), 100 randomizations, we get the following results:
 28

k:	RdRp					rORF				
	1	2	3	4	5	1	2	3	4	5
WT	482	331	233	172	130	261	105	41	21	12
randomized (means)	482	331	233	172	130	261	90.6	32.7	12.1	4.7

36 As noted above, only the last 4 of the 10 columns can have different values between WT and the
 37 randomizations. The result is that, randomizing the sequence decreases the number of conserved
 38 amino acid 2-, 3-, 4- and 5-mers between the two sequences in the rORF polypeptide. This might
 39 be taken as evidence that the rORF polypeptide is subject to some purifying selection at the amino
 40 acid level. Nonetheless, we can think of various caveats with this approach that may render this
 41 observation an artifact.
 42

43 Thus we have not included any change in the ms for this point (see also our response to the
 44 related point 3 below).
 45

46 **(3) related to 2 above, is there any detectable homology at all between different 'clusters' of**
 47 **rORFs (i.e. clades of the the RdRp tree that encode the rORF)? If so, wouldn't this imply**
 48 **multiple losses rather than gains? What about similar properties for the rORF protein**
 49 **(predicted secondary structure? conserved amino-acid properties or composition (after**
 50 **having accounted for any that is enforced by encoding the RdRp on the other strand)?**

51 Regarding the first part:
 52

53 We split the rORF sequences into five main clades (based on Figure 2), aligned each clade with
 54 MUSCLE, and then used HHpred to compare each clade with each other clade. The 10 clade-
 55 versus-clade comparisons had e-values ranging from 0.019 to 8.9e-16. Thus, at first glance, this
 56 analysis indicates that there is homology between different clusters of rORFs. However, to what
 57 extent this reflects constraints imposed by the RdRp and/or common ancestry as opposed to true
 58 purifying selection on the rORF amino acid sequence is difficult to ascertain. As a kind of control,
 59 we did the same analysis for seven non-rORF containing clades (again based on Figure 2), where
 60

we generate “fake” rORF amino acid sequences by translating the RdRp ORF region in the reverse reading direction, and change any stop codons to the ambiguous amino acid code “X”. Many of the non-rORF versus non-rORF interclade HHpred comparisons have comparable e-values to the rORF versus rORF interclade comparisons. Thus the apparent homology between different clusters of rORFs may indeed be due to constraints on the RdRp and/or common ancestry. Actually, in these analyses, it is difficult to control for differences in the number and divergence of sequences in the within-clade alignments, and differences in the divergences between the different clades of the different comparisons. Thus, we do not think this analysis is conclusive one way or the other, and so we have not added these analyses to the manuscript.

Regarding the second part:

We searched for evidence of known motifs (ELM - Eukaryotic Linear Motif server) and transmembrane regions (TMHMM, TMpred) in individual sequences, and between related sequences at various levels of identity, but did not find anything of particular note.

We have now also looked at various sequence properties: pI, predicted helix/sheet/turn/coil composition, and hydrophobicity plots, using the 26 representative RdRp and rORF sequences.

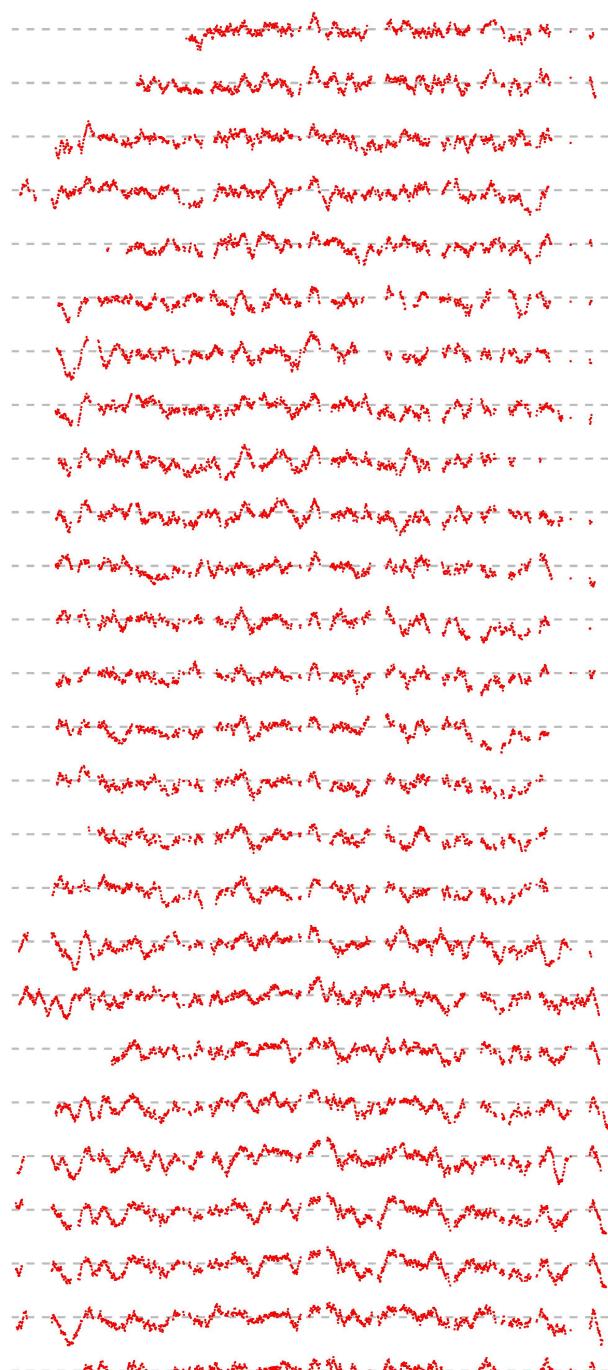
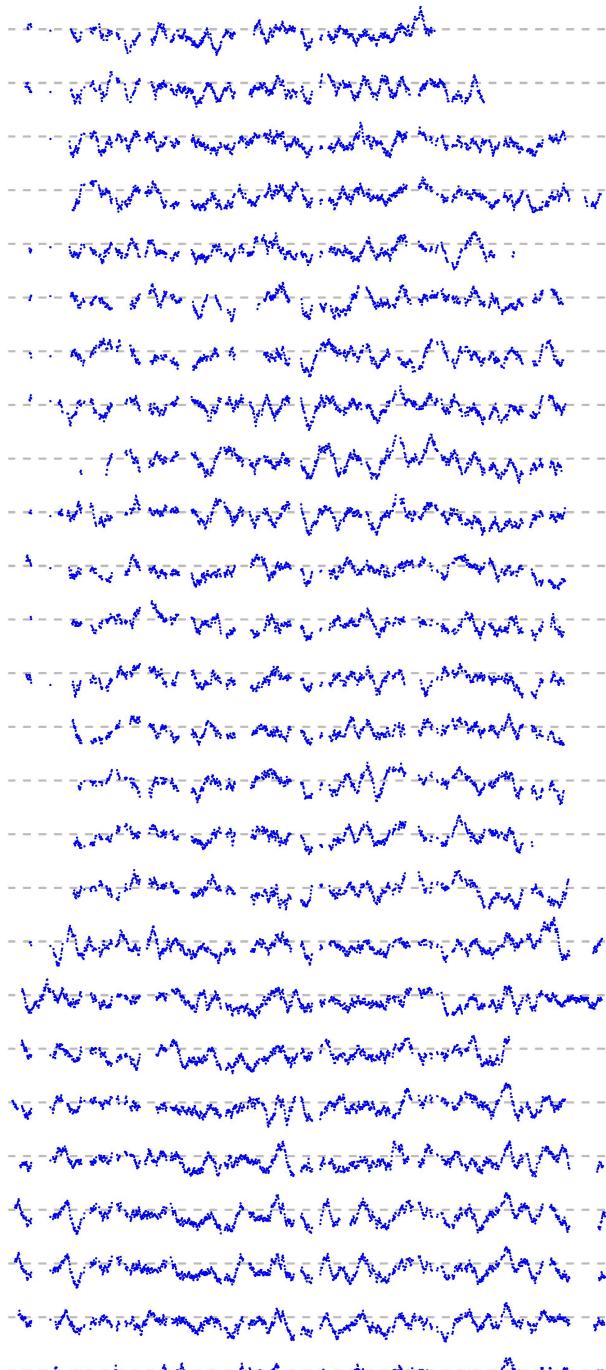
Isoelectric point (pI) is potentially interesting. Both the rORF polypeptide and the RdRp are quite basic (rORFs 9.82 ± 0.98 , RdRps 10.06 ± 0.59 ; cytoplasmic proteins have an average pI value of ~5–6). This is consistent with a hypothesis that the rORF protein might, like the RdRp, bind the viral RNA to protect it and maintain it in a single-stranded form in the cytoplasm. We have now added to the ms the text "Both the RdRp and rORF amino acid sequences are highly basic (isoelectric points: rORFs – 9.82 ± 0.98 , RdRps – 10.06 ± 0.59 ; means \pm standard deviations). The high isoelectric point of the RdRp is consistent with it binding viral nucleic acid to form ribonucleoprotein complexes. It is possible that the rORF protein might also bind viral nucleic acid, perhaps (given its expected low expression level) with a specificity for the negative-strand."

We predicted helix/sheet/turn/coil composition with the garnier program in the EMBOSS suite. This has a fairly low prediction accuracy (~65%), but we prefered to use a single-sequence method, and then simply look at overall proportions of positions called as helix, sheet, turn or coil. The results are (means \pm standard deviations over the 26 representative rORFs and RdRPs):

	rORFs	RdRPs
helix	0.191 ± 0.048	0.314 ± 0.051
sheet	0.259 ± 0.038	0.242 ± 0.034
turns	0.310 ± 0.033	0.229 ± 0.033
coil	0.240 ± 0.037	0.214 ± 0.027

Thus the RdRPs have more predicted helix than the rORFs and the rORFs have more predicted turns. Even so, we decided not to add this to the ms as it does not obviously lead anywhere.

We calculated hydrophobicity with the pepwindow program (19-aa window) in the EMBOSS suite, which produces Kyte-Doolittle plots. We then transformed the plot for each sequence to alignment coordinates, so that profiles could be compared across the 26 different rORF and RdRp sequences. We also plotted the mean (for positions in the alignment with <20% gaps), with the idea being, perhaps, that if the rORF has a less-conserved hydrophobicity profile, then the mean should be flatter. These data are shown in the plot on the following page (red = RdRp, blue = rORF, pink = RdRp mean, purple = rORF mean). However, once again, there does not seem to be anything to be gained by adding this analysis to the ms.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60**RdRp – individual sequences****RdRp – average****rORF – individual sequences****rORF – average**

1
2 Finally, I note that a substantial amount of the material in lines 340-370 describe results,
3 and I think it would be better if all results appeared in the results section (with
4 accompanying methods in the methods section).
5

6 The parts describing the SRA taxonomy analyses are simply reporting data shown in the NCBI
7 databases, so these are not really something that can be described in our Results/Methods
8 sections. Similarly much of the other material here is a discussion of previously published results.
9 We have however rearranged the EVE/WGS material as suggested, adding a new (small) Results
10 section ("EVEs related to rORF-containing narnaviruses are found in insect WGS datasets"),
11 besides corresponding details in Methods.
12

13 We also added another sentence to the Discussion in light of a very recent paper: "On the other
14 hand, it has also been suggested that arthropod-associated narnaviruses may instead be infecting
15 trypanosomatids – common intracellular parasites of arthropods (Harvey et al., 2019)
16 (trypanosomatid DNA – including EVEs – could potentially also contaminate WGS datasets of their
17 hosts)."
18

19 -----
20 **Reviewer: 1**
21

22 It's a pleasure to read this paper. The authors presented strong bioinformatics evidence
23 that in some alphananavirus, the negative strand encode functional ORFs. The performed
24 phylogenetic analysis which separate nanaviruses into two proposed genera. The showed
25 that in about half of the alphananaviruses, the negative strand contained an ORF that
26 spanned most of the sequences. The length of the ORFs, their R0 reading frame in relation
27 to the reading frame of RdRp, and the codon avoidance of UUA, CUA, and UCA, provide
28 strong evidence that these ORFs are functional.
29

30
31 It will be exciting if the authors or someone else can provide experimental evidence that the
32 rORFs are indeed expressed at the protein level. Since their sequences are divergent, and
33 they only exist in some of the alphaviruses, it's difficult to imagine that they have essential
34 function. However, if they lack of essential function, how do the viruses maintain such
35 rORFs? I look forward to follow-up studies.
36

37 A few minor comments:
38

39 **1. Do any of the rORFs-encoded proteins have any similarity to known proteins?**
40

41 To check for significant homology to known proteins, we performed blastp searches against the
42 NCBI non-redundant (nr) protein database, using each of the protein sequences predicted to be
43 encoded by an rORF separately as a query. Apart from other rORF-containing narnaviruses, no
44 significant homology was detected. We also applied HHpred, using rORF amino acid alignments
45 as queries (five alignments, one for each of the main rORF clades in Figure 2, with – for
46 definiteness – IABX01132835, KX883605, KX883500, excluding KX883602 being one of these
47 clades), searching against PDB_mmcIF70_11_Oct and Pfam-A_v32.0, but again found no
48 potential homologues (best e-value = 56). This is not surprising, since overlapping genes normally
49 have to evolve via overprinting of an ancestral gene (in this case, clearly the RdRp), and not via
50 gene duplication and divergence, and thus they nearly always exhibit novel protein folds (Keese &
51 Gibbs, 1992, PMID 1329098; Rancurel et al, 2009, PMID 19640978). We have added details of
52 these analyses to the ms Results (top of "Analysis of amino acid and nucleotide conservation in
53 the rORF" section) and Methods.
54

55 **2. Page 9 lines 194-196: The authors restricted the rORFs finding to within 200 nt of 5' end.
56 Though the provided justification, it would be beneficial if they repeat this part of analysis
57 without this restriction and present it as a supplemental figure.**
58

We have included the additional figure in the Supplementary Material (new Supplementary Figure 5), showing the longest stop codon-free regions (i.e. ORFs), as a percentage of sequence length, regardless of the 5' start position of the ORF. Similar to the original version (Figure 4), there is a clear segregation of rORF-containing and rORF-lacking taxa, based on R0-frame ORF lengths. As would be expected, when the 5' start position of the ORF is not restricted, the maximal relative ORF lengths increase for many sequences. In particular, two additional sequences have an R0-frame ORF occupying ~36% of the respective sequences. However, this value (36%) is still considerably closer to the range for non-rORF-containing sequences (previously 5–25%) than it is to the range for rORF-containing sequences (94–100%). Hence, our argument – that a clear distinction can be drawn between rORF-containing and rORF-lacking taxa, based on R0-frame ORF length – is unchanged.

3. Page 10, lines 221-222, double check the codon locations. Since they are in the same reading frame, the differences should be multiples of 3. The numbers presented are not.

We apologise for the confusion here. Actually these numbers are in units of codons rather than nucleotides (i.e. a difference of 1 = 1 codon). Therefore, the differences in positions need not be a multiple of three. To help clarify this, we have amended the text as follows:

- "at positions 105 and 373" → "at **codons** 105 and 373"
- "at positions 81, 134 and 157" → "at **codons** 81, 134 and 157"
- "at position 72" → "at **codon** 72"

4. Page 12, lines 262-265, the authors used the virus GC contents to justify their conclusion. The GC contents of their hosts (if known) should be used instead since the viruses use their host machinery for translation.

This is not possible, since the hosts are only really known for a very small subset of narnaviruses (specifically, those for which direct infection has been shown in a laboratory setting). Others might come from fungi or protozoans that infect the main transcriptomic target such as insects or plants. However, the point of this section in the ms is that the presence or absence of the rORF cannot be explained by variation in overall GC3 content of the virus genome. This conclusion is based on a direct comparison of rORF-containing and non-rORF-containing alphanarnaviruses, and it stands regardless of the host taxa.

Reviewer: 2

This paper is close to perfection, it is brilliantly written and its technical level is without question. The authors cannot do miracles, though: the message of the paper boils down to finding long rORF, which arise due to the avoidance of stop codons. There is no information whatsoever as to what the function of the encoded proteins could be. However, I believe that these thoroughly documented findings as such are of great interest and "Virus Evolution" is a perfect journal for presenting them.

We greatly appreciate the reviewer's kind words.

ARTICLE**A case for a negative-strand coding sequence in a group of positive-sense RNA viruses**

Adam M. Dinan¹, Nina I. Lukhovitskaya¹, Ingrida Olendraite¹, and Andrew E. Firth¹

1. Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road,
Cambridge, CB2 1QP, UK

Correspondence to: Andrew E. Firth, aef24@cam.ac.uk

ABSTRACT

Positive-sense single-stranded RNA viruses form the largest and most diverse group of eukaryote-infecting viruses. Their genomes comprise one or more segments of coding-sense RNA that function directly as messenger RNAs upon release into the cytoplasm of infected cells. Positive-sense RNA viruses are generally accepted to encode proteins solely on the positive strand. However, we previously identified a surprisingly long (~1000-codon) open reading frame (ORF) on the negative strand of some members of the family *Narnaviridae* which, together with RNA bacteriophages of the family *Leviviridae*, form a sister group to all other positive-sense RNA viruses. Here, we completed the genomes of three mosquito-associated narnaviruses, all of which have the long reverse-frame ORF. We systematically identified narnaviral sequences in public data sets from a wide range of sources, including arthropod, fungal and plant transcriptomic datasets. Long reverse-frame ORFs are widespread in one clade of narnaviruses, where they frequently occupy >95% of the genome. The reverse-frame ORFs correspond to a specific avoidance of CUA, UUA and UCA codons (i.e. stop codon reverse complements) in the forward-frame RNA-dependent RNA polymerase ORF. However, absence of these codons cannot be explained by other factors such as inability to decode these codons or GC3 bias. Together with other analyses, we provide the strongest evidence yet of coding capacity on the negative strand of a positive-sense RNA virus. As these ORFs comprise some of the longest known overlapping genes, their study may be of broad relevance to understanding overlapping gene evolution and *de novo* origin of genes.

29 INTRODUCTION

Traditionally, viruses have been divided between seven Baltimore classes based on the nature of the nucleic acid of their genomes and their replicative intermediates. The seven classes are positive-sense, negative-sense and double-stranded RNA viruses, single-stranded and double-stranded DNA viruses, retroviruses and pararetroviruses (Baltimore 1971). Of these, the single-stranded positive-sense or (+)ssRNA viruses comprise the largest and most diverse group of eukaryote-infecting viruses (Dolja and Koonin 2011). The group includes many important human and animal pathogens (such as dengue, Zika, yellow fever, hepatitis C, foot-and-mouth disease, polio, chikungunya, SARS and MERS viruses) besides the majority of plant viruses.

To fully understand the molecular biology of viruses, it is crucially important to know their coding capacity. In recent years, a number of “hidden” protein-coding open-reading frames (ORFs) have been discovered in the genomes of various (+)ssRNA viruses (Chung et al. 2008; Loughran et al. 2011; Fang et al. 2012; Firth 2014; Smirnova et al. 2015; Napthine et al. 2017; Lulla et al. 2019). Such genes tend to be very short and/or to overlap previously known coding ORFs, explaining why they have escaped detection prior to the application of sensitive comparative genomic methods. However, all of these experimentally verified novel coding ORFs are in the positive sense. To our knowledge, no negative-sense coding ORF has *ever* been demonstrated in *any* (+)ssRNA virus. Thus, an unanswered question in virology is whether the negative strand of (+)ssRNA viruses can encode proteins.

In 2013, by acquiring and analyzing RNA transcriptomic datasets for several mosquito and other dipteran species, we identified two “mosquito-associated” narnavirus-like viruses (Cook et al. 2013). Similar to other narnaviruses, one strand contains a single long ORF that encodes a protein inferred by homology to be the RNA-dependent RNA polymerase (RdRp). However, we were surprised to

1
2
3
4 52 notice that both sequences also contain a reverse-frame ORF (rORF) covering nearly the entire
5 sequence of ~3000 nt. We also identified related rORF-containing sequences in public *Puccinia*
6
7 54 *striiformis* and *Uromyces appendiculatus* transcriptome shotgun assembly (TSA) data sets. The
8 extreme divergences between these sequences and the mosquito-associated sequences (~22%
9 amino acid identity in the RdRp sequence) effectively rule out a region of such length being
10 preserved free of stop codons by chance, and thus we hypothesized that the rORF represented a
11
12 56 *bona fide* protein-coding sequence.
13
14
15
16
17
18
19

20
21 59 As currently defined by the International Committee on Taxonomy of Viruses (ICTV), the family
22
23 60 *Narnaviridae* contains two genera: *Mitovirus* and *Narnavirus*. Both genera contain single-stranded
24 positive-sense RNA viruses which are non-encapsidated and hence are expected to be transmitted
25 either vertically through cell division or horizontally during host mating. Mitoviruses replicate in
26 the mitochondria of host cells whereas narnaviruses replicate in the cytoplasm. These viruses were
27 originally described as infecting fungi, but related viruses have since been observed in
28 transcriptomic data sets derived from diverse organisms. The narnaviral positive strand normally
29 contains a single long ORF that covers most of the genome and encodes an RdRp which catalyses
30 viral replication. The narnaviral RdRp is highly divergent from those of other eukaryotic RNA
31 viruses, and shows closer homology to the RdRps of RNA bacteriophages in the family *Leviviridae*
32 (Rodriguez-Cousiño et al. 1991; Esteban et al. 1992; Wolf et al. 2018). Detailed comparative
33 genomic and phylogenetic analyses suggest that the *Narnaviridae* are descended from a levivirus-
34 like bacteriophage which may have been carried within the bacterial progenitor of mitochondria at
35 the point of eukaryogenesis, followed by loss of the capsid protein giving rise to capsidless
36 (“naked”) RNA elements (Koonin et al. 2015; Wolf et al. 2018). The escape of a group of these
37 viruses into the cytosol may have given rise to the *Narnavirus* genus (Koonin and Dolja 2014;
38 Wolf et al. 2018).
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 76 The prototypical and best studied narnaviruses are the *Saccharomyces cerevisiae* 20S and 23S
5
6 77 RNA viruses (ScNV-20S and ScNV-23S, respectively) (reviewed in Wickner et al. 2013). ScNV-
7
8 78 20S persistently infects most laboratory strains of the yeast *Saccharomyces cerevisiae*, whereas
9
10 79 fewer strains carry ScNV-23S. Their genomes have no 3' poly(A) tail and it is not known if they
11
12 80 have a 5' cap structure (Rodríguez-Cousiño et al. 1998). The genomes do not encode capsids, but
13
14 81 form ribonucleoprotein complexes with the RdRp in a 1:1 stoichiometry in the host cell cytoplasm;
15
16 82 the RdRp interacts with both the 5' and 3' ends which may help protect the viral RNA from
17
18 83 degradation by host exonucleases (Solórzano et al. 2000; Fujimura and Esteban 2004; Fujimura
19
20 84 and Esteban 2007). Under suitably inducing conditions such as heat shock and nitrogen starvation,
21
22 85 copy numbers of either virus can reach 100,000 copies per cell (Kadowaki and Halvorson 1971;
23
24 86 Wejksnora and Haber 1978; Esteban et al. 1992). Approximately 98–99% of viral RNA in the cell
25
26 87 is in a single-stranded positive-sense form, whereas the remainder exists as single-stranded
27
28 88 negative-sense replication intermediates (Rodríguez-Cousiño et al. 1991; Fujimura et al. 2005).
29
30 89 When cells are grown at high temperature, double-stranded forms (known as W for ScNV-20S and
31
32 90 T for ScNV-23S) accumulate, but these appear to represent by-products and not replication
33
34 91 intermediates (Wesolowski and Wickner 1984; Rodríguez-Cousiño et al. 1998; Fujimura et al.
35
36 92 2005).

37
38
39
40
41
42
43
44 93 To further investigate the presence of the rORF, and now that many more sequences are available
45
46 94 in public sequence databases, in this work we present a comprehensive comparative analysis of
47
48 95 narnaviral genomes, in which we assess the prevalence, distribution and sequence features of
49
50 96 rORFs. Two major clades of narnavirus are identified, for which we propose the establishment of
51
52 97 the genera *Alphanarnavirus* and *Betanarnavirus*, with the former clade containing all sequences
53
54 98 with long rORFs. Overall codon usage is similar in alphanarnaviruses with and without long
55
56 99 rORFs, but the former display a highly specific avoidance of CUA, UUA and UCA codons over

1
2
3
4 100 large regions of the RdRp ORF, corresponding to an absence of stop codons in the rORF. We
5
6 101 explore possible reasons for the avoidance of CUA, UUA and UCA and conclude that the most
7
8 102 plausible explanation is selection to maintain an rORF, indicating that the rORF is functional.
9
10
11
12 103
13
14
15 104 **RESULTS**
16
17
18 105
19
20
21
22 106 **Completion of narnaviral genomic sequences**
23
24
25 107 We completed three genomic sequences for previously described rORF-containing narnaviruses
26
27 108 (Cook et al. 2013), and used these genomes as a reference set for comparative analyses. The
28
29 109 sequences with GenBank accession numbers KF298275.1, KF298276.1 and KF298284.1 were
30
31 110 extended at their 5' termini by 479 nt, 10 nt and 364 nt, respectively, and at their 3' termini by 43
32
33 111 nt, 9 nt and 9 nt, respectively (**Figure 1A**). The latter two sequences are the same length (3093 nt)
34
35 112 and encode RdRps which are 97.95% identical at the amino acid level; hence, they were considered
36
37 113 to represent a single viral species. Given that the samples from which these sequences were
38
39 114 originally obtained were isolated from mosquitoes in the *Ochlerotatus* genus, we name the viruses
40
41 115 Ochlerotatus-associated narna-like virus 1 (ONLV1; sequence KF298275.1) and ONLV2
42
43 116 (sequences KF298276.1 and KF298284.1) (**Figure 1A**). The AUG codons of the three RdRp ORFs
44
45 117 each start at the 7th nucleotide of the respective sequences (contexts: GUCAUGA, GUUAUGG
46
47 118 and GUUAUGG), the same position as that of the RdRp AUG in ScNV-23S (GenBank:
48
49 119 NC_004050.1). The RdRp-encoding ORFs are 3045, 3075 and 3075 nucleotides in length,
50
51 120 respectively. The RdRp stop codons (UGA, UAA and UAA) end at positions 14 nt, 13 nt and 13 nt,
52
53
54
55
56
57
58
59
60

1
2
3
4 121 respectively, from the 3' termini of the genomes. Hence, the length of the 3' UTR mirrors closely
5
6 122 that of ScNV-20S (12 nt; GenBank: NC_004051.1) but is shorter than that of ScNV-23S (59 nt).
7
8
9
10 123
11
12
13 124 **An expanded narnaviral phylogeny**
14
15
16 125 We employed tblastn (Camacho et al. 2009) and hmmsearch (Finn et al. 2011), using hidden
17
18 126 Markov model based profiles (pHMM) to identify sequences encoding proteins closely related to
19
20 127 narnaviral RdRps in the NCBI non-redundant nucleotide (nr/nt) and TSA databases. In total, 124
21
22 128 unique such sequences were identified – 46 in the nr/nt database and 78 in the TSA database. The
23
24 129 holobiont sources of these TSA sequences were phylogenetically diverse and included 21
25
26 130 arthropod, 8 fungal and 10 plant species. To place these sequences in context, we also selected a
27
28 131 number of mitovirus and ourmiavirus sequences as outgroups, besides additional narna-like virus
29
30 132 sequences from Shi et al. (2016), giving 141 sequences in total (**Supplementary Table S1**). We
31
32 133 oriented all sequences (by reverse complementing if necessary) so that the RdRp ORF was on the
34
35 134 forward strand.
36
37
38
39
40 135 To assess the relationships among sequences, an alignment of the predicted RdRp proteins was
41
42 136 generated using MUSCLE (Edgar 2004) (**Supplementary File S1**), and a Bayesian Markov chain
43
44 137 Monte Carlo (MCMC)-based tree was constructed from this alignment using MrBayes
45
46 138 (Huelsenbeck and Ronquist 2001; **Figure 1B**). As expected, the free-floating genus *Ourmiavirus*,
47
48 139 which includes plant-associated viruses with tripartite genomes, grouped unambiguously with the
50
51 140 narnaviruses (Turina et al. 2017; Rastgou et al. 2009; Wolf et al. 2018), whereas mitoviruses
53
54 141 formed a sister clade (**Figure 1B**). A relatively small, but highly diverse set of fungal- and
55
56 142 arthropod-derived sequences clustered with the ourmiaviruses (**Figure 1B**). Recent evidence shows
58
59 143 that these “ourmia-like” viruses display a range of genomic architectures, which can be segmented
60

1
2
3
4 144 (as in the case of the *bona fide* ourmiaviruses) or non-segmented, occurring in mono- and di-
5 cistronic forms (reviewed in Dolja and Koonin 2018). For example, botrytis ourmia-like virus,
6 which appears to have a non-segmented genome, clusters unambiguously with ourmiaviruses rather
7 than with the narnaviruses (Donaire et al. 2016; **Figure 1B**). Unlike classical narnaviruses and
8 mitoviruses, ourmiaviruses encode a capsid protein on one of their genome segments, and some of
9 the “ourmia-like” viruses encode a capsid protein in a separate ORF on the RdRp-encoding RNA
10 (Dolja and Koonin 2018).

11 151
12
13
14
15
16
17
18 150 (Dolja and Koonin 2018).
19
20
21
22
23
24
25 152 **Two major clades of narnaviruses**
26
27
28 153 The narnaviral sequences predominantly fell into one of two major clades, both of which had
29
30 154 Bayesian posterior probabilities of 0.95 (**Figure 1B**). Based on these data, we propose that the
31 genus *Narnavirus* be subdivided to form two new genera, which we name *Alphanarnavirus* and
32 155
33
34 156 *Betanarnavirus*, besides additional unclassified sequences.

35
36
37
38 157 The alphanarnaviral clade contains the prototypical narnaviruses, ScNV-20S and ScNV-23S, as
39
40 158 well as sequences associated with a range of other fungal taxa, including members of the divisions
41
42 159 Ascomycota, Basidiomycota and Entomophthoromycota, and the divergent subphylum
43
44 159
45 160 Mucoromycotina (**Figure 2**). The alphanarnaviral sequences ranged in length from 1805 nt to 3874
46
47 161 nt (**Supplementary Figure S1**), and the pairwise amino acid identities of the corresponding set of
48
49 162 RdRps ranged from 16.8% to 99.8% (**Supplementary File S2**). Putative viral sequences containing
50
51 163 an rORF all clustered within the alphanarnaviral clade, and none in the betanarnaviral clade.
52
53
54 164 However, the rORF-containing sequences appear not to form a monophyletic clade, but instead
55
56 165 cluster in several regions of the phylogeny, and are found in sequences derived from fungi,
57
58 166 arthropods and plants (**Figure 2: red bars**). The core RdRp catalytic regions – motifs A to E in the
59
60

1
2
3
4 167 palm domain and motifs F and G in the fingers (Wu et al. 2015; te Velthuis 2014) – are well-
5
6 168 conserved despite the overall high degree of sequence divergence (**Supplementary Figure S2**).
7
8
9

10 169 The betanarnaviral clade includes several viruses of unicellular eukaryotes, including the
11
12 170 oomycete-infecting Phytophthora infestans RNA virus 4 (PiRV4; Cai et al. 2012) and the
13
14 171 protozoan-associated Leptomonas seymouri narna-like virus 1 (Lye et al. 2016), as well as
15
16 172 sequences from red algae, brown algae, myxozoa and arthropods (**Figure 3**). The betanarnaviral
17
18 173 sequences ranged in length from 2215 nt to 3610 nt (**Supplementary Figure S1**), and the pairwise
19
20 174 RdRp amino acid sequence identities ranged from 15.2% to 99.9% (**Supplementary File S3**). In
21
22 175 the case of Leptomonas seymouri narna-like virus 1, the genome may be bipartite, with the RdRp
23
24 176 being encoded on the longer (L) segment, although the functional association of the putative
25
26 177 segments has not been shown experimentally (Lye et al. 2016).
27
28
29
30

31 178 Notably, in both clades, a number of TSA sequences are highly divergent from other sequences in
32
33 179 the phylogenies (**Figure 2** and **Figure 3**), suggesting that further sampling will continue to reveal
34
35 180 new clades. Conversely, other groups were disproportionately well sampled (e.g. nine Zhejiang
36
37 181 mosquito virus 3 sequences with >95% pairwise amino acid identity were found; Shi et al. 2017).
38
39 182 In none of the sequences did we observe an additional ORF that might encode a capsid protein.
40
41 183 Nonetheless, it is possible that some of these viruses might encode a capsid protein on an
42
43 184 undetected genome segment (cf. Lye et al. 2016) or in an extra ORF 5' or 3' of the RdRp ORF that
45
46 185 has been missed as a result of an incomplete assembly, or might exploit a helper virus for the
47
48 186 provision of a capsid protein.
49
50
51
52
53 187
54
55
56 188 **Genomic architecture and terminal regions**
57
58
59
60

1
2
3
4 189 The terminal sequences of alphanarnaviral and betanarnaviral genomes were found to be dissimilar
5 – the former group having short runs of G and C residues at the 5' and 3' termini, respectively, and
6
7 190 the latter having A/U-rich termini with considerably longer 3' untranslated regions (UTRs)
8
9 191 (Supplementary Figure S3). Local RNA secondary structures are predicted to occur at the 5' ends
10
11 192 of the ONLV1 and ONLV2 genomes, coincident with a reduction in synonymous site variation in
12
13 193 the RdRp ORF (see below). The putative structures are large (92 nt and 103 nt), and the G-rich 5'
14
15 194 terminus forms an integral component of the stem duplex (Supplementary Figure S4A), as it does
16
17 195 in the *Saccharomyces* narnaviruses (Rodríguez-Cousiño et al. 1998; Fujimura and Esteban 2007).
18
19 196 Meanwhile, the stop codons of the RdRp ORFs are predicted to be situated within shorter RNA
20
21 197 stem-loop structures at the genomic 3' termini (Supplementary Figure S4, B and C).
22
23
24
25 198
26
27
28 199
29
30
31 200 **Identification of ORFs in narnaviral genomes**

32
33
34
35 201 We next determined for each sequence the longest stop codon-free region (ORF) in each of the
36
37 202 three possible reading frames on each of the positive and negative strands, restricting to ORFs that
38
39 203 begin within 200 nt of the 5' end of the positive or negative strand, as appropriate (Figure 4; see
40
41 204 Supplementary Figure S5 for a similar analysis without the 5'-proximity restriction). We used this
42
43 approach because, under normal circumstances, non-5'-proximal ORFs are not expected to be
44
45 205 translated, and the inclusion of spurious long non-5'-proximal ORFs might dilute the signal from
46
47 206 translatable 5'-proximal ORFs. We designated the reading frame of the RdRp ORF as frame “F0”
48
49 207 (for forward orientation, 0 frame), followed by frames “F+1” and “F+2” in the same orientation;
50
51 208 whereas the reverse complement of the set of codons in frames F0, F+1 and F+2 were designated
52
53 209 “R0” (reverse orientation, 0 frame), followed by “R+1” and “R+2” respectively.
54
55
56 210
57
58
59
60

1
2
3
4 211 In line with the relatively short lengths of narnaviral untranslated regions (UTRs), the RdRp-
5 encoding (i.e. frame F0) stop codon-free regions occupied 93.8–100.0% (median 99.5%) of
6 alphanarnaviral sequences and 70.5–100.0% (median 97.0%) of betanarnaviral sequences (**Figure**
7
8 213
9 214 **4**). Several sequences in the data set are likely to be incomplete and, as a result, to have
10 misannotated start codons; for example, the genome of Hubei narna-like virus 18 (KX883517.1;
11
12 215
13 Shi et al. 2016) is annotated as containing a 155-nt 5' UTR, but this UTR contains no stop codons
14
15 216 in-frame with the RdRp, and a blastx search of the sequence shows that it encodes amino acid
16
17 217 sequence that shows close homology to the RdRp of the related Hubei narna-like virus 19 (e-value
18
19 218 = 4×10^{-12}).
20
21
22
23
24
25

26 220 Strikingly, other than the RdRp ORF, 5'-proximal stop-codon free regions occupying >25% of the
27 genome were present only in the R0 frames and only in a subset of alphanarnaviral sequences,
28 whereas stop (UAG, UAA, UGA) codons in all other frames were relatively common (**Figure 4**).
29
30 222
31 223 The absence of stop codons in frame R0 in this subset directly mirrors an absence of the reverse-
32 complementary codons (CUA, UUA, UCA) in the RdRp-encoding (F0) frame.
33
34
35 224
36
37

38 225 In a single case – Wenling narna-like virus 7 (GenBank: KX883602.1; Shi et al. 2016) – an rORF
39 of intermediate length (531 codons; 58.5% of the sequence) was found (**Figure 4**), compared with
40 an 890-codon stop-free RdRp-encoding region on the forward strand. The 5' half of this RdRp ORF
41 includes several codons corresponding to R0-frame stops: including two CUA (reverse-strand =
42 UAG) codons, at **codons** 105 and 373; three UCA (reverse-strand = UGA) codons at **codons** 81,
43 134 and 157; and one UUA (reverse-strand = UAA), at **codon** 72. However, the closely related
44 sequences of Wenling narna-like virus 8 (KX883605.1), Beihai narna-like virus 24 (KX883500.1),
45 and the TSA sequence from *Caridina multidentata* (IABX01132835.1) all contain rORFs
46 occupying >90% of the respective sequences (**Figure 2**).
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

234

235 **Avoidance of CUA, UUA and UCA codons in RdRp-encoding ORFs**

236 To enable a representative assessment of codon usage in narnaviral ORFs, sequences predicted to
237 encode RdRp proteins with >90% amino acid identity were clustered using CD-HIT (Fu et al.
238 2012) and a single sequence from each cluster was retained for further analysis. After this step, 53
239 alphanarnavirus and 29 betanarnavirus representative sequences were left. An rORF occupying at
240 least 90% of the sequence was present in 26 of the 53 alphanarnavirus sequences.

241 Among alphanarnaviruses, codon usage (as a proportion of total codons) in the RdRp ORF was
242 broadly similar in sequences with and without long rORFs, with the exception of the three forward-
243 orientation codons that introduce reverse-orientation stops (i.e. CUA, UUA and UCA) (**Figure 5**).
244 Effectively by definition, each of these three codons is excluded from large portions of sequences
245 with long rORFs (mean usage per associated amino acid in RdRp ORF = 0.0037, 0.0054 and
246 0.0056, respectively), but are 17- to 26-fold more common in those alphanarnaviruses without long
247 rORFs (mean usage per associated amino acid = 0.0973, 0.0896, and 0.1277, respectively) (**Figure**
248 **6A**). These three codons were also relatively common in the RdRp ORFs of the betanarnaviruses
249 (mean usage per associated amino acid = 0.1323, 0.1452 and 0.1784, respectively). CUA and UUA
250 both encode leucine (Leu) whereas UCA encodes serine (Ser). Examination of codon usage bias
251 for these two amino acids shows that sequences with long rORFs specifically avoid these codons
252 and use proportionately more of each alternative codon to encode these two amino acids (**Figure**
253 **6A**).

254

255 **Comparison of codon usage across species**

1
2
3
4 256 Like all (+)ssRNA viruses, narnaviruses are dependent upon host tRNA pools and translation
5 machinery for their gene expression. This raises the possibility that the paucity of CUA, UUA and
6
7 257 UCA codons in the rORF-containing sequences could at least partially reflect an adaptation to
8 particular host codon usage patterns.
9
10
11
12

13
14 260 Using the RefSeq database from NCBI, as tabulated in the latest release of the codon usage table
15 database (CUTD) (Athey et al. 2017), we assessed the relative proportions of the different leucine
16 and serine codons across species. We found that global usage of CUA, UUA and UCA (per
17 associated amino acid) scales directly and inversely with GC content at the third position of codons
18 (GC3) and only species with extremely high GC3 had bias against CUA, UUA and UCA as
19 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
20
21 263 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
22 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
23
24 264 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
25 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
26
27 265 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
28 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
29
30 266 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
31 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
32
33 267 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
34 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
35
36 268 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
37 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
38
39 269 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
40 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Most taxa with mean overall
41
42 270 GC3 and only species with extremely high GC3 had bias against CUA, UUA and UCA as
43 extreme as that in the rORF-containing narnaviruses (**Figure 6B**). Thus, selection
44
45 271 against CUA, UUA and UCA codons in rORF-containing alphanarnaviruses cannot readily be
46 explained by GC bias or host codon usage bias.
47
48

49 274 To investigate whether codon usage could be used as an indicator of the likely host taxonomic
50 group or groups for narnaviruses, and motivated by previous work (Kapoor et al. 2010), we
51 performed a principal component analysis (PCA) of codon usage statistics for host groups that
52 frequently co-occur with narnaviruses in transcriptomic datasets. The PCA allowed clear
53 segregation of arthropods from ascomycetes and basidiomycetes (two major fungal phyla) and
54 streptophytes (land plants and some green algae), but the latter three groups did not clearly
55
56
57
58
59 279
60

1
2
3
4 280 segregate from each other (**Supplementary Figure S6**). Narnaviral codon usage statistics were
5 projected onto the resulting principal component space, and did not clearly segregate with a single
6 host group, regardless of whether or not leucine and serine codons were included in the analysis
7
8 282 (**Supplementary Figures S6 and S7**). Chordates were excluded from this analysis, as their codon
9 bias is heavily influenced by the avoidance of CpG- and UpA-ending codons, and they group
10 distinctly from other phyla and from the narnaviruses.
11
12
13
14
15
16
17
18
19 286
20
21
22 287 **Analysis of amino acid and nucleotide conservation in the rORF**
23
24
25
26 288 To check for homology to known proteins, we queried rORF amino acid sequences and alignments
27 with blastp and HHpred respectively but, apart from other rORF-containing narnaviruses, no
28 significant homologies were detected. This is not surprising, since overlapping genes normally
29 have to evolve via overprinting of an ancestral gene (in this case, the RdRp), and not via gene
30 duplication and divergence, and thus they nearly always exhibit novel protein folds (Keese and
31 Gibbs 1992; Rancurel et al. 2009).
32
33
34
35
36
37
38
39
40 294 Both the RdRp and rORF amino acid sequences are highly basic (isoelectric points: rORFs – 9.82
41 ± 0.98, RdRps – 10.06 ± 0.59; means ± standard deviations). The high isoelectric point of the
42 RdRp is consistent with it binding viral nucleic acid to form ribonucleoprotein complexes. It is
43 possible that the rORF protein might also bind viral nucleic acid, perhaps (given its expected low
44 expression level) with a specificity for the negative-strand.
45
46
47
48
49
50
51
52
53 299 To assess conservation within the rORF of the representative narnavirus sequences, the predicted
54 RdRp amino acid sequences were aligned and then back-translated to RNA. Overall, the amino
55 acid sequences predicted to be encoded by rORFs were rather more divergent than the
56
57
58
59
60

1
2
3
4 302 corresponding set of RdRps (mean pairwise identities 20.1% and 26.6%, respectively). Codon-
5 based alignments of RdRp ORFs and rORFs were generated, based in both cases on the RdRp
6
7 303 amino acid alignments, and amino acid conservation and synonymous site variation of codons
8
9 304 (Firth 2014) were assessed (**Figure 7**).
10
11
12

13
14 306 The core RdRp functional motifs (A to G; see above) localise to a region of fewer than 400 amino
15 acids; with additional short conserved motifs (R&UP and Pxx[L/V]GGx[G/N]xP; U = I, L, V or
16
17 307 M, & = I, L, V, M, A, P, G, F, W or Y) being found further upstream and downstream, respectively
18
19 308 (**Figure 7A**). Increases in synonymous site conservation were observed in the rORF directly
20 opposite RdRp motifs A and E, and – to a lesser extent – motifs B, C, D and F, and at the
21
22 309 Pxx[L/V]GGx[G/N]xP motif (**Figure 7B**) indicating that these conserved motifs in the RdRp
23
24 310 constrain synonymous site variation in the rORF.
25
26
27 311
28
29 312

30
31 313 The largest increase in synonymous site conservation in the RdRp ORF lies between codons 23 and
32
33 314 37 (**Figure 7A**), coinciding with the location of a known RdRp interaction site and a *cis*-acting
34
35 315 replication signal in the ScNV-20S genome (Fujimura and Esteban 2007). This region also contains
36
37 316 the predicted 5' structures shown in **Supplementary Figure S4A**. Increased synonymous site
38
39 317 conservation was independently observed in this region for betanarnaviruses (not shown). The
40
41 318 most highly conserved region in the putative rORF-encoded amino acid sequence was also directly
42
43 319 opposite the RdRp motif B, in a 15-amino acid window of the alignment centred on position 510
44
45 320 (**Figure 7B**), where the rORF-encoded consensus sequence is QvxExExxPREREAH.
46
47
48 321
49
50
51
52
53
54 322 **EVEs related to rORF-containing narnaviruses are found in insect WGS datasets**
55
56
57
58
59
60

1
2
3
4 323 Endogenized virus elements (EVEs) represent virus fragments which have been spuriously reverse-
5 transcribed and integrated into host genomes; EVEs can provide evidence as to the true host(s) of
6 virus groups (Katzourakis and Gifford 2010). To identify endogenized narnavirus-derived
7 sequences, we queried the ONLV1 and ONLV2 RdRp sequences against NCBI arthropod whole
8 genome shotgun (WGS) datasets. Nine matches were found, all with e-values $\leq 5 \times 10^{-15}$:
9 324 BHEC01027060.1 (*Coccinella septempunctata*), JJNS02035057.1 (*Homalodisca vitripennis*),
10 325 SJPC01005419.1 (*Cataglyphis niger*), NJRP01000653.1 and NJRP01005509.1 (*Aphaenogaster*
11 326 *floridana*), LBMM01010399.1 (*Lasius niger*), ADOQ01003831.1 and ADOQ01001783.1
12 327 (*Linepithema humile*) and NJRK01001360.1 (*Aphaenogaster rudis*), all of which derive from
13 insect datasets. To confirm that these EVEs clustered most closely with rORF-containing
14 narnaviruses, we performed reciprocal tblastn analysis against the NCBI virus nr/nt database. In all
15 328 nine cases, the top virus match was one of KX883548, MH213236, KX883539 or KP642119, all of
16 329 which are indeed insect-associated rORF-containing narnaviruses.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39 336
40
41 337 **DISCUSSION**
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 346 having AU-rich termini. These variations might, to some extent, reflect divergence in host taxa.
5
6 347 Although some of the available sequences are likely incomplete, there is evidence that members of
7
8 348 both proposed genera exhibit complementarity between the genomic termini. In ScNV-20S and
9
10 349 ScNV-23S, the 5'-GGGG and CCCC-3' genomic termini, besides 5' and 3' RNA structures, are
11
12 350 essential for efficient virus replication (Esteban and Fujimura 2003; Fujimura and Esteban 2004;
13
14 351 Esteban et al. 2005). Since narnaviral RNAs are non-polyadenylated and probably non-capped,
15
16 352 they are vulnerable to host mRNA degradation pathways. However, the viral RdRp binds to the 3'-
17
18 353 terminal CCCC and an adjacent RNA stem-loop structure in the positive strand and this is thought
19
20 354 to stabilize the genome and protect the 3' end from host 3' exonuclease degradation (Fujimura and
21
22 355 Esteban 2004; Fujimura and Esteban 2007). Similar elements also exist at the 3' end of the negative
23
24 356 strand (Esteban et al. 2005; Fujimura and Esteban 2007). Meanwhile, the positive-strand 5'-
25
26 357 proximal RNA structure protects the genome from the host SKI1/XRN1 5' exonuclease (Esteban et
27
28 358 al. 2008). These stability elements are thought to be particularly important for narnaviruses due to
29
30 359 the lack of capsids or membrane-associated replication that other (+)ssRNA viruses use to protect
31
32 360 their genomes (Fujimura and Esteban 2007).

33
34
35
36 361 The diversity of sources of co-clustering narnaviral genomes suggests either that horizontal transfer
37
38 362 of narnaviruses between divergent hosts has occurred, or that at least some of these sources are not
39
40 363 the *bona fide* host species, potentially reflecting contamination from parasitic or commensal
41
42 364 organisms, gut contents, or external debris. For example, in **Figure 2**: two TSA sequences from the
43
44 365 fly-infecting fungus *Entomophthora muscae* (GENC01006608.1 and GEND01011317.1) cluster
45
46 366 within a clade of arthropod-derived sequences; a TSA sequence (GFKT011160020.1) from the
47
48 367 spider *Nephila clavipes* has 95% nucleotide identity to ScNV-20S (AF039063.1); and two TSA
49
50 368 sequences (GGCO01105932.1 and GGCO01034162.1) from the barley plant *Hordeum vulgare*
51
52 369 cluster with sequences derived from the Basidiomycota obligate plant pathogenic fungi *Uromyces*

1
2
3
4 370 *appendiculatus* and *Puccinia striiformis*. Since ScNV-20S is a well-characterized virus of
5
6 371 *Saccharomyces cerevisiae*, it is likely that the *Nephila clavipes* sequence derives from fungal
7
8 372 contamination. The *Entomophthora muscae* sequences were obtained from fungus-infected *Delia*
9
10 373 *radicum* cabbage flies, and indeed the NCBI SRA taxonomy analysis webpages (alpha version, 30
11
12 374 May 2019) for the corresponding RNA-Seq libraries show 4–16 times as much fly RNA as fungal
13
14 375 RNA (within the only 8–9% of reads that were taxonomically identified); thus these TSAs may
15
16 376 derive from infected insect cells. Similarly, the *Hordeum vulgare* TSAs may derive from infected
17
18 377 fungal cells since the NCBI SRA taxonomy analysis webpages show contamination with
19
20 378 Opisthokonta (i.e. animal/fungi) RNA at ~4% the level of plant RNA. Thus, it is important to apply
21
22 379 caution when attempting to infer the hosts of TSA sequences.
23
24
25
26
27
28 380 Given the co-occurrence of fungal sequences in arthropod transcriptomic data sets, it was originally
29
30 381 suggested that the identified narnavirus sequences might derive from fungal contaminants
31
32 382 (Chandler et al. 2015; Cook et al. 2013). More recently, evidence was put forward that at least
33
34 383 some narnaviruses **may be** true arthropod viruses: for example, the group of alphanarnaviruses
35
36 384 comprising KF298275.1, KF298276.1, KF298284.1, Zhejiang mosquito virus 3, KP642119.1 and
37
38 385 KP642120.1 (**Figure 2**) come from four different studies and five different Culicinae mosquito
39
40 386 species (Chandler et al. 2015; Cook et al. 2013; Shi et al. 2016; Shi et al. 2017); moreover, in
41
42 387 several samples, viral RNA accounts for >0.1% (in one case, >2%) of total non-ribosomal RNA
43
44 388 reads, which **may be** unlikely if the virus is infecting a contaminant (Shi et al. 2016; Shi et al.
45
46 389 2017). Very recently, another narnavirus in this clade (MK628543.1; Culex narnavirus 1; 97% nt
47
48 390 identity to KP642120.1) was found to persistently infect a *Culex tarsalis* cell culture and also give
49
50 391 rise to typical 21-nt viral siRNAs with equal coverage of both strands, indicative of active infection
51
52 392 (Göertz et al. 2019). **An ONLV2-related EVE has been previously reported for arthropods (Shi et**
53
54 393 **al. 2016), and we also identified sequences that clustered with rORF-containing narnaviruses in**

1
2
3
4 394 WGS assembly data sets from several species of ant, the seven spot ladybird and the glassy-winged
5 sharpshooter. On the other hand, it has also been suggested that arthropod-associated narnaviruses
6
7 395 may instead be infecting trypanosomatids – common intracellular parasites of arthropods (Harvey
8 et al., 2019) (trypanosomatid DNA – including EVEs – could potentially also contaminate WGS
9 datasets of their hosts). Other rORF-containing alphanarnaviruses come from fungal (or plant)
10 samples that have no obvious association with arthropods (< 0.8% metazoan-mapping reads in total
11 per sample), such as the *Uromyces appendiculatus* and *Puccinia striiformis* samples, indicating that
12 presence of the rORF is likely not a unique adaptation to arthropod-association.
13
14

15
16 402 Since our original identification of long rORFs in the ONLV1 and ONLV2 genomes, besides
17 sequences from *Uromyces appendiculatus* and *Puccinia striiformis* transcriptomes (Cook et al.
18 2013), a number of other studies have discovered narnaviral genomes containing similar rORFs
19 (Chandler et al. 2015; Shi et al. 2016; Shi et al. 2017; Viljakainen et al. 2018; Göertz et al. 2019).
20
21 406 These rORF-containing sequences appear not to form a monophyletic group, although they do
22 show considerable phylogenetic clustering. The rORF occurs exclusively in the R0 frame relative
23 to the RdRp ORF, and large sections of the latter therefore exhibit a specific exclusion of the three
24 codons CUA, UUA and UCA, which are however widely used in the RdRp ORF of non-rORF-
25 containing narnaviruses. Nonetheless, these codons are still used occasionally in rORF-containing
26 narnaviruses, exclusively within the 5'-most extremity (3%) of the RdRp coding region (at or
27 upstream of the rORF stop codon). This indicates that selection against these codons is not due to
28 an inability for these codons to be decoded by the host translational machinery (e.g. due to the lack
29 of cognate tRNAs). This is also supported by our analysis of cellular organism codon usage which
30 showed that the vast majority of organisms with GC3 in the same range as rORF-containing
31 narnaviruses have CUA, UUA and UCA mean codon usage values well above those of the rORF-
32
33

1
2
3
4 417 containing narnaviruses (**Figure 6B**). Thus, the presence of the long rORF in highly divergent
5
6 418 alphanarnaviruses cannot be explained as an artefact of RdRp ORF codon usage or GC bias.
7
8
9

10 419 To our knowledge there is no experimental evidence for protein-coding ORFs in the negative
11 strand of any (+)ssRNA virus. A recent bioinformatic study used single-sequence randomization
12 procedures to identify ORFs overlapping previously annotated ORFs in RNA virus sequences that
13 are statistically significantly longer than expected by chance, with the assumption being that such
14 ORFs are likely to be functional (Schlub et al. 2018). The authors identified statistically
15 significantly long negative-strand stop-codon-free regions overlapping positive-strand ORFs in
16 sweet potato virus 2 (NC_017970.1; family *Potyviridae*), Macrophomina phaseolina tobamo-like
17 virus (NC_025674.1; family *Virgaviridae*), Nhumirim virus (NC_024017.1; family *Flaviviridae*),
18 hibiscus chlorotic ringspot virus (NC_003608.1; family *Tombusviridae*), hydrangea ringspot virus
19 (NC_006943.1; family *Alphaflexiviridae*), and Scrophularia mottle virus (NC_011537.1; family
20 *Tymoviridae*). However it remains unclear how these ORFs would be translated. In eukaryotes,
21 translation normally relies on recruitment of pre-initiation ribosomes to the 5' end of mRNAs,
22 followed by scanning and initiation at the first AUG in a good initiation context. Thus, the majority
23 of *de novo* protein-coding ORFs tend to evolve towards the 5' ends of transcripts, frequently
24 overlapping the 5' end of an ancestral protein-coding ORF where they can be translated via a
25 process known as leaky scanning (Firth and Brierley 2012). If protein-coding ORFs exist in the
26 negative strand of (+)ssRNA viruses, one might expect the majority to have initiation sites close to
27 the 5' end of the negative strand. However for the six ORFs above there are, respectively, 171, 101,
28 115, 8, 3 and 100 intervening AUG codons between the 5' end of the negative strand and the first
29 in-frame AUG codon in the ORF, which would appear to rule out 5'-end-dependent scanning as a
30 translation mechanism. 5'-distal ORFs on positive-sense transcripts are often translated via
31 ribosomal frameshifting or stop codon readthrough, but that still requires 5'-proximal initiation in
32

1
2
3
4 441 the pre-fameshift/pre-readthrough ORF (Firth and Brierley 2012). 5'-distal ORFs are occasionally
5
6 442 translated via an internal ribosome entry site (IRES), but IRESes are normally complex elements
7
8 443 and it would be difficult for them to evolve within protein-coding sequences. Simple inefficient
9
10 444 IRESes are a possibility and, compounded with the low availability of negative strand, would lead
11
12 445 to extremely low expression levels of any resulting proteins. Splicing is unknown in (+)ssRNA
13
14 446 viruses as they all replicate cytoplasmically. Subgenome-sized negative-sense transcripts, where
15
16 447 they are produced, are normally 5' co-terminal with the full-length negative strand so would not
17
18 448 normally provide access to internal ORFs, although the production of other classes of negative-
19
20 449 sense transcripts is plausible (Sztuba-Solińska et al. 2011). Given the pronounced evolutionary
21
22 450 adaptations that might be required to express such negative-strand ORFs, in the absence of
23
24 451 experimental data they would be more plausible if they were conserved in related species. However
25
26 452 this has not been demonstrated. For example, the Nhumirim virus rORF is not conserved in the
27
28 453 next most closely related flavivirus genome sequence, that of Barkedji virus (MG214906.1; 71% aa
29
30 454 identity in the forward ORF) where the rORF region is disrupted by eight stop codons. Similarly,
31
32 455 the sweet potato virus 2 rORF is not conserved in sweet potato virus G (NC_018093.1; 74% aa
33
34 456 identity in the forward ORF) where the rORF region is disrupted by nine stop codons. Indeed it is
35
36 457 not even conserved in other isolates of sweet potato virus 2 – e.g. KP729268.1 and KP115618.1
37
38 458 each have a stop codon. Thus we contend that the alphanarnavirus rORF is currently the only
39
40 459 plausible candidate for a negative-strand coding ORF in (+)ssRNA viruses.

41
42
43
44
45
46
47
48
49 460 The replication of most (+)ssRNA viruses is membrane-associated (den Boon and Ahlquist 2010;
50
51 461 Romero-Brey and Bartenschlager 2014; Shulla and Randall 2016; Ertel et al. 2017). Infection
52
53 462 typically results in extensive rearrangement of host membranes to produce membrane-bound mini-
54
55 463 organelles within which genome replication occurs. These structures co-localize viral RNA and
56
57 464 proteins and may also shield viral positive:negative strand duplexes from recognition by host

1
2
3
4 465 dsRNA-recognizing antiviral factors such as RIG-I, MDA-5, PKR and the RNAi machinery.
5
6 466 Sequestering replication within membranous compartments may also play a role in separation of
7
8 467 translation from replication: positive-sense mRNAs are extruded into the cytoplasm for translation,
9
10 468 whereas the negative strand remains protected within the membranous compartment for replication.
11
12 469 This prevents ribosomes translating 5' to 3' from colliding with RdRps tracking 3' to 5' on the same
13
14 470 template. Thus, at least at later time points, the negative strand is expected to be generally
15
16 471 sequestered away from the translational machinery. Further, the negative strand of (+)ssRNA
17
18 472 viruses often lacks the 5' structures (such as a 5' cap, 5' covalently linked viral protein of the
19
20 genome, or IRES structure) required to efficiently recruit ribosomes. Narnavirus replication is,
21
22 473 however, quite atypical among eukaryote-infecting (+)ssRNA viruses, and this may explain why
23
24 474 they – possibly uniquely – appear to have evolved negative-strand coding capacity. Narnavirus
25
26 475 replication is thought to occur entirely within the cytoplasm, not associated with cellular
27
28 476 membranes (Solórzano et al. 2000; Fujimura et al. 2005). As mentioned above, the narnavirus
29
30 477 RdRp is more closely related to the RdRp of bacteriophages in the family *Leviviridae* than it is to
31
32 478 the RdRp of other (+)ssRNA viruses. Similar to leviviruses, the RNA in narnavirus replication
33
34 479 intermediates is essentially single-stranded (Blumenthal and Carmichael 1979; Dobkin et al. 1979;
35
36 480 Zinder 1980; Takeshita and Tomita 2012; Fujimura et al. 2005; Wickner et al. 2013). Each newly
37
38 481 synthesized RNA is released before a new round of replication commences and resting complexes
39
40 482 comprising a negative-sense or positive-sense single-stranded RNA bound to a single copy of the
41
42 483 viral RdRp are present within the cytoplasm (Solórzano et al. 2000). Thus, for narnaviruses, the
43
44 484 positive and negative strands likely have similar accessibility to the host cell translational
45
46 485 machinery.

54
55
56 487 The rORF is present in many but not all alphanarnaviruses. Surprisingly, the distribution of the
57
58 488 rORF does not appear to be monophyletic. Thus it may have evolved multiple times, or it may have
59
60

been present ancestrally and lost from some lineages. The observation that the rORF is always in the R0 frame relative to the RdRp ORF may be evidence against independent evolution in multiple lineages, although it could also be a consequence of codon bias, e.g. if long ORFs are by chance more likely to occur in the R0 frame as a result of RdRp ORF codon usage. Such random ORFs could provide “seeds” for the evolution of longer overlapping genes (Belshaw et al. 2007). Another surprising feature is that the rORF, where present, is nearly always full-length. If the sole function of the rORF were to encode an additional protein, one might expect to see a variety of rORF lengths in different lineages whereas we only see a single example of an rORF beginning within 200 nt of the 5' end of the negative strand with length in the range 24–94% of the full sequence length (**Figure 4**). In contrast to the RdRp, we found little evidence for conserved amino acid motifs in the rORF protein (**Figure 7**). It is possible that the very high divergence between most of our sequences might obscure a weak conservation signature (in contrast to the RdRp which is the most highly conserved RNA virus protein known). However it is also possible that the protein product of the rORF is not itself functionally important. A possible alternative explanation is that rORF translation in itself might facilitate replication e.g. by disassociating dsRNA, or by increasing negative strand RNA stability (Bicknell and Ricci 2017). Alternatively, if the negative strand is unavoidably accessible for translation, exon-junction-complex-independent nonsense mediated decay (Kurosaki et al. 2019) may provide strong selection pressure against non-full-length rORFs.

Although, as discussed above, the negative strands of most (+)ssRNA viruses are normally expected to be occluded in membranous compartments, it is currently unclear whether at early timepoints – before extensive membranous compartments have formed – negative strands produced in a first round of replication might be exposed to the cytoplasm and potentially available for translation (Shulla and Randall 2015). It is also not inconceivable that some (+)ssRNA viruses might have evolved mechanisms to allow some negative strands access to the translational

1
2
3
4 513 machinery even at late timepoints. If so, it is possible that some (+)ssRNA viruses might have
5 514 evolved ribosome recruitment elements within their negative strands. Thus it is perhaps too early to
6 515 dismiss the possibility of negative-strand ORFs in other (+)ssRNA virus lineages. In most cases,
7 516 however, any proteins encoded on the negative strand would be expected to be expressed at much
8 517 lower levels than positive-strand encoded proteins simply because of the huge disparity in
9 518 positive:negative RNA abundance during virus infection (typically of order 100:1; Novak and
10 519 Kirkegaard 1991; Kopek et al. 2007; Irigoyen et al. 2016). Even for narnaviruses, less than 1–2%
11 520 of ScNV-20S viral RNA in infected cells is negative-sense (Rodríguez-Cousío et al. 1991;
12 521 Fujimura et al. 2005), indicating that the rORF, where present, is likely to be expressed at a much
13 522 lower level than the RdRp.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46 523 In conclusion, we have provided the first strong evolutionary evidence for reverse-**strand** coding
47 capacity in a group of positive-sense RNA viruses. The alphanarnavirus rORFs are exceedingly
48 long compared to most known overlapping genes (Pavesi et al. 2018; Brandes and Linial 2016).
49 Thus their study is not only of interest to the evolution and molecular biology of viruses, but also
50 of broad relevance to understanding the evolution of overlapping genes and the *de novo* origin of
51 genes. During revision of this manuscript, a related study was also published (DeRisi et al. 2019).
52
53
54
55
56
57
58
59
60

530 METHODS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50 531 The NCBI non-redundant nucleotide (nr/nt) and transcriptome shotgun assembly (TSA) databases
51 532 were downloaded on August 31, 2018. Narnaviral sequences were identified in these databases
52 533 using, tblastn (version 2.8.0) (Camacho et al. 2009), with the RdRp protein sequence of ScNV-20S
53 534 and the updated RdRp sequence of ONLV1 (completed in this study) as queries. In total, 81
54 535 sequences were found by combined searches using these queries (46 in the nr database and 35 in
55
56
57
58
59
60

1
2
3
4 536 the TSA database). Hidden Markov model (HMM)-based searches were carried out using hhsearch
5
6 537 (Finn et al. 2011), based on an alignment of RdRps from rORF-containing narnaviruses, leading to
7
8 538 the identification of a further 43 TSA database sequences.
9
10
11

12 539 To identify narnavirus-derived EVEs, we queried ONLV RdRp sequences AGW51766.2 and
13
14 540 AGW51768.2 against the NCBI WGS database on 18 December 2019, using the NCBI tblastn
15
16 interface, with default parameters except the taxonomy ID was set to “Arthropoda” and the expect
17
18 541 threshold and word size were set to 0.1 and 6 respectively. Match sequences were downloaded, the
19
20 match region extracted, translated, and reciprocally queried against the NCBI nr/nt database with
21
22 543 tblastn, with the taxonomy ID set to “viruses” and other parameters as above.
23
24 544
25
26

27 545 Phylogenetic trees were generated with MrBayes (version 3.2.7) (Huelsenbeck and Ronquist 2001),
28
29 546 using a mixed substitution model with sampling across fixed amino acid rate matrices (aamodelpr
30
31 547 = mixed) and 5,000,000 generations. All other parameters were set as defaults.
32
33
34

35 548 To check for homology to known proteins, we performed blastp searches against the NCBI non-
36
37 549 redundant (nr) protein database, using each of the protein sequences predicted to be encoded by an
38
39 550 rORF separately as a query. We also applied HHpred (Zimmermann et al. 2018), using rORF
40
41 551 amino acid alignments as queries (five alignments, one for each of the main rORF clades in **Figure**
42
43 552 **2**, with – for definiteness – IABX01132835, KX883605, KX883500, excluding KX883602 being
44
45 553 one of these clades), searching against PDB_mmCIF70_11_Oct and Pfam-A_v32.0.
46
47
48

49 554 To check for potential capsid-encoding ORFs, all positive-strand ORFs ≥ 150 codons (between
50
51 555 flanking stop codons and/or sequence ends), excluding the RdRp ORF itself, were identified and
52
53 556 used as queries in HHpred. Among the 29 ORFs found in 124 narnaviral sequences, none were
54
55 557 found to have matches to viral capsid proteins; indeed none had significant e-value matches to any
56
57 558 known proteins (the minimum e-value was 10).
58
59 559
60

1
2
3
4 559 RNA structures in the genomic terminal regions were predicted by scanning full-length narnaviral
5 genomes with RNALfold (version 2.4.9) (Lorenz et al. 2011), allowing a maximum base-pair span
6
7 560 of 150 nt.
8
9
10
11

12 562 For all comparisons of alphanarnaviruses “with” and “without” rORFs, sequences in which the
13 longest stop codon-free region (beginning within the 5'-most 200 nt) occupied at least 90% of the
14 genomic sequence were assigned to the former group, whereas other sequences were assigned to
15
16 564 the latter group.
17
18
19 565
20
21

22 566 A codon-based multiple sequence alignment of RdRp ORFs was produced by aligning the RdRp
23 amino acid sequences with MUSCLE (version 3.8.31) (Edgar 2004) and then backtranslating to a
24 nucleotide sequence alignment using the trnalnalign program in the EMBOSS suite (version 6.6.0.0)
25 (Rice et al. 2000). The alignment was then mapped to ONLV2 sequence coordinates by removing
26 alignment positions that contained a gap character in the ONLV2 sequence. Synonymous site
27 variation was assessed using synplot2 (Firth 2014), with a window size of 15 codons. Comparison
28 of amino acid sequences was performed using plotcon from the EMBOSS suite (version 6.6.0.0)
29 (Rice et al. 2000) with a BLOSUM62 substitution matrix and a window size of 15 amino acids.
30
31 570
32
33 571
34
35 572
36
37 573
38
39 574 When analysing synonymous site **conservation** in the rORF, the reverse complement of the RdRp
40 alignment was used, and aligned nucleotides on the reverse strand were converted to the
41 corresponding protein sequences using AMAS (Borowiec 2016).
42
43
44
45
46

47
48 577 Principal component analysis (PCA) of codon usage across species was performed using relative
49 codon abundances (i.e. normalized by the abundances of the associated amino acids). The principal
50 components were calculated using the RefSeq sequences of cellular organisms, and narnaviral
51 codon abundances were projected onto the resulting principal component space.
52
53
54
55
56
57
58
59 581
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

582 ACKNOWLEDGEMENTS

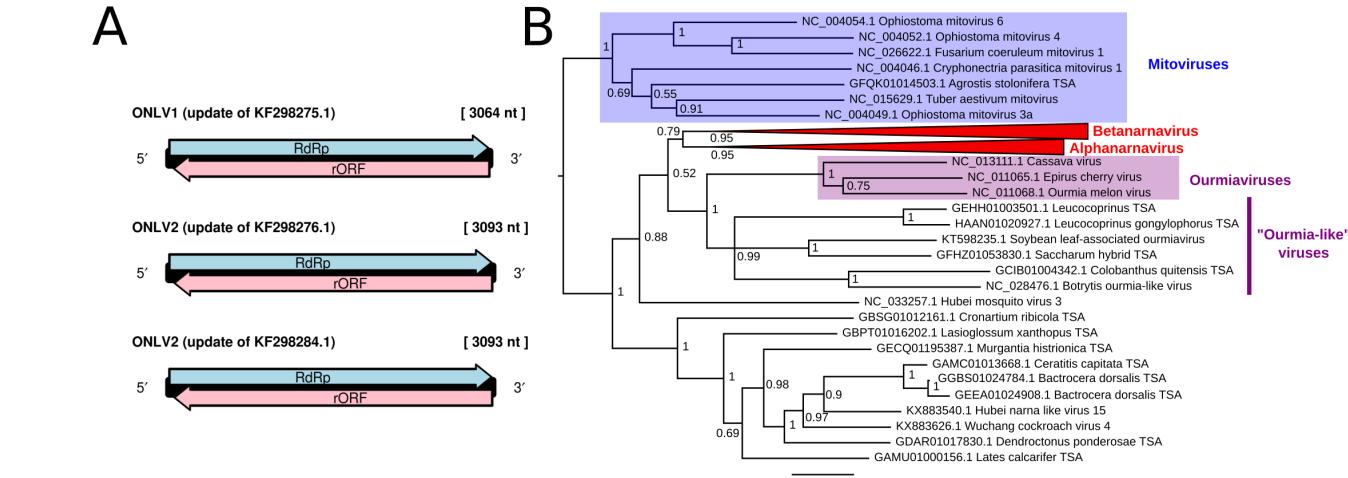
583 We thank Chris McCormick, Katherine Brown, Valeria Lulla and Hazel Stewart for useful
584 discussions. Work was supported by Wellcome Trust grant [106207] and European Research
585 Council grant [646891] to A.E.F.

586

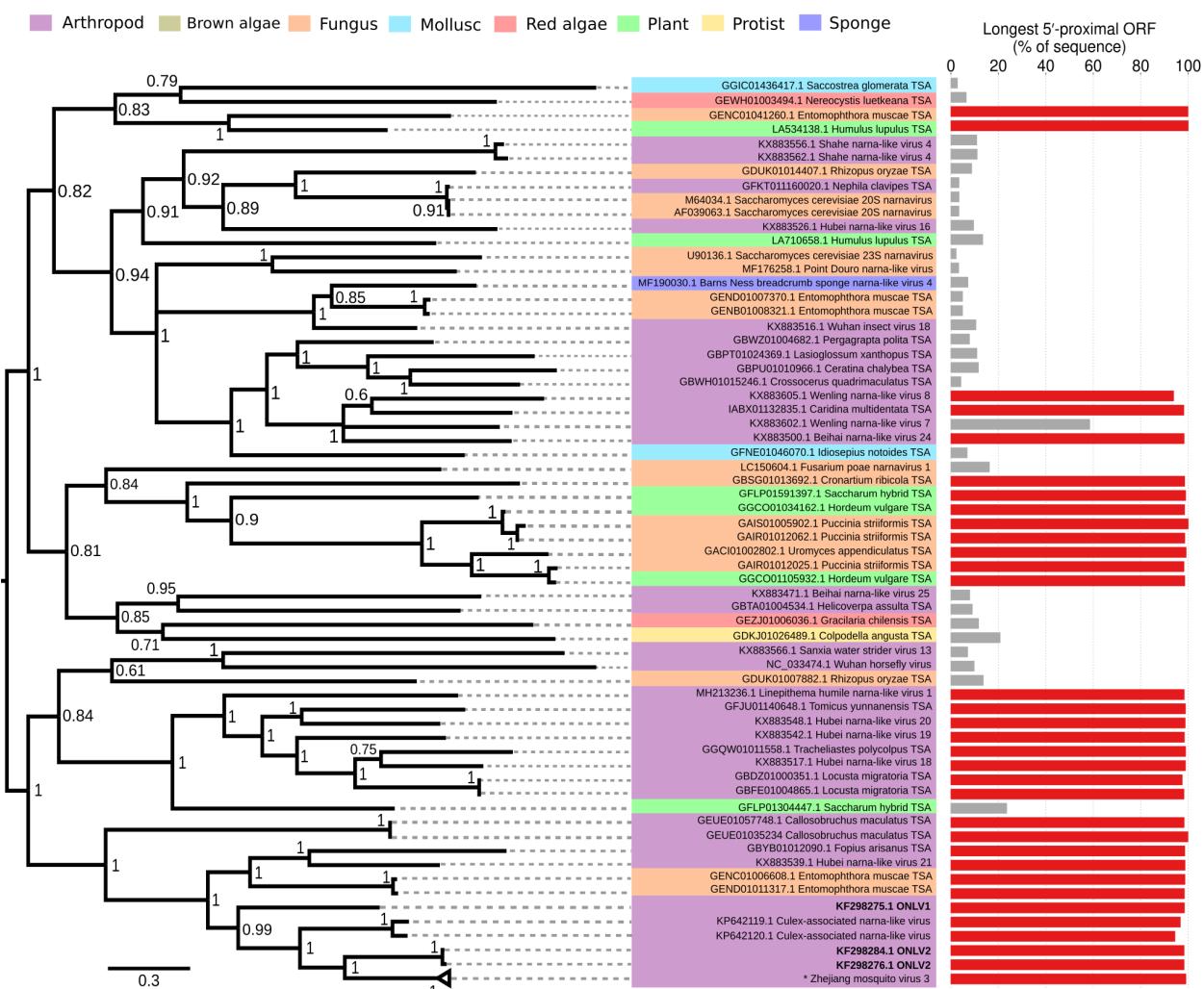
587 DATA AVAILABILITY

588 Data available in supplementary material and in GenBank accession numbers KF298275.2,
589 KF298276.2 and KF298284.2.

590

591 **FIGURES**

592 **Figure 1.** Narnavirus genome structure and taxonomy. **(A)** Updated narnaviral genome sequences:
593 genome structures of ONLV1 and ONLV2. **(B)** Bayesian phylogeny of narnaviruses and selected
594 outgroup sequences. Alphanarnaviruses and betanarnaviruses form clades with posterior
595 probabilities of 0.95. The tree is rooted with the mitoviral clade as an outgroup. Accession numbers
596 are for the nucleotide sequences from which the corresponding protein sequences were derived.



598 **Figure 2.** Midpoint-rooted Bayesian phylogenetic tree of alphanarnaviruses. The longest 5'-
 599 proximal ORF in the negative-strand R0 frame is shown in the bars to the right. Sequences with
 600 R0-frame ORFs occupying over 90% of the sequence are indicated with red bars. Nine highly
 601 similar sequences for Zhejiang mosquito virus 3 (indicated with an asterisk) are collapsed to a
 602 single taxon. Accession numbers are for the nucleotide sequences from which the corresponding
 603 protein sequences were derived.

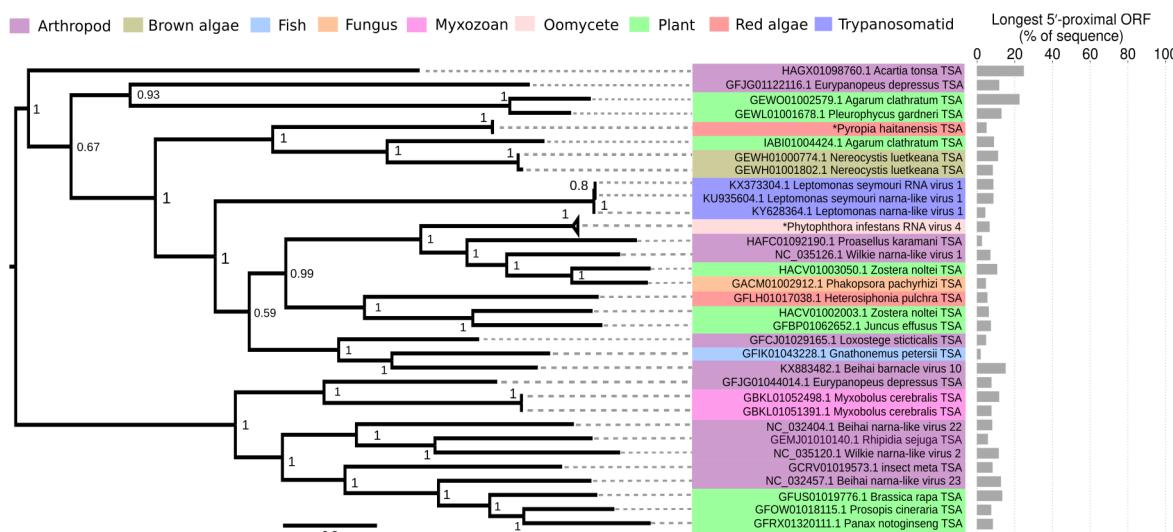


Figure 3. Midpoint-rooted Bayesian phylogenetic tree of betanarnaviruses. As in Figure 2, the longest 5'-proximal ORF in the negative-strand R0 frame is shown in the bars to the right. Five highly similar TSA sequences for *Pyropia haitanensis* and six highly similar sequences for *Phytophthora infestans* RNA virus 4 (both indicated with asterisks) are collapsed to single nodes. Accession numbers are for the nucleotide sequences from which the corresponding protein sequences were derived.

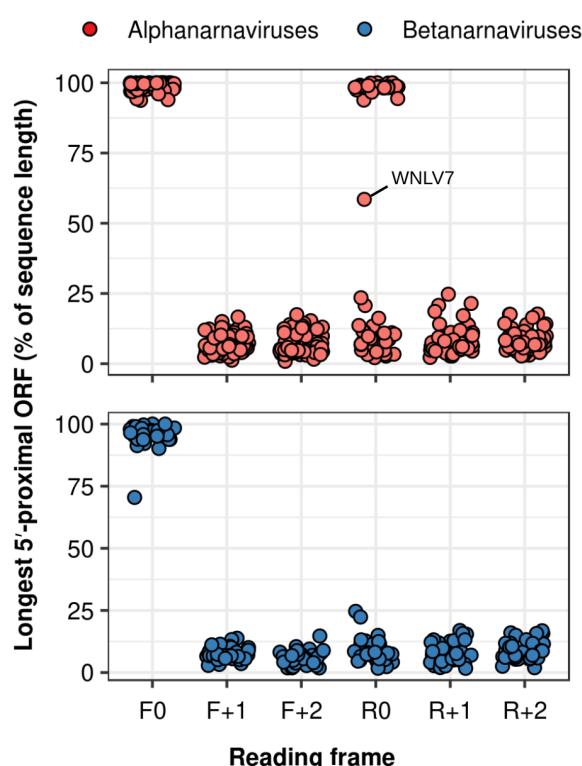


Figure 4. The longest 5'-proximal stop codon-free regions in each of the three possible positive-strand and negative-strand reading frames, for alphanarnaviruses (red) and betanarnaviruses (blue), as a percentage of the sequence length. Wenling narna-like virus 7 (WNLV7) has an intermediate-length R0-frame ORF, as indicated. Mean values are plotted for nodes with high levels of representation in the underlying data set (i.e. Zhejiang mosquito virus 3, Phytophthora infestans RNA virus 4, and TSA sequences from *Pyropia haitanensis*).

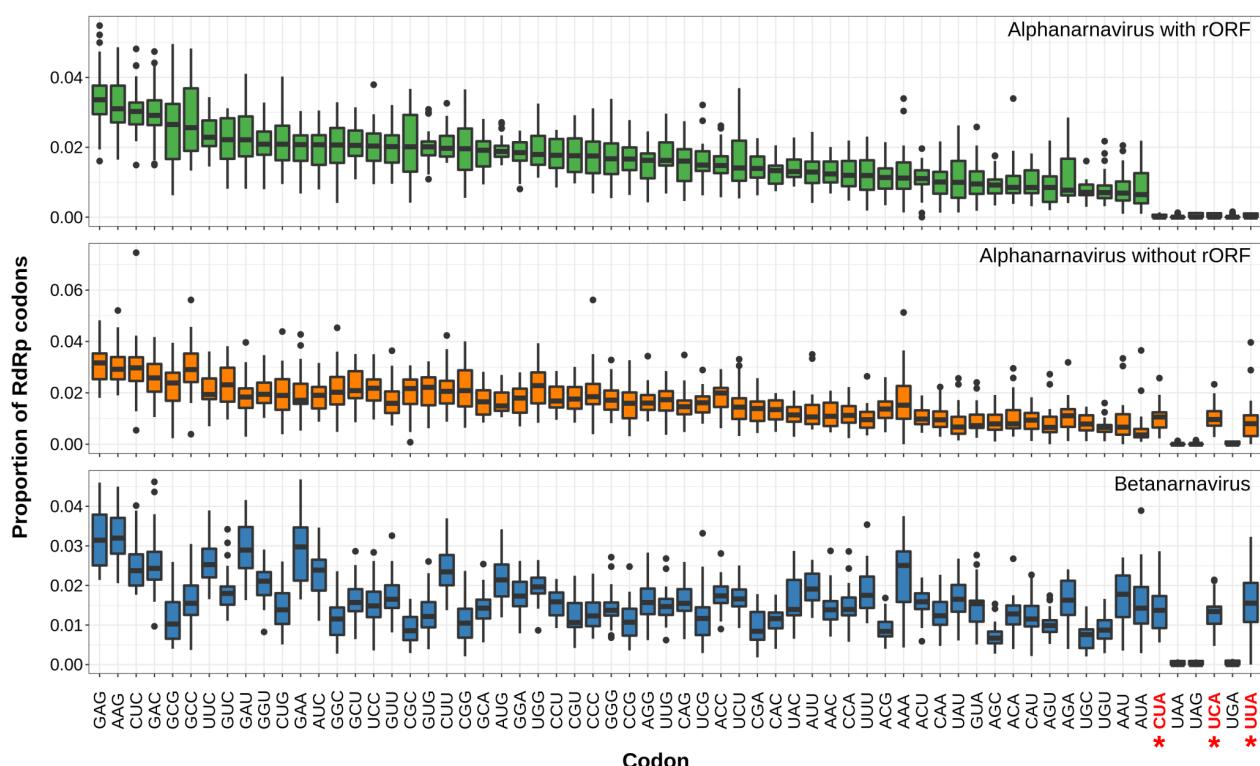


Figure 5. Box plots of codon usage (proportion of total codons) in RdRp ORFs. The upper and lower hinges are located at the first and third quartiles, respectively; the median values are indicated as horizontal lines; and the whiskers extend from the hinges to the furthest data points within 1.5 times the interquartile range (IQR) of the hinges. Data more than 1.5 times the IQR from the hinges are drawn as individual points (outliers). Codons are ordered according to their median frequency in alphanarnaviruses that contain the rORF. Alphanarnaviruses that contain the rORF show a specific avoidance of the three codons marked with red asterisks (CUA, UUA and UCA), which correspond to the reverse complements of stop codons.

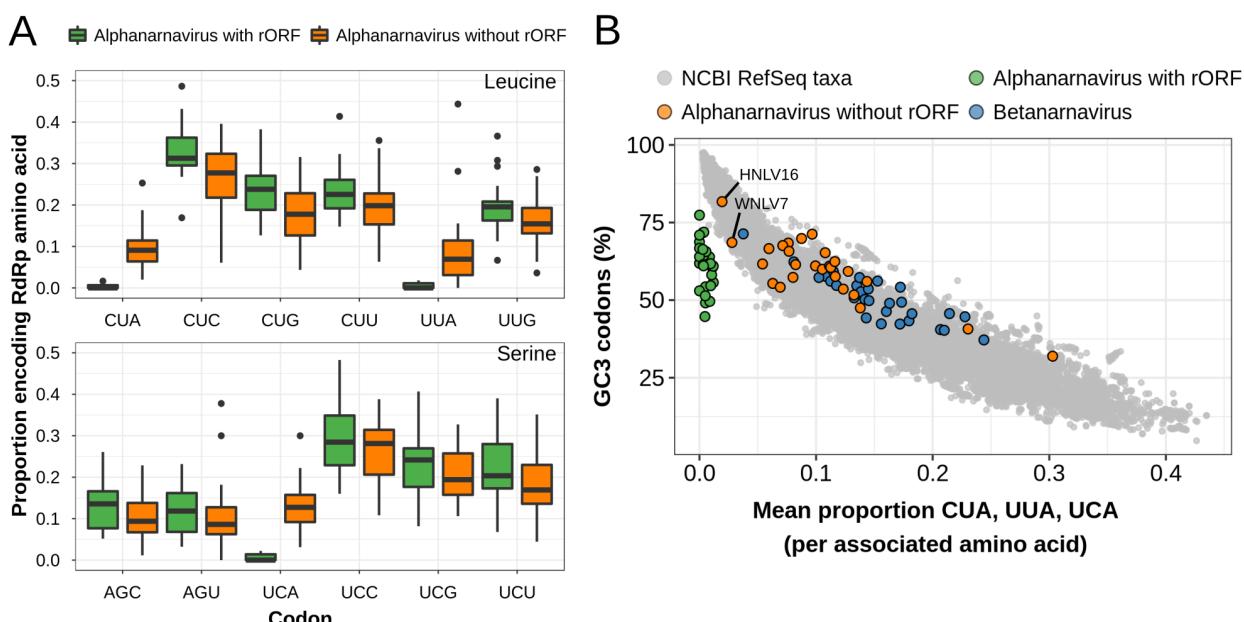
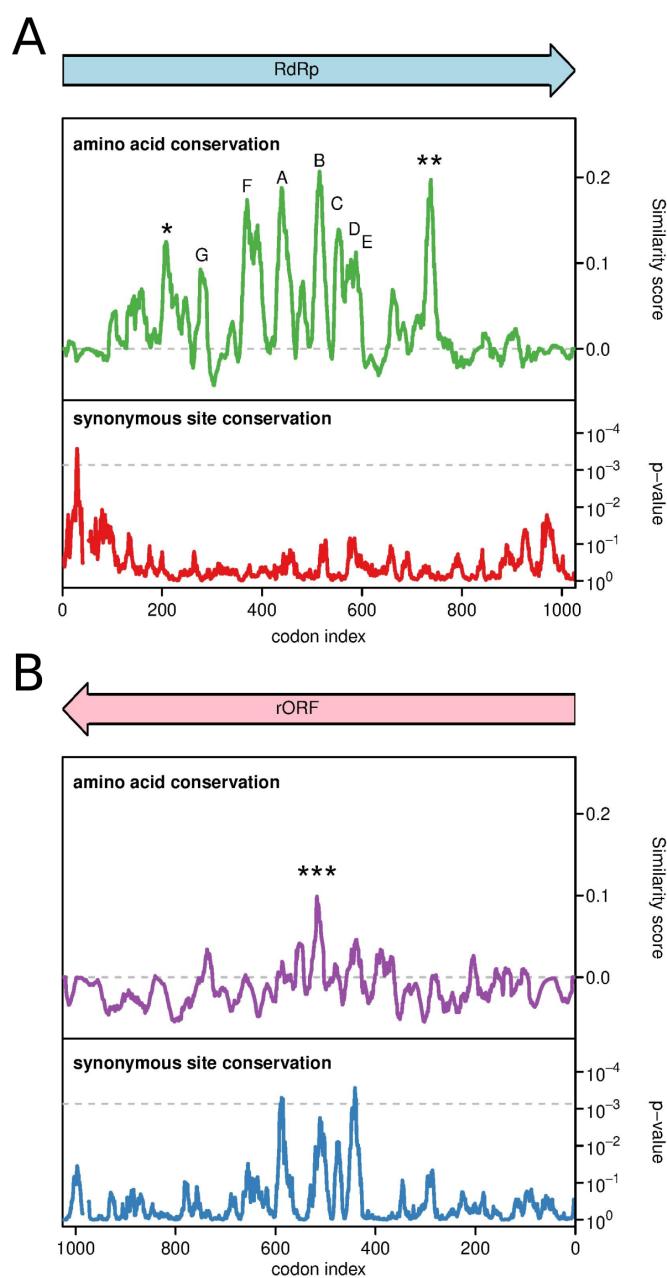


Figure 6. Specific selection against CUA, UUA and UCA codons in rORF-containing alphanarnaviruses. **(A)** Box plots of codon usage bias (proportion of codons encoding the amino acid) for leucine and serine. Box plots are drawn as in Figure 5. **(B)** Comparison of mean usage of CUA, UUA and UCA (per associated amino acid) with the third-position GC (GC3) content of codons. Prokaryotic and eukaryotic taxa in NCBI RefSeq are shown as grey points. The RdRp ORFs of non-rORF alphanarnaviruses, rORF alphanarnaviruses and betanarnaviruses are indicated with orange, green and blue points, respectively. The Wenling narna-like virus 7 (WNLV7) and Hubei narna-like virus 16 (HNLV16) sequences are indicated.



45 638 **Figure 7.** Amino acid conservation and synonymous site conservation in rORF-containing
46 639 alphanarnaviruses. **(A)** RdRp amino acid conservation, and the corresponding synonymous site
47 640 conservation analysis. In the latter, the grey dashed line indicates an approximate 5% false positive
48 641 threshold after correcting for multiple tests (i.e. ~69 non-overlapping/independent 15-codon
49 642 windows in the 1027-codon ORF). **(B)** The equivalent plots for the rORF. In each case,
50 643 conservation was assessed in a 15 codon/amino acid window. Motifs A–G in the RdRp amino acid
51 644 sequence are indicated with letters. Additional consensus motifs indicated with asterisks are as

1
2
3
4 645 follows: * = R&UP, ** = Pxx[L/V]GGx[G/N]xP and *** = QvxExExxPREREAH; U = I, L, V or
5
6 646 M, & = I, L, V, M, A, P, G, F, W or Y.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

647 **REFERENCES**

- Athey J, Alexaki A, Osipova E, Rostovtsev A, Santana-Quintero LV, Katneni U, Simonyan V, Kimchi-Sarfaty C. 2017. A new and updated resource for codon usage tables. *BMC Bioinformatics* 18:391–391.
- Baltimore D. 1971. Expression of animal virus genomes. *Bacteriol Rev* 35:235.
- Belshaw R, Pybus OG, Rambaut A. 2007. The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res* 17:1496–1504.
- Bicknell AA, Ricci EP. 2017. When mRNA translation meets decay. *Biochem Soc Trans* 45:339–351.
- Blumenthal T, Carmichael GG. 1979. RNA replication: function and structure of Qbeta-replicase. *Annu Rev Biochem* 48:525–548.
- den Boon JA, Ahlquist P. 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu Rev Microbiol* 64:241–256.
- Borowiec ML. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4:e1660–e1660.
- Brandes N, Linial M. 2016. Gene overlapping and size constraints in the viral world. *Biol Direct* 11:26–26.
- Cai G, Myers K, Fry WE, Hillman BI. 2012. A member of the virus family *Narnaviridae* from the plant pathogenic oomycete *Phytophthora infestans*. *Arch Virol* 157:165–169.

1
2
3
4 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009.
5
6 BLAST+: architecture and applications. BMC Bioinformatics 10:421.
7
8
9

10 Chandler JA, Liu RM, Bennett SN. 2015. RNA shotgun metagenomic sequencing of northern
11 California (USA) mosquitoes uncovers viruses, bacteria, and fungi. Front Microbiol 6:185–
12 185.
13
14
15

16 Chung BY-W, Miller WA, Atkins JF, Firth AE. 2008. An overlapping essential gene in the
17 19 *Potyviridae*. Proc Natl Acad Sci USA 105:5897.
20
21
22

23 Cook S, Chung BY-W, Bass D, Moureau G, Tang S, McAlister E, Culverwell CL, Glücksmann E,
24 Wang H, Brown TDK, et al. 2013. Novel virus discovery and genome reconstruction from
25 field RNA samples reveals highly divergent viruses in dipteran hosts. PLoS One 8:e80720–
26 e80720.
27
28
29
30
31
32

33 DeRisi JL, Huber G, Kistler A, Retallack H, Wilkinson M, Yllanes D. 2019. An exploration of
34 ambigrammatic sequences in narnaviruses. Sci Rep 9:17982.
35
36
37

38 Dobkin C, Mills DR, Kramer FR, Spiegelman S. 1979. RNA replication: required intermediates
39 and the dissociation of template, product, and Q beta replicase. Biochemistry 18:2038–
40 2044.
41
42
43
44

45 Dolja VV, Koonin EV. 2011. Common origins and host-dependent diversity of plant and animal
46 viromes. Curr Opin Virol 1:322–331.
47
48
49

50 Dolja VV, Koonin EV. 2018. Metagenomics reshapes the concepts of RNA virus evolution by
51 revealing extensive horizontal virus transfer. Virus Research 244:36–52.
52
53
54
55
56
57
58
59
60

1
2
3
4 Donaire L, Rozas J, Ayllón MA. 2016. Molecular characterization of Botrytis ourmia-like virus, a
5 mycovirus close to the plant pathogenic genus *Ourmiavirus*. *Virology* 489:158–164.
6
7
8
9

10 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
11
12 Nucleic Acids Res 32:1792–1797.
13
14

15 Ertel KJ, Benefield D, Castaño-Diez D, Pennington JG, Horswill M, den Boon JA, Otegui MS,
16 Ahlquist P. 2017. Cryo-electron tomography reveals novel features of a viral RNA
17 replication compartment. *Elife* 6:e25940.
18
19
20
21
22

23 Esteban LM, Rodriguez-Cousío N, Esteban R. 1992. T double-stranded RNA (dsRNA) sequence
24 reveals that T and W dsRNAs form a new RNA family in *Saccharomyces cerevisiae*.
25 Identification of 23 S RNA as the single-stranded form of T dsRNA. *J Biol Chem*
26 267:10874–10881.
27
28
29
30
31
32

33 Esteban R, Fujimura T. 2003. Launching the yeast 23S RNA narnavirus shows 5' and 3' cis-acting
34 signals for replication. *Proc Natl Acad Sci U S A* 100:2568–2573.
35
36
37

38 Esteban R, Vega L, Fujimura T. 2005. Launching of the yeast 20 S RNA narnavirus by expressing
39 the genomic or antigenomic viral RNA in vivo. *J Biol Chem* 280:33725–33734.
40
41
42

43 Esteban R, Vega L, Fujimura T. 2008. 20S RNA narnavirus defies the antiviral activity of
44 SKI1/XRN1 in *Saccharomyces cerevisiae*. *J Biol Chem* 283:25812–25820.
45
46
47

48 Fang Y, Treffers EE, Li Y, Tas A, Sun Z, van der Meer Y, de Ru AH, van Veelen PA, Atkins JF,
49 Snijder EJ, et al. 2012. Efficient –2 frameshifting by mammalian ribosomes to synthesize
50 an additional arterivirus protein. *Proc Natl Acad Sci USA* 109:E2920.
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity
5 searching. Nucleic Acids Res 39:W29–W37.
6
7
8
9 Firth AE. 2014. Mapping overlapping functional elements embedded within the protein-coding
10 regions of RNA viruses. Nucleic Acids Res 42:12425–12439.
11
12
13
14 Firth AE, Brierley I. 2012. Non-canonical translation in RNA viruses. J Gen Virol 93:1385–1409.
15
16
17
18 Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation
19 sequencing data. Bioinformatics 28:3150–3152.
20
21
22
23 Fujimura T, Esteban R. 2004. The bipartite 3'-cis-acting signal for replication is required for
24 formation of a ribonucleoprotein complex in vivo between the viral genome and its RNA
25 polymerase in yeast 23 S RNA Virus. J Biol Chem 279:44219–44228.
26
27
28
29
30
31 Fujimura T, Esteban R. 2007. Interactions of the RNA polymerase with the viral genome at the 5'-
32 and 3'-ends contribute to 20S RNA narnavirus persistence in yeast. J Biol Chem
33 282:19011–19019.
34
35
36
37
38
39 Fujimura T, Solórzano A, Esteban R. 2005. Native replication intermediates of the yeast 20 S RNA
40 virus have a single-stranded RNA backbone. J Biol Chem 280:7398–7406.
41
42
43
44 Göertz GP, Miesen P, Overheul GJ, van Rij RP, van Oers MM, Pijlman GP. 2019. Mosquito small
45 RNA responses to West Nile and insect-specific virus infections in *Aedes* and *Culex*
46 mosquito cells. Viruses 11:271.
47
48
49
50
51
52
53 Harvey E, Rose K, Eden JS, Lawrence A, Doggett SL, Holmes EC. 2019. Identification of diverse
54 arthropod associated viruses in native Australian fleas. Virology 535:189–199.
55
56
57
58
59
60

1
2
3
4 Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees.
5
6 Bioinformatics 17:754–755.
7
8

9
10 Irigoyen N, Firth AE, Jones JD, Chung BY-W, Siddell SG, Brierley I. 2016. High-resolution
11 analysis of coronavirus gene expression by RNA sequencing and ribosome profiling. PLoS
12 Pathog 12:e1005473–e1005473.
13
14

15 Kadowaki K, Halvorson HO. 1971. Appearance of a new species of ribonucleic acid during
16 sporulation in *Saccharomyces cerevisiae*. J Bacteriol 105:826.
17
18

19 Kapoor A, Simmonds P, Lipkin WI, Zaidi S, Delwart E. 2010. Use of nucleotide composition
20 analysis to infer hosts for three novel picorna-like viruses. J Virol 84:10322–10328.
21
22

23 Katzourakis A, Gifford RJ. 2010. Endogenous viral elements in animal genomes. PLoS Genet
24 6:e1001191–e1001191.
25
26

27 Keese PK, Gibbs A. 1992. Origins of genes: "big bang" or continuous creation? Proc Natl Acad Sci
28 U S A 89:9489–9493.
29
30

31 Koonin EV, Dolja VV. 2014. Virus world as an evolutionary network of viruses and capsidless
32 selfish elements. Microbiol Mol Biol Rev 78:278.
33
34

35 Koonin EV, Dolja VV, Krupovic M. 2015. Origins and evolution of viruses of eukaryotes: The
36 ultimate modularity. Virology 479–480:2–25.
37
38

39 Kopek BG, Perkins G, Miller DJ, Ellisman MH, Ahlquist P. 2007. Three-dimensional analysis of a
40 viral RNA replication complex reveals a virus-induced mini-organelle. PLoS Biol 5:e220–
41 e220.
42
43

- 1
2
3
4 Kurosaki T, Popp MW, Maquat LE. 2019. Quality and quantity control of gene expression by
5 nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol* 20:406–420.
6
7
8
9 Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL.
10 2011. ViennaRNA Package 2.0. *Algorithms Mol Biol* 6:26.
11
12
13
14 Loughran G, Firth AE, Atkins JF. 2011. Ribosomal frameshifting into an overlapping gene in the
15 2B-encoding region of the cardiovirus genome. *Proc Natl Acad Sci USA* 108:E1111.
16
17
18
19
20 Lulla V, Dinan AM, Hosmillo M, Chaudhry Y, Sherry L, Irigoyen N, Nayak KM, Stonehouse NJ,
21 Zilbauer M, Goodfellow I, et al. 2019. An upstream protein-coding region in enteroviruses
22 modulates virus infection in gut epithelial cells. *Nat Microbiol* 4:280–292.
23
24
25
26
27
28 Lye L-F, Akopyants NS, Dobson DE, Beverley SM. 2016. A narnavirus-like element from the
29 trypanosomatid protozoan parasite *Leptomonas seymouri*. *Genome Announc* 4:e00713–16.
30
31
32
33
34 Napthine S, Ling R, Finch LK, Jones JD, Bell S, Brierley I, Firth AE. 2017. Protein-directed
35 ribosomal frameshifting temporally regulates gene expression. *Nat Commun* 8:15582.
36
37
38 Novak JE, Kirkegaard K. 1991. Improved method for detecting poliovirus negative strands used to
39 demonstrate specificity of positive-strand encapsidation and the ratio of positive to negative
40 strands in infected cells. *J Virol* 65:3384–3387.
41
42
43
44
45
46
47 Pavesi A, Vianelli A, Chirico N, Bao Y, Blinkova O, Belshaw R, Firth A, Karlin D. 2018.
48 Overlapping genes and the proteins they encode differ significantly in their sequence
49 composition from non-overlapping genes. *PLoS One* 13:e0202513–e0202513.
50
51
52
53
54
55
56
57
58
59
60

Rancurel C, Khosravi M, Dunker AK, Romero PR, Karlin D. 2009. Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. *J Virol* 83:10719–10736.

Rastgou M, Habibi MK, Izadpanah K, Masenga V, Milne RG, Wolf YI, Koonin EV, Turina M. 2009. Molecular characterization of the plant virus genus *Ourmiavirus* and evidence of inter-kingdom reassortment of viral genome segments as its possible route of origin. *J Gen Virol* 90:2525–2535.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* 16:276–277.

Rodríguez-Cousío N, Esteban LM, Esteban R. 1991. Molecular cloning and characterization of W double-stranded RNA, a linear molecule present in *Saccharomyces cerevisiae*. Identification of its single-stranded RNA form as 20 S RNA. *J Biol Chem* 266:12772–12778.

Rodríguez-Cousío N, Solórzano A, Fujimura T, Esteban R. 1998. Yeast positive-stranded virus-like RNA replicons: 20 S and 23 S RNA terminal nucleotide sequences and 3' end secondary structures resemble those of RNA coliphages. *J Biol Chem* 273:20363–20371.

Romero-Brey I, Bartenschlager R. 2014. Membranous replication factories induced by plus-strand RNA viruses. *Viruses* 6:2826–2857.

Schlub TE, Buchmann JP, Holmes EC. 2018. A simple method to detect candidate overlapping genes in viruses using single genome sequences. *Mol Biol Evol* 35:2572–2581.

- 1
2
3
4 Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, Li C-X, Qin X-C, Li J, Cao J-P, Eden J-S, et al.
5
6 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539.
7
8
9
10 Shi M, Neville P, Nicholson J, Eden J-S, Imrie A, Holmes EC. 2017. High-resolution
11 metatranscriptomics reveals the ecological dynamics of mosquito-associated RNA viruses
12 in Western Australia. *J Virol* 91:e00680–17.
13
14
15
16
17 Shulla A, Randall G. 2015. Spatiotemporal analysis of hepatitis C virus infection. *PLoS Pathog*
18 11:e1004758–e1004758.
19
20
21
22
23 Shulla A, Randall G. 2016. (+) RNA virus replication compartments: a safe home for (most) viral
24 replication. *Curr Opin Microbiol* 32:82–88.
25
26
27
28 Smirnova E, Firth AE, Miller WA, Scheidecker D, Brault V, Reinbold C, Rakotondrafara AM,
29 Chung BY-W, Ziegler-Graff V. 2015. Discovery of a small non-AUG-initiated ORF in
30 poleroviruses and luteoviruses that is required for long-distance movement. *PLoS Pathog*
31 11:e1004868–e1004868.
32
33
34
35
36
37
38 Solórzano A, Rodríguez-Cousío N, Esteban R, Fujimura T. 2000. Persistent yeast single-stranded
39 RNA viruses exist in vivo as genomic RNA·RNA polymerase complexes in 1:1
40 stoichiometry. *J Biol Chem* 275:26428–26435.
41
42
43
44 Sztuba-Solińska J, Stollar V, Bujarski JJ. 2011. Subgenomic messenger RNAs: Mastering
45 regulation of (+)-strand RNA virus life cycle. *Virology* 412:245–255.
46
47
48
49 Takeshita D, Tomita K. 2012. Molecular basis for RNA polymerization by Q β replicase. *Nat Struct
50 Mol Biol* 19:229–237.
51
52
53
54
55
56
57
58
59
60

1
2
3
4 Turina M, Hillman BI, Izadpanah K, Rastgou M, Rosa C, ICTV Report Consortium. 2017. ICTV
5
6 virus taxonomy profile: *Ourmiavirus*. J Gen Virol 98:129–130.
7
8

9
10 te Velthuis AJW. 2014. Common and unique features of viral RNA-dependent polymerases. Cell
11
12 Mol Life Sci 71:4403–4420.
13
14

15 Viljakainen L, Holmberg I, Abril S, Jurvansuu J. 2018. Viruses of invasive Argentine ants from the
16 European Main supercolony: characterization, interactions and evolution. J Gen Virol
17 99:1129–1140.
18
19

20 Wejksnora PJ, Haber JE. 1978. Ribonucleoprotein particle appearing during sporulation in yeast. J
21
22 Bacteriol 134:246.
23
24

25 Wesolowski M, Wickner RB. 1984. Two new double-stranded RNA molecules showing non-
26 mendelian inheritance and heat inducibility in *Saccharomyces cerevisiae*. Mol Cell Biol
27 4:181.
28
29

30 Wickner RB, Fujimura T, Esteban R. 2013. Viruses and prions of *Saccharomyces cerevisiae*. Adv
31
32 Virus Res 86:1–36.
33
34

35 Wolf YI, Kazlauskas D, Iranzo J, Lucía-Sanz A, Kuhn JH, Krupovic M, Dolja VV, Koonin EV.
36
37 2018. Origins and evolution of the global RNA virome. MBio 9:e02329–18.
38
39

40 Wu J, Liu W, Gong P. 2015. A structural overview of RNA-dependent RNA polymerases from the
41
42 *Flaviviridae* family. Int J Mol Sci 16:12943–12957.
43
44

45 Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN,
46
47 Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred
48 server at its core. J Mol Biol 430:2237–2243.
49
50
51
52
53
54
55
56
57
58
59
60

Zinder ND. 1980. Portraits of viruses: RNA phage. Intervirology 13:257–270.

For Review Only

SUPPLEMENTARY MATERIAL**A case for a negative-strand coding sequence in a group of positive-sense RNA****viruses**

Adam M. Dinan, Nina I. Lukhovitskaya, Ingrida Olendraite and Andrew E. Firth

Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road,
Cambridge, CB2 1QP, UK

1
2 **Supplementary Table S1.** Full list of sequences included in this study. Taxonomic information is
3 taken from NCBI GenBank, where applicable.
4

Organism	Database	Accession	Taxonomy of source
Barns Ness breadcrumb sponge narna-like virus 4	GenBank	MF190030.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Beihai barnacle virus 10	GenBank	KX883482.1	Viruses, unclassified RNA viruses
Beihai narna-like virus 22	GenBank	NC_032404.1	Viruses, unclassified RNA viruses
Beihai narna-like virus 23	GenBank	NC_032457.1	Viruses, unclassified RNA viruses
Beihai narna-like virus 24	GenBank	KX883500.1	Viruses, unclassified RNA viruses
Beihai narna-like virus 25	GenBank	KX883471.1	Viruses, unclassified RNA viruses
Botrytis ourmia-like virus	GenBank	NC_028476.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Ourmiavirus
Cassava virus C	GenBank	NC_013111.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Ourmiavirus
Cryphonectria parasitica mitovirus 1-NB631	GenBank	NC_004046.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Epirus cherry virus	GenBank	NC_011065.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Ourmiavirus
Fusarium coeruleum mitovirus 1	GenBank	NC_026622.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Fusarium poae narnavirus 1	GenBank	LC150604.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Hubei mosquito virus 3	GenBank	NC_033257.1	Viruses, unclassified RNA viruses
Hubei narna-like virus 15	GenBank	KX883540.1	Viruses, unclassified RNA viruses
Hubei narna-like virus 16	GenBank	KX883526.1	Viruses, unclassified RNA viruses
Hubei narna-like virus 18	GenBank	KX883517.1	Viruses, unclassified RNA viruses
Hubei narna-like virus 19	GenBank	KX883542.1	Viruses, unclassified RNA viruses
Hubei narna-like virus 20	GenBank	KX883548.1	Viruses, unclassified RNA viruses
Hubei narna-like virus 21	GenBank	KX883539.1	Viruses, unclassified RNA viruses
Leptomonas Narna-like virus 1	GenBank	KY628364.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Leptomonas seymouri Narna-like virus 1	GenBank	KU935604.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Leptomonas seymouri RNA virus	GenBank	KX373304.1	Viruses, unclassified RNA viruses
Linepithema humile narna-like virus 1	GenBank	MH213236.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Narnaviridae environmental sample	GenBank	KP642119.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Narnaviridae environmental sample	GenBank	KP642120.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Ochlerotatus-associated narna-like virus 1	GenBank	KF298275.2	Viruses, environmental samples
Ochlerotatus-associated narna-like virus 2	GenBank	KF298276.2	Viruses, environmental samples
Ochlerotatus-associated narna-like virus 2	GenBank	KF298284.2	Viruses, environmental samples
Ophiostoma mitovirus 3a	GenBank	NC_004049.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Ophiostoma mitovirus 4	GenBank	NC_004052.1	Viruses, ssRNA viruses, ssRNA positive-strand

			viruses, no DNA stage, Narnaviridae
Ophiostoma mitovirus 6	GenBank	NC_004054.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Ourmia melon virus	GenBank	NC_011068.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Ourmiavirus
Phytophthora infestans RNA virus 4	GenBank	JN400241.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Phytophthora infestans RNA virus 4	GenBank	JN400242.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Phytophthora infestans RNA virus 4	GenBank	JN400243.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Phytophthora infestans RNA virus 4	GenBank	KU295722.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Phytophthora infestans RNA virus 4	GenBank	KU295726.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Phytophthora infestans RNA virus 4	GenBank	KU295727.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Point-Douro narna-like virus	GenBank	MF176258.1	Viruses, unclassified viruses
Saccharomyces 20S RNA narnavirus	GenBank	AF039063.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Saccharomyces 23S RNA narnavirus	GenBank	U90136.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Saccharomyces cerevisiae	GenBank	M64034.1	Eukaryota, Fungi, Dikarya, Ascomycota
Sanxia water strider virus 13	GenBank	KX883566.1	Viruses, unclassified RNA viruses
Shahe narna-like virus 4	GenBank	KX883556.1	Viruses, unclassified RNA viruses
Shahe narna-like virus 4	GenBank	KX883562.1	Viruses, unclassified RNA viruses
Soybean leaf-associated ourmiavirus 1	GenBank	KT598235.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Ourmiavirus
Tuber aestivum mitovirus	GenBank	NC_015629.1	Viruses, ssRNA viruses, ssRNA positive-strand viruses, no DNA stage, Narnaviridae
Wenling narna-like virus 7	GenBank	KX883602.1	Viruses, unclassified RNA viruses
Wenling narna-like virus 8	GenBank	KX883605.1	Viruses, unclassified RNA viruses
Wilkie narna-like virus 1	GenBank	NC_035126.1	Viruses, unclassified viruses
Wilkie narna-like virus 2	GenBank	NC_035120.1	Viruses, unclassified viruses
Wuchang cockroach Virus 4	GenBank	KX883626.1	Viruses, unclassified RNA viruses
Wuhan horsefly Virus 3	GenBank	NC_033474.1	Viruses, unclassified RNA viruses
Wuhan insect virus 18	GenBank	KX883516.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	KX883461.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	KX883537.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	KX883538.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	MF176257.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	MF176278.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	MF176306.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	MF176344.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	MF176365.1	Viruses, unclassified RNA viruses
Zhejiang mosquito virus 3	GenBank	MF176385.1	Viruses, unclassified RNA viruses
Acartia tonsa	TSA	HAGX01098760.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda

1	Agarophyton chilense	TSA	GEZJ01006036.1	Eukaryota, Rhodophyta, Florideophyceae, Rhodymeniophycidae
2	Agarum clathratum	TSA	GEWO01002579.1	Eukaryota, Stramenopiles, PX clade, Phaeophyceae
3	Agarum clathratum	TSA	IABI01004424.1	Eukaryota, Stramenopiles, PX clade, Phaeophyceae
4	Agrostis stolonifera	TSA	GFQK01014503.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
5	Bactrocera dorsalis	TSA	GEEA01024908.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
6	Bactrocera dorsalis	TSA	GGBS01024784.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
7	Brassica rapa subsp. chinensis	TSA	GFUS01019776.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
8	Callosobruchus maculatus	TSA	GEUE01035234.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
9	Callosobruchus maculatus	TSA	GEUE01057748.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
10	Caridina multidentata	TSA	IABX01132835.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
11	Ceratina chalybea	TSA	GBPU01010966.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
12	Ceratitis capitata	TSA	GAMC01013668.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
13	Colobanthus quitensis	TSA	GCIB01004342.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
14	Colpodella angusta	TSA	GDKJ01026489.1	Eukaryota, Alveolata, Colpodellidae, Colpodella
15	Cronartium ribicola	TSA	GBSG01012161.1	Eukaryota, Fungi, Dikarya, Basidiomycota
16	Cronartium ribicola	TSA	GBSG01013692.1	Eukaryota, Fungi, Dikarya, Basidiomycota
17	Crossocerus quadrimaculatus	TSA	GBWH01015246.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
18	Dendroctonus ponderosae	TSA	GDAR01017830.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
19	Entomophthora muscae	TSA	GENB01008321.1	Eukaryota, Fungi, Zoopagomycota, Entomophthoromycotina
20	Entomophthora muscae	TSA	GENC01006608.1	Eukaryota, Fungi, Zoopagomycota, Entomophthoromycotina
21	Entomophthora muscae	TSA	GENC01041260.1	Eukaryota, Fungi, Zoopagomycota, Entomophthoromycotina
22	Entomophthora muscae	TSA	GEND01007370.1	Eukaryota, Fungi, Zoopagomycota, Entomophthoromycotina
23	Entomophthora muscae	TSA	GEND01011317.1	Eukaryota, Fungi, Zoopagomycota, Entomophthoromycotina
24	Eurypanopeus depressus	TSA	GFJG01044014.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
25	Eurypanopeus depressus	TSA	GFJG01122116.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
26	Fopius arisanus	TSA	GBYB01012090.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
27	Gnathonemus petersii	TSA	GFIK01043228.1	Eukaryota, Metazoa, Chordata, Craniata
28	Helicoverpa assulta	TSA	GBTA01004534.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
29	Heterosiphonia pulchra	TSA	GFLH01017038.1	Eukaryota, Rhodophyta, Florideophyceae, Rhodymeniophycidae
30	Hordeum vulgare	TSA	GGCO01034162.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
31	Hordeum vulgare	TSA	GGCO01105932.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
32	Humulus lupulus var. lupulus	TSA	LA534138.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
33	Humulus lupulus var. lupulus	TSA	LA710658.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
34	Idiosepius notoides	TSA	GFNE01046070.1	Eukaryota, Metazoa, Lophotrochozoa, Mollusca
35	insect metagenome	TSA	GCRV01019573.1	unclassified sequences, metagenomes, organismal metagenomes
36	Juncus effusus	TSA	GFBP01062652.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta

1	Lasioglossum xanthopodus	TSA	GBPT01016202.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
2	Lasioglossum xanthopodus	TSA	GBPT01024369.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
3	Lates calcarifer	TSA	GAMU01000156.1	Eukaryota, Metazoa, Chordata, Craniata
4	Leucoagaricus gongylophorus	TSA	HAAN01020927.1	Eukaryota, Fungi, Dikarya, Basidiomycota
5	Leucocoprinus sp. HH-2015a	TSA	GEHH01003501.1	Eukaryota, Fungi, Dikarya, Basidiomycota
6	Locusta migratoria	TSA	GBFE01004865.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
7	Locusta migratoria manilensis	TSA	GBDZ01000351.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
8	Loxostege sticticalis	TSA	GFCJ01029165.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
9	Murgantia histrionica	TSA	GECQ01195387.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
10	Myxobolus cerebralis	TSA	GBKL01051391.1	Eukaryota, Metazoa, Cnidaria, Myxozoa
11	Myxobolus cerebralis	TSA	GBKL01052498.1	Eukaryota, Metazoa, Cnidaria, Myxozoa
12	Nephila clavipes	TSA	GFKT011160020.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
13	Nereocystis luetkeana	TSA	GEWH01000774.1	Eukaryota, Stramenopiles, PX clade, Phaeophyceae
14	Nereocystis luetkeana	TSA	GEWH01001802.1	Eukaryota, Stramenopiles, PX clade, Phaeophyceae
15	Nereocystis luetkeana	TSA	GEWH01003494.1	Eukaryota, Stramenopiles, PX clade, Phaeophyceae
16	Panax notoginseng	TSA	GFRX01320111.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
17	Pergagrapta polita	TSA	GBWZ01004682.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
18	Phakopsora pachyrhizi Thai1	TSA	GACM01002912.1	Eukaryota, Fungi, Dikarya, Basidiomycota
19	Pleurophytus gardneri	TSA	GEWL01001678.1	Eukaryota, Stramenopiles, PX clade, Phaeophyceae
20	Proasellus karamani	TSA	HAFC01092190.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
21	Prosopis cineraria	TSA	GFWO01018115.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
22	Puccinia striiformis f. sp. tritici	TSA	GAIR01012025.1	Eukaryota, Fungi, Dikarya, Basidiomycota
23	Puccinia striiformis f. sp. tritici	TSA	GAIR01012062.1	Eukaryota, Fungi, Dikarya, Basidiomycota
24	Puccinia striiformis f. sp. tritici	TSA	GAIS01005902.1	Eukaryota, Fungi, Dikarya, Basidiomycota
25	Pyropia haitanensis	TSA	GADD01004427.1	Eukaryota, Rhodophyta, Bangiophyceae, Bangiales
26	Pyropia haitanensis	TSA	GFOL01000028.1	Eukaryota, Rhodophyta, Bangiophyceae, Bangiales
27	Pyropia haitanensis	TSA	GFOL01000185.1	Eukaryota, Rhodophyta, Bangiophyceae, Bangiales
28	Pyropia haitanensis	TSA	GFOL01002916.1	Eukaryota, Rhodophyta, Bangiophyceae, Bangiales
29	Pyropia haitanensis	TSA	GFOL01002977.1	Eukaryota, Rhodophyta, Bangiophyceae, Bangiales
30	Rhipidia juga	TSA	GEMJ01010140.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
31	Rhizopus oryzae	TSA	GDUK01007882.1	Eukaryota, Fungi, Fungi incertae sedis, Mucromycotina
32	Rhizopus oryzae	TSA	GDUK01014407.1	Eukaryota, Fungi, Fungi incertae sedis, Mucromycotina
33	Saccharum hybrid cultivar	TSA	GFHZ01053830.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
34	Saccharum hybrid cultivar	TSA	GFLP01304447.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
35	Saccharum hybrid cultivar	TSA	GFLP01591397.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
36	Saccostrea glomerata	TSA	GGIC01436417.1	Eukaryota, Metazoa, Lophotrochozoa, Mollusca
37	Tomicus yunnanensis	TSA	GFJU01140648.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda

1	Tracheliastes polycolpus	TSA	GGQW01011558.1	Eukaryota, Metazoa, Ecdysozoa, Arthropoda
2	Uromyces appendiculatus	TSA	GACI01002802.1	Eukaryota, Fungi, Dikarya, Basidiomycota
3	Zostera noltei	TSA	HACV01002003.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta
4	Zostera noltei	TSA	HACV01003050.1	Eukaryota, Viridiplantae, Streptophyta, Embryophyta

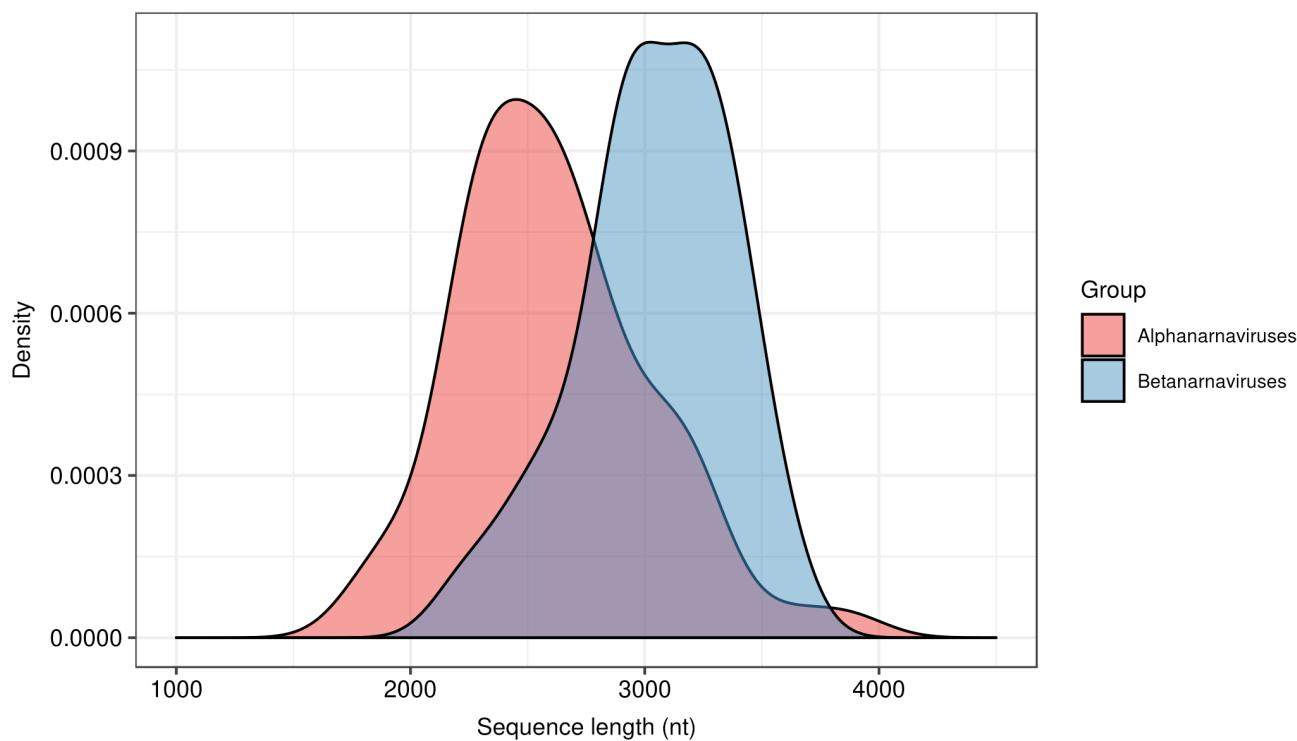


Figure S1. Gaussian kernel density estimates of the sequence length distributions for alphanarnaviruses ($n = 64$) and betanarnaviruses ($n = 33$) included in the analysis.

		motif G		motif F		
2	GBSG01013692.1	SCFERSTTOGG-----VDGHLR	//	HCLRTPGYKCRVIGPDAFLTVFEGTWVRWSSQLPR		Cronr
3	MH213236.1	STFMTMSSEG-----RLKEMM	//	IGVPEYGWKRTRLFPNVYALTPGDLIR--QQLWPL		LhuNLV1
4	GFJU01140648.1	ATLNPHASOGG-----RMAELI	//	LVVPERGYKARVLVKFPASALLVGDIVR--RQLWPQ		Tomyu
5	KX883548.1	ASLFNKASEGG-----RLAELI	//	IAVSEKGYKARVLVFPASTLLPGDIIR--RQLWPM		HNLV20
6	KX883542.1	ASQGATILEGG-----RMEELI	//	MAIAEKGYKARVLTKFPAAALVVGDVIR--RQMAM		HNLV19
7	GGQW01011558.1	ASLEFTEKAGG-----RLADVM	//	VAVAEPGNKVRVCKFRSPVLLSDIIR--RQLFPI		Trapo
8	GBFE01004865.1	ASRDYTIERGG-----RVQELL	//	MAAAEPGNKARVLCKFPAPVALVPGDIIR--RQLWPI		Lmigr
9	KX883517.1	ASRDVPASAGG-----RLRDVL	//	LAIPEPGFKARVLCKFPATALLAGDIIR--RQLWPI		HNLV18
10	IABX01132835.1	ATLETTRAOGG-----FSEETR	//	IALAERGKTRVVTAKAPEWIVLGHFLR--SWLLDG		Cmult
11	KX883500.1	ACKEVPKRAKGG-----LGTRSV	//	MALPERGGKVRIVTCKPCWPALVYLGHFLR--VWLLEG		BNLV24
12	KX883605.1	ATFERTRTEGG-----FAQSQV	//	LAIPERGLKARVITCKPCWPALVYLGHFLR--SWLLQG		WNLV8
13	GFLP01591397.1	AVGERSKARGG-----YNAHYI	//	TGIGEQGDKCRRIITVPPASLFAAGDVCR--SRIWPR		Sacch
14	GGC001034162.1	ACYENPRGAGGYAYIKKLGD	//	AVIPERGYKRNRTVAPPASILSMGEVVR--SSIFPY		Horvu
15	GAC101002802.1	ACYE6PGRSRGG-----YLGHIK	//	STIPERGYKRNRTVAPPGSVLSGEVIR--SIIFFP		Uroma
16	GGC001105932.1	SCFEGPRSKGG-----YFGYIR	//	GVVPERGYKRNRTVSPPSGVLSAGEVVR--HVLFPP		Horvu
17	GENC01041260.1	--	//	VALRERGFKARIWTKSPVELVECGHLLR--SLVWPM		Enmus
18	LA534138.1	ATLDYSRRLKGG-----MRSDLK	//	TTVKERGYKCRVVTKSPADVVEVGHLVR--SVWPM		Humlu
19	GEUE01057748.1	ASVEFSRLKGG-----QTAELF	//	TTVPETGGKARVVTAGDPMLVLAGNALR--RAWPPI		Calma
20	GBYB01012090.1	ASEGMSRAKGG-----QRAELE	//	TALPELGKARVVTAPPAGHGIIDAMR--KVLWPL		Fopar
21	KX883539.1	AAESVSRARGG-----QREELR	//	AVVPELGAKSRIVTAGEAHWWVIGDAIR--KCLWPA		HNLV21
22	GENC010066608.1	AAEVSRSRADGG-----QKADLL	//	SAIPEYGSKVRVTCAPGPVVAGDACK--KIWGL		Enmus
23	KF298275.1	AAAEFAASAGG-----QQAELR	//	TIVREQGMKARVVTACPWAHAVVCGDACK--KTLWPL		ONLV1
24	KP642119.1	ASASVSAADDG-----QLAELR	//	TTVLELGKARVVTKPPAWAVVAGDACK--QSVWPL		CNLV1
25	MF176385.1	ASATVSAKLKGG-----QLTELR	//	TVISELGMKARVVTKPPAWAVVAGDACK--KTVWPL		ZJMV3
26	KF298284.1	ASATVAMSAGG-----QLAELS	//	TVISELGMKARVVTKPPAWAVVAGNACR--KTVWPL		ONLV2
27			//	*	:	:
28				***	***	:
29						:
30						:
31						:
32						:
33						:
34						:
35						:
36						:
37						:
38						:
39						:
40						:
41						:
42						:
43						:
44						:
45						:
46						:
47						:
48						:
49						:
50						:
51						:
52						:
53						:
54						:
55						:
56						:
57						:
58						:
59						:
60						:
61						:
62						:
63						:
64						:
65						:
66						:
67						:
68						:
69						:
70						:
71						:
72						:
73						:
74						:
75						:
76						:
77						:
78						:
79						:
80						:
81						:
82						:
83						:
84						:
85						:
86						:
87						:
88						:
89						:
90						:
91						:
92						:
93						:
94						:
95						:
96						:
97						:
98						:
99						:
100						:
101						:
102						:
103						:
104						:
105						:
106						:
107						:
108						:
109						:
110						:
111						:
112						:
113						:
114						:
115						:
116						:
117						:
118						:
119						:
120						:
121						:
122						:
123						:
124						:
125						:
126						:
127						:
128						:
129						:
130						:
131						:
132						:
133						:
134						:
135						:
136						:
137						:
138						:
139						:
140						:
141						:
142						:
143						:
144						:
145						:
146						:
147						:
148						:
149						:
150						:
151						:
152						:
153						:
154						:
155						:
156						:
157						:
158						:
159						:
160						:
161						:
162						:
163						:
164						:
165						:
166						:
167						:
168						:
169						:
170						:
171						:
172						:
173						:
174						:
175						:
176						:
177						:
178						:
179						:
180						:
181						:
182						:
183						:
184						:
185						:
186						:
187						:
188						:
189						:
190						:
191						:
192						:
193						:
194						:
195						:
196						:
197						:
198						:
199						:
200						:
201						:
202						:
203						:
204						:
205						:
206						:
207						:
208						:
209						:
210						:
211						:
212						:
213						:
214						:
215						:
216						:
217						:
218						:
219						:
220						:
221						:
222						:
223						:
224						:
225						:
226						:
227						:
228						:
229						:
230						:
231						:
232						:
233						:
234						:
235						:
236						:
237						:
238						:
239						:
240						:
241						:
242						:
243						:
244						:
245						:
246						:
247						:
248						:
249						:
250						:
251						:
252						:
253						:
254						:
255						:
256						:
257						:
258						:
259						:
260						:
261						:
262						:
263						:
264						:
265						:
266						:
267						:
268						:
269						:
270						:
271						:
272						:
273						:
274						:
275						:
276						:
277						:
278						:
279						:
280						:
281						:
282						:
283						:
284						:
285						:
286						:
287						:
288						:
289						:
290						:
291						:
292						:
293						:
294						:
295						:
296						:
297						:
298						:
299						:
300						:
301						:
302						:
303						:
304						:
305						:
306						

Figure S2. Alignment of RdRp protein-coding sequences of rORF-containing alphanarnaviruses. Known RdRp functional motifs are indicated and key functional residues previously described are highlighted. Key: Cronr = *Cronartium ribicola* TSA; LhuNLV1 = Linepithema humile narna-like virus 1; Tomyu = *Tomicus yunnanensis* TSA; HNLV = Hubei narna-like virus; Trapo = *Tracheliastes polycolpus* TSA; Lmigr = *Locusta migratoria* TSA; BNLV = Beihai narna-like virus; Cmult = *Caridina multidentata* TSA; WNLV = Wenling narna-like virus; Sacch = *Saccharum* hybrid TSA; Horvu = *Hordeum vulgare* TSA; Uroma = *Uromyces appendiculatus* TSA; Enmus = *Entomophthora muscae* TSA; Humlu = *Humulus lupulus* TSA; Calma = *Callosobruchus maculatus* TSA; Fopar = *Fopius arisanus* TSA; ONLV = Ochlerotatus-associated narna-like virus; CNLV = Culex-associated narna-like virus; ZJMV3 = Zhejiang mosquito virus 3.

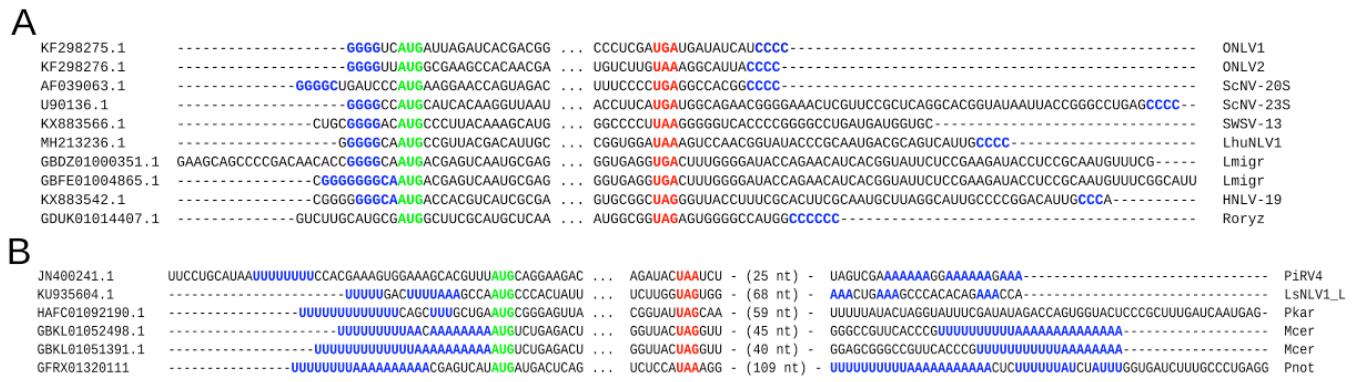


Figure S3. Alignments of **(A)** alphanarnaviral and **(B)** betanarnaviral genomic termini. The start and stop codons of the RdRp ORFs are shown in green and red, respectively. Complementary G/C- and A/U-rich regions are indicated in blue. Key: ONLV = Ochlerotatus-associated narna-like virus; ScNV-20S = *Saccharomyces cerevisiae* 20S narnavirus; ScNV-23S = *Saccharomyces cerevisiae* 23S narnavirus; SWSV-13 = Sanxia water strider virus strain 13; LhuNLV1 = *Linepithema humile* narna-like virus 1; Lmigr = *Locusta migratoria* TSA; HNLV-19 = Hubei narna-like virus 19; Roryz = *Rhizopus oryzae* TSA; PiRV4 = Phytophthora infestans RNA virus 4; LsNLV1_L = *Leptomonas seymouri* narna-like virus 1 segment L; Pkar = *Proasellus karamani* TSA; Mcer = *Myxobolus cerebralis* TSA; Pnot = *Panax notoginseng* TSA.

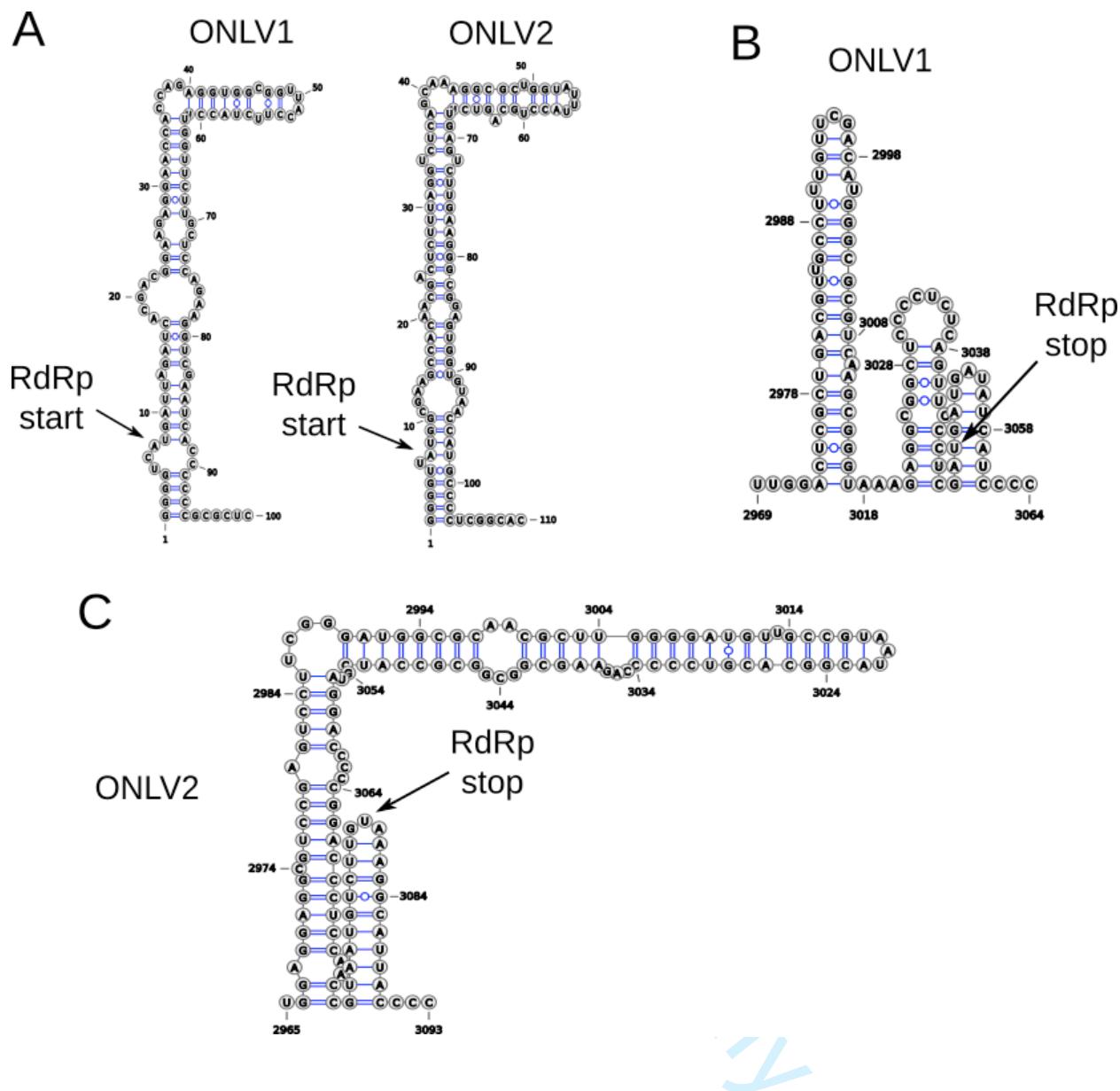


Figure S4. Predicted RNA secondary structures at the terminal regions of alphanarnaviral genomes. **(A)** 5' termini of *Ochlerotatus*-associated narna-like virus (ONLV) 1 and ONLV2; **(B)** 3' terminus of ONLV1; and **(C)** 3' terminus of ONLV2.

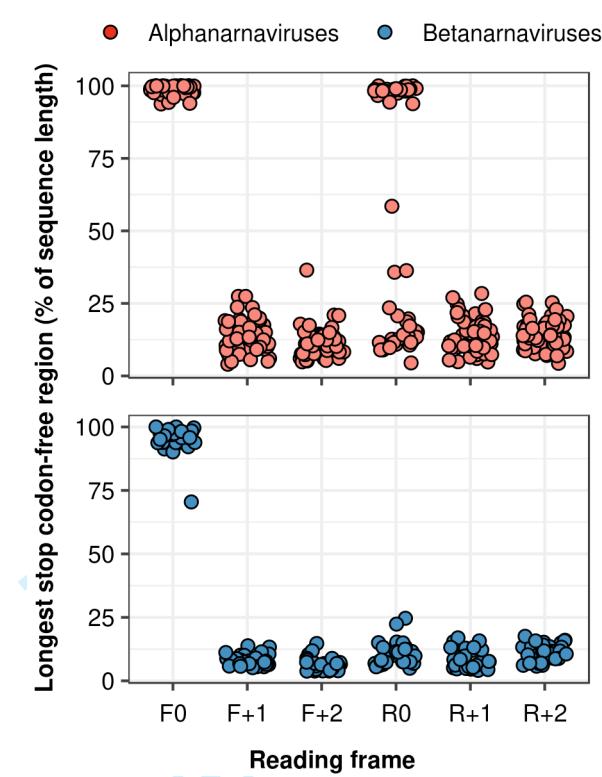


Figure S5. The longest stop codon-free regions in each of the three possible positive-strand and negative-strand reading frames (without the 5'-proximity restriction), for alphanarnaviruses (red) and betanarnaviruses (blue), as a percentage of the sequence length.

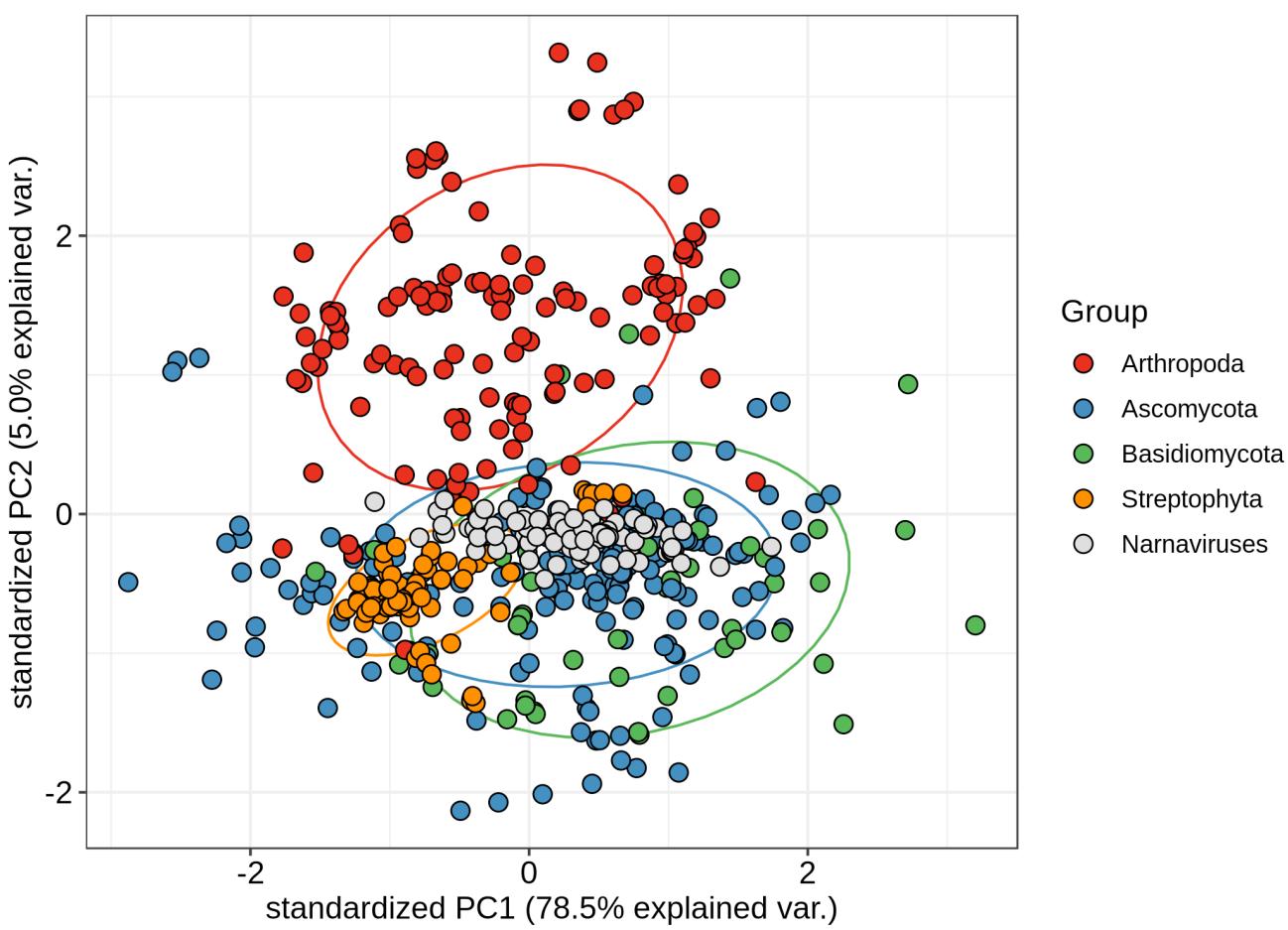


Figure S6. Principal component analysis of codon usage (per associated amino acid) in phyla that frequently co-occur with narnaviruses. Codon usage across NCBI RefSeq genomes was extracted from the latest release of the codon usage table database (CUTD). Ellipse contours are drawn for each group at a normal probability of 0.68.

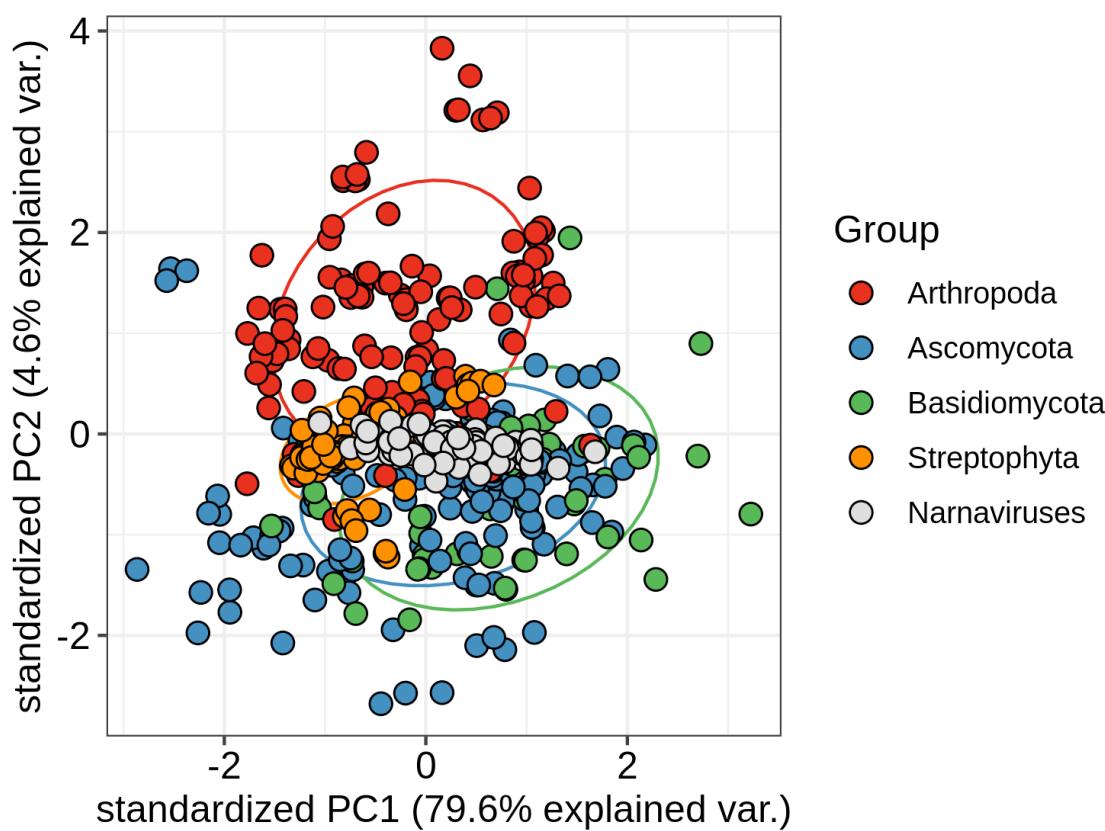


Figure S7. Principal component analysis of codon usage (per associated amino acid) in phyla that frequently co-occur with narnaviruses, excluding leucine and serine codons. Codon usage across NCBI RefSeq genomes was extracted from the latest release of the codon usage table database (CUTD). Ellipse contours are drawn for each group at a normal probability of 0.68.

1
2
3 >NC_004046.1
4 -----
5 MAMIIHDPVYKLWYA-----XRAKSNLPG
6 LAPH-----KTESITTVLSLTNVKNKQAT
7 -----NSILLNKITDLVLGLKAFAFSIKS
8 KG-----VRTT-----LRVNKHWTIPKEFPRFVKLVVWC-----TR
9 TQEHEDSFMKIIG-KCDHIWQT-----AGPNFLFKYLK-----EVMRLSVRRIA---
10 -----NIELEPSKKIFVKLNK-----FRFPNIPLPICDQIIRDQ
11 -----NDQVLWASKR-----LIICLLTILSVHRV-LPTKVVPDYSTIVDP----
12 -----FTGVSKTIDQK--LLRKAIHLLNIKRVKQL-----KLKIT--GSMKAG
13 PN--GKIS--LLT-----SSVDALSFITQPTKIFTYLDFSVRVYKFRGLLLWMWM--
14 -----M
15 CILL-----ITLPYAI--VSFML
16 GALI-----
17 -----PIMGKLSV--VYDQA-GKARIVAITNSWIQTA-FYSLH-----LHVFK-LL
18 KNIDQ-----DGTFDQERPFKLLI-----KWLNEPTQ-----
19 -----KF-YGF-----DLTAATDRLPI--DLQVDILNIIFKN-----
20 -SPGS-SWRSSLRIKYKS--PQGF-LTY-----
21 -----AVGQPMGAYSSFAMLALTHVI---V---
22 QVA-----ALNSGFTTRF-----
23 -----TDY--CILGDD--IVIA--HDTVA--SEYLKLMET-LGLSISSGKSV
24 IS--SEF-TEFAKKL-----KGRNNF--DIFYR
25 SWF-----
26 SIIHFEKQ-----
27 -----ILHLCTV-----FELLRRGVCELY-----DLY
28 PQYINKL-----PKIYLR-----YNLLID-----W
29 VVVAFTNQILIGDRPRA--DGIR--L-----FDYFV-----
30 -----GLEVIP-PL-----LRIMLHTIKKDWNGLWNSIKYTLNKGFFVVSQVRVGLPD-
31 -----WTEFL-----LPILPSTYIIRDY-----
32 -----CRSFNDLTKLFGEWWL
33 LRF-----ESESYQVSILD-----VIDRLAH
34 TSIPNLDIHDKKKVKLTLD---NLYKLSLIVNIPS-----GG
35 ARRY-----IEFLRFNGLKSPLIVERYIKDGI-----
36 RIEKPLTL-----
37 -----QGLHRSGDIQLGFKIS-----
38 -----
39 >NC_004054.1
40 -----
41 -----
42 -----MKLKQLKLMKN-----
43 -----
44 -----
45 -----
46 -----
47 -----
48 -----KTYQ-----IIRILLIVFFPSI-----KR
49 QTVLNIFMSKIN-KMIKN-----NGTLFTVKYLK-----ELRLHITKYISGEP
50 YRN-----SLNRVSVDKDGFP-----TLCKELKVLVNGTY--LEK
51 ---RFVLTIITLSKLLIP--QKSESIPFSTKSITDHWSG-IDNISNEE---LD---
52 -----KSCSELNISTREVQWDVK-----NF-----KLLTKA
53
54
55
56
57
58
59
60

1
2
3 GPHG-----PQSLTWYHTIKLYDFNQWLGII-----
4 -----G
5 ILPK-----SVLDLFT--ETLSY
6 ASKL-----VLPEIKS
7 NIKSTKLIRRLSI--VHDPE-CKERVIAIFDYGSQMV-LKPIA-----DVLFD-LL
8 RNIPS-----DRTFTQSP-FFTHT-----DLDNKS-----
9 -----KF-WSI-----DLSSATDRFPI--VFQKRVLQKILGK-----
10 -----QMTDSWERIMIG---SKFL--A-----
11 -----PDGD-----TVSYNCQPMGAQSSWPMFTLAHHVI----V---
12 RVA-----ANRCGLSNF-----
13 -----DKY--IILGDD--IVIN--NDNVA--LKymeimnd-FKVEISRNKTH
14 VS--NDT-YEFAKRW---I-----K---NKMEFF--PLPIR
15 -----
16 GIVDNINN-----
17 -----KYIIFNILYSFFVEKGNTFLNKDTLL----VCV
18 SKFIQLH--SLTLKKPIG-----LNKVKG-----ILYPFNFMRLR
19 Y-RQNLCCTNEE--IRIFL--GSST---C-----KRDDYML-----
20 -----PISAKDVSLELT---RVISAALVG----MAYNAEKLKNIYFDLDKLS--P-
21 -----WIGDGF-----KTGKHPK-----
22 -----VM-IQSIYNSVKSLSDFGL
23 KMA-----QNKLTLSAAMD-----
24 -----SLLLV--DLDSISSLERIKY-----
25 -----IQMKQNICLSQKVRKELR-----FD
26 PLQMEQKA-----RA-----
27 -----MMLVK--HMKDLEG-----
28 -----
29 >NC_004052.1
30 -----
31 -----
32 -----
33 -----
34 -----
35 -----
36 -----MKRNNNLQIIIKR-----
37 -----
38 -----
39 -----LILH-----IFKINLSVEIDKF-----
40 -----LGFINHLRKS-----NGLLYTIKYM-----AAKLHITRYMCGKP
41 LYS-----NNENVALDKTGFP-LRF-----WYLKRLVNDNPRAL-----
42 -----LTLLTYTRRI--VP-----NKSESKARIVKLST-ITDPYKGK-----
43 -----VYTIPKWFIELD--FISK--YNLSSTKPI-----YTDND--HYLSIK
44 GSPN-GKASMSSLYS-----IISFNSSNIRYLFNIVGDYQLVLNKFYQ-----
45 -----
46 -----
47 -----DLSQFYT--KYINR
48 DKLG-----
49 -----LGKLSI--VHDPE-LKERVIAVDYTTQFA-LRPIH-----NILLN-NL
50 SKLPCD-----RTFTQDPF-HKWND-----DHKERY-----
51 -----HSL-----DLSSATDRFPI--FLQQKLISLIFND-YE-----
52 -----FGKNWRNL-LV---DRNY--D-Y-----
53 -----QGIS-----YRYSVGQPMGAYTSWAFTLTHHLV----V---
54 HWA-----AELAGLKNF-----
55 -----
56
57
58
59
60

1
2
3 -----KDY--IILGDD--IVIK---NNKVA--QIYINLMTK-WGVDISLSKTH
4 VS--YDT-YEFAKRW-----IKNGKE--ISGIS
5 LK-----
6 GLTNIRH-----
7 -----IHVVYM-----NIFTYLQRIPSLNVNDILTC--VGKLYGY-----LLI
8 RNRIKSP---NTIKRSLY-----DFHHSI-----RYSFGL-----LNYEEIRNY
9 L-HNKFPFDNYYAWPERLVHSKLN--EI-----FKLEMVE-----
10 ---SAKSFSKDFMN----QSTMLINTVTDNEIMVQWPLYK---GFMNHIEKLK--D-
11 ----YIKSKQ----NQHDIDLLDLMQNL-----RFQNLDSIVKK-----
12 -----LRNSYTNLIMLDKFWK
13 SA-----FNREYRDLERESILTIEKQESNMMMSRIWDMALSYR
14 TSPMSYSTLTTFETDSEFYM---MPSIWDMASSSTTGPKP-----FT
15 TATF-----KTDFVGLSTFDDKLLKDLENLKID-----IT
16 LRTGKYTN-----KTQPLE
17 S-----KIHQH--PIECIEDNKVPNFNNIK-----
18 -----
19 >NC_026622.1
20 -----
21 -----
22 -----
23 -----
24 -----MP-----LRDSHWSFN
25 SISN-----YSVRLPPDKWMRSLIKNYNS
26 -----
27 -----IMKN-----IKLHLLLVKRLITCIFNI-----DH
28 TIIITLFIKEIYRLWDK-----NGFTFMIKYMK-----SVRLHITRYICKKP
29 IHS-----NSSNVSVDKSGFP-SKF---LYLKPLIKEVKDIKI----
30 ---LLTLLSFTRG-----LKP--SKKEDSNIKYDLSS-ITSP-----
31 -----FKGTSKGSPVQ--WFIK--SFVETNKLYKKIP-----EYSVKDH----YL
32 STKG-GPCG-----KSTWSSTKSHLFYKQDLILNIQNIFREGFKELFFT---
33 -----P
34 FLKN-----MHLSY
35 GVTR-----
36 -----
37 -----WPNGKLSI--VKDPE-GKRRVIAMVDYHSQLA-LKSIH-----NDLLD-LL
38 SKFKCD-----RTFTQDPL-HNWSN-----NKE-----
39 -----SY-FSL-----DLSSATDRFPV--ELQKRLLAEIYQD-DK-----
40 -FATN--WMELL---LN---RDYI--G-----
41 -PGGE-----ICRYSVGQPMGAYSSWAFTLTHHLV----V---
42 SWS-----AYKALKTKNF-----
43 -----DQY--IILGDD--IVIK---NNEIA--NIYRGQMMR-MGVDISMPKTH
44 IS--KDT-YEFAKRW-----IKDGKE--ISGIP
45 LK-----
46 GLINNIKH-----
47 -----LKIVYTIINDYLIKVP-----SN-----VSL
48 SSWLIFE-K-----IFVGFP-----IYSKNT-----SVKRYISRKY
49 I-RSLKDFALS--VRFSM-----GLTT-----
50 -----PYELRQ----YLSQMWIODESK---TEFKAIKPENIIRQYFEGILVNG-
51 -----LADQAK----KTILQINDQLYSF-----EKLEQKERKSL-----
52 -----VY-SGVIFGLMNRL-----
53
54
55
56
57
58
59
60

1
2
3 -----TKLEEQCTKWK-----ED
4 ESTIVDILNFFSTPSVDSL--SRKDRDCMIRLDF-----L
5 DSLW-----KKSFKKHFSEQRFPDSYYRDLE-----TK
6 TLYGSLDS-----LDITSSGIEVR-
7 -----PIWNT--ELKRFNSRSIGALEMFISDNE-----
8 -----
9
10 >GFQK01014503.1
11 -----
12 -----PLDVISLLRSK-----YVVIDKMVS
13 LSDK-----IRTTLHVPFAINNMLAHGRV
14 -----GRLARGIRRTNDF
15 VS-----YVLK-----LWKAHGTTFTIKWLKASHVAIAK-----CL
16 GQNKLSSLRTLEQDLNLPRLV-----NGLPKIIPFEDRRRMQGHVPTIRFWLGLFN
17 LY-----RILKAPAVPKLAT-ITS-----PWTGQFKP-----
18 ---FLELCRIARR---P---DFNFFCFLKGYRHIDLQHVQPTE-----
19 ---LVSSSSASPSNKISWQGFLTDLHLLREHN-----PVVWS--AIKRYI
20 HLTG-GHTF-LSYLEHALEIDDIFSKYDFYRDSSSEGHHLSTGKKIL-----
21 -----
22 -----VADHSMT--KAAIA
23 QHGL-----
24 -----PTGDGLSQFAIKEAAGKVRVFALVDLITQST-LKPLH-----DQLFS-LL
25 RKIPC-----DGTFNQEA-ALDRA-----MFKA
26 -----NCAFSF-----DLSAATDRIPV--LLTQNILANILNS-SE-----
27 -----AARLWKII-MT---HREF--G-F-----L
28 QPTAE-----KLGIDHNVGFRYAVGQPMGALSSWAGLAITHHWI-----V---
29 QVA-----AREAYYDLEGWY-----
30 -----ENY--EVLGDD--IVIF--DSLVA--DKYLGIMSS-IGTDINITKSI
31 VSRNRPV-FEFAKRI---C-----
32 -----WGS-----
33 GIVSGISL-----
34 -----NQVRAS-----WN--VGSRLAS-----V
35 LQFARLG-----LLETGS-----SLMVAL-----LSRDAFS-K
36 G-KAIPNVKTE--SPRSQKARALG--L-----LALLG-----
37 -----ERYQKGILSL--KEVTRILINP--KDYLDEGNAIA-IP-
38 -----LQFAAQ-----LAYRAL-----VDPQKPHPV-----
39 -----PSS-VEERDRIWGFR-EPFL
40 -----KMELMEAITSE-----
41 -----KQLLISEMV-----SL
42 SFNW-----TDKL-----FYPLCYDSSKE-----RVP-----ND
43 DLPVDYLN-----
44 -----AYWEL--SE-----
45 -----
46 -----
47 >NC_015629.1
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----MMNPFIRLGSLLMGKTTK--
56
57
58
59
60

1
2
3
4 -----ELIIGLYSCVQTLHKIYL-----SQ
5 GSRGLAIRLKASSTYLMKSWS-----GDPLLDCKAFG-----
6 -----PTMSLTPRGIPIW-IPS-----AWRCQISSRNLAIF-----
7 -RVTMTLCNLRYRV--LP--YQG---KLKLSTITDDRKG-STPDLMFS-----
8 -----FIKDKFAPNITNRVPYPG--IFSWDPKPLL-----TRGPG--AQKEKN
9 TMSGFGMAL-----RAIHETGMLDHLKAYALYTGLNP-----
10 -----
11 -----QLESLKVIWEGTRK
12 HCSI-----
13 -----GPIGKLAf--KQEPM--KVRVFAMVDCITQWF-LHPLH-----KYLFS-VL
14 RTTKE-----DATFDQEKGINLVR-----LALSKKLDKSV-----
15 -----FSF-----DL SAATDRLPM--DIQMVLILNSLTPFWLAEKAVRH
16 GVKGG--LGDAWADL-LV---DRDY---YLP-----
17 -WSGYARDS-----KVRYAVGQPMGALSSWAMLALTHMI----V---
18 QFA-----AASVGVTGWF-----
19 -----KEY--MVLGDD--IVIY--NSEVA--KAYSTLMGT-LGVGISDTKSL
20 TS--KIGVFEFAKRL---M-----D--LEGPCQ--GPLA
21 EFAAARF-----NLS-----
22 ILFQSFRS-----
23 -----RTLYPKISTFMRFLGFGYKVLGSLGMRLGDMRKRSYFETVAYSPLVTD-----KSV
24 NKWSEFF--KRFSTYDLY-----PLVEAC-----VYKAYH-----LIPALT-K-F
25 D-RWN-----LWRTL-----F-----PHFDDKK-----
26 -----ILLIP-EVAG---RHLESLFTSMLDNKVAKYKADHSVYVNRLNRAEVPTND-
27 -----WEAIDT---LLSIMETSPPTS DLLQDP-----DMNELGFGLSV-----
28 -----SN-YDEDNKSCTCSLSLPY
29 TVFTKI-----PQLLDSIDDWV-----EPIIPDKH
30 PEVSADVNRPLIHKPGKVF--VKGIATGKNENN-----
31 -----VEITKMLQDKKFLAYIIIESGKITI-----YS
32 PAFQAMIR-----
33 -----ASWAK--AQSGSK-----
34 -----
35 >NC_004049.1
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----MKRLTLSQNKSQNLTNNNDLS
44 -----
45 -----NVGYITKQLFPHWIRLLV-----WS
46 LQLSPAPYKKFGS-RIAILWKA-----NGVSFTVQYLN-----ECTRIVQHFVSGHP
47 VFV-----TDVMPIGLAGG--LPT-----IIPGTLRTLLRSKD--SST
48 IRGVVLSTLAVYRIMK-MP-----CVLKLESITDPFKG-ISDTLPKS-----E-----
49 -----IINGLASLGFEIPKGRSK--HLLTLSNPII-----YLLSA--GPNHSI
50 SMMG-----IW-----KDIYAWVVSPLFPTLLSFIGRMNRGNVL-----
51 -----I
52 DLLR-----AEVSYWE--ATGVK
53 PSVS-----
54
55
56
57
58
59
60

1
2
3 -----PLDLKLGKLAIKEAAGKARVFAMADSITQSV-MAPLN-----SWVFS-KL
4 KDLPMD-----GTFNQQAPLNRLVQ-----LYQDGLLHDV-----
5 -----EF-YSY-----DLSSATDRLPM--AFQKQIISVLFGS-----
6 -----KFAKDWTLLVG---RDW---Y-----
7 -----LKDI-----PYRYSVGVQPMGALSSWAMLALSHVI----V---
8 QIA-----AMRVGKLPTNY-----
9 -----ALLGDD--IVIA--DKAVA--TSYHMIMTQILGVEINLSKSL
10 VS--NNs-FEFAKRL---V-----T---MDGEVS--AVGAK
11 -----
12 -----
13 NLLVALKS-----
14 -----RWGIS-----SVI
15 LDLYNKG--LALS-----EQDLRQ-----RFSSIP-----T
16 V-SKQFGVDKL--WLVL--GPFG--F-----IPSKDGLS-----
17 -----AFMKNRSSL-----VDMHILLSCVDE--AKFDLDK-----KT-
18 -----WEANIQ-----ETVHTLLRFG-----MLSEPAGFEVF-----
19 -----SDFTSSP-LYSFIRGQFGN-KLSA
20 LV-----QDKPVRRLIFD-----GPLLHFNFY
21 TEGWCDGLMEHLTKKIQSD---SQETVSPSNPKD-----
22 -----DKVILPLRGNIKGIFFKHVLALMAE-----RD
23 PATVMRW-----
24 -----
25 -----
26 -----
27 -----
28 >HAGX01098760.1
29 -----
30 -----
31 -----
32 APPARLPLE-----
33 -----
34 -----
35 -----
36 -----FLEGVRЛАQPKSDG-----
37 -----HLSLAASATV--EYPRSKGG-----
38 -----RAAAAYFTWAKA-----RMKELGLKAP-----
39 -----ETVA-----ETATHWLYSQEAAQEYYSECLGQ-----
40 -----
41 -----AEAWL
42 -----
43 ESGE-----
44 -----PVPHRLAV--VGELG--KARVITVAPTWFVQV-LITMQ-----KILMH-SL
45 RFIPEV-----SAGVFSRD-DAYKA-----YAEIPDVST-----
46 -----GDYHF-FMS-----DLKDATN-GPPKSGIVDTLPWLFDQ-VGA-ESP-
47 -IPQR--DLEIALSLLMS---ERKI--T-V-KN-----
48 -D-YG--IP-----REFITRRGILMGDPMTKVILTIAYYLI----V---
49 KAV--C-----ADVEREHPGCHTWF-----
50 -----RIKGDDVVILVVGVSPEASSTEKHFLRRISD-AQFRVSELDTF
51 VS--HA-AFYCEGL---F-----LAPRGL--EIEKI
52 SRARRFGH-----
53 AYADPVKG-----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
 2
 3 RYMVPF-----WKVDPGGKDPLVLP--LSKLESI----VKR
 4 LSWFVPD-----SVQWRR----TYS CAM-----TATKS
 5 F-RLERDPSLF--LPPKQ--GGIG--LL-----PDAGIRE-----
 6 ----LSTADWNQFIWSAVDLRATVL RDKIPEFSDPGVPRFVARSSNHA FSHHLRV DPR
 7 QVIPAEWEETFL-----HLGDVDERLSTR-----LAEYLGRSPNYSSDISV
 8 ASYAESVDRFSTV-----HL-GRPLIKRLKTQGSYLG
 9 SLEKVGSI PPPTDLIEAVCQYRSNEARHRRIDLRFMRSDLIGPRLVVFKPEDPEVIDLEA
 10 IKQASSPGEARTLVQAGDP---FMDPDAQILRALYLA DFKS VYPIVTND RRLCFEARRTL
 11 SSRW-----PGRDFVVARSRTPLLD SRSGLEML---SRHLRTLVEGSDSHLWW D
 12 PLRTVELF-----PSLSVT DGERFLR RTSD SDWSE SLMDITNVQY
 13 ELVKSQ LA PLIQR GDQH WEE--ILA QYETVG VVFPESSVQD LE LTPGP FQL SWR PRL REP
 14 RRLKV NH LQ KIW RL RP-----
 15 >GFJG01044014.1
 16 -----
 17 -----
 18 -----
 19 IRWKP-----PIPSIGRDG
 20 PNVLLSGDETPTQKLEIWSLT YGDVC RSMFEKR-----IPRT ALRRA GIPWT NKR PNT
 21 -----RRLANLTADQRSHFILRQ
 22 LYPRMSGAFKA-----SLKKQSLGQIKSWFHTLDGLILPLL-----LE
 23 GAEDLRLIDRVTK-HVMENCA-----QNYRAFLDRLK-----SLRKFRKWL SHV
 24 NTP-----YKGFPGECHDLAP-ILD-----LWWSILQPHQGDRR--KFI
 25 Q-VTCMCTQTRATGLVP-----ARMMQESVKKFEAT-VSKEGPTI-----
 26 ----QLDGEVLMETV-----AACRLVN PETAKL-----SGGPA--ACFEKP
 27 RSKG-GQTQ--AIL-----EICKQTINYHYEPEGSTLRRVETPTHE-----
 28 -----V
 29 NLPD-----DILDYCI--EQLRM
 30 SPES-----
 31 -----MKTRV SV--VNEPS--KARVITAASV SYQVY-MSMVA-----HMLQP-VI
 32 RSPRII-----SGLRSSR-HLWNF-----LRKTMHPQNPEWDR LAGQ-----
 33 ---GAIFALST-----DLEEATDFGNI--SVADQIWKAIVET-----VPE
 34 SVKPY--FILAHEAY-TR--ERDL--L-M-----
 35 -PDGR--WV-----KKRGWLMGD PCTKAILS AQEYIFRVATK---
 36 RWG--K-----
 37 -----VFMG--SLVGDDCIMLSR--YKEAL--EDYLQ EYTR-CGF KISEKDTF
 38 IT--TNA-MYYCEEA--S-----L--VPSDSM-VDNLTV
 39 QIKR-----GMK-----
 40 DAILYIDV-----
 41 -----PRLRLM-----VATRPDQPSRHSYSDT--GRTE LL-----GRE
 42 SNWAYRW--SEKLYL-----PFR LGT-----LLQRAL-----V
 43 P-RSVDTISPF--LPISM--GGDG--AFYADPQ-----LTRDLIE-----
 44 -----RFAWSI-GEVR----YRIRR VLSGQRNP YLRTSNPIGGDKFHIIMDMASIE-
 45 -----ELVPRE-----CRIPVDDMEHCSML-----MSFKGT FIMEPIRTAYRA
 46 LQSRYWESILIEKRI PSD-----ISIPH RVTKT-GHTEGA VPLSAQE IAD
 47 FLKLI-----EDRP KMVQEWPFLVDMR KAYIMDY MHL GWKRER DP
 48 TPQEEWDLYLSGQIDL MGA--SRALYN FITRGCTDS GEPIPKVIR DR LNFFSTD PV
 49 VKLE-----VRRMELKPE SHLILVSRD YRLA FE-----IQ
 50 KMKPETYI-----TLVDPIIHLLGFNPW
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 TDDGDKTRVIEDSGAILWAD--
4 LNLFKDGYPVNRRVELVYSACMNCEPTRYGYSRADWR
5 RLHLNSRIGTYFAYRHGGWILQKVRERRPSHEAE-----
6 >GBKL01052498.1
7 -----
8 -----QKKMSETIRWKP-----PLPSIGRNG
9 PTVVVSKASTLTEKLKVYTLETGICVSFFEKR-----IPRSQLVRLNISWRETEPES
10 -----RRLSSLSPDQKTAVILSNIWPKM
11 SP-----HFIT-----LVRAAPVTAIRNWFQTLDGLLIPAIL-----FG
12 LDNEYGLIDRITK-SCLERMA-----QNYSLFLREIK-----EVRKRFRLFFFESD
13 QRF-----GTLRYPGPSYGLTSILE-----FWRRILTPALIGNQREF---
14 ---LRAVVTIGQT--RATGLVD--GATLQKSLDKFVRT-VTTPPEET----PC----
15 ----EVDEVAVRLRG-----PLFARADPSRAKL-----SGGPA--SCFESV
16 RSQG-GQTG---AIV-----EICKRSINF SYGIDLVRTDIPEHRICTSA-----
17 -----
18 -----DLFSWAV--EELKI
19 NPET-----
20 -----SFAVRAVV--VNEPG--KGRVVTAAASIAYLLV-MGVMG-----HIFQP-TI
21 WTERT-----SSGLKKTR-HLWT-----LKEGLHTEGLAWEKLGTP-----
22 -----VLALST-----DLEECTDFGHP--RLSTAVMQVMAKF-----
23 -CPEV--LRPLVEFA-TTVYHSERPI-H-L-----
24 -PDGS-----IRTKRRGWLMGDPFTKFMLSAMQEYVYQRMLI---
25 RWG---P-Q-----
26 -----FVT--TVVGDD--LVVL-SSNKPLL--EDYLSQYKV-AGFKVSILDTM
27 IS--DQI-MYYCEEA---S-----Q--IPQKAS-KAL-TV
28 SLK-----RGE-----
29 GEFCYLDY-----
30 -----PRFRLL-----IPTHPDVDRHSYTD--TGRVELL-----NRE
31 FSWVVVKW-----KTGSEE-----AFERA V-----LLQKII V--
32 P-RSSETSHPF--LPVWL--GGDG--AY---H-----PDPGYVF-----
33 -----EYVRSH-SWNVGEYAWRLQRYLEDTVAYTRTSRGPSVHRYHNIMALAEIDG-
34 -----LVPEEL-----RFPVVDEEREMW-----LAFKGTHILT PPEAAFKF
35 LKARYWLT VLEGEIPEET-----KISKVIIP-LGNVAPELTRK-SISE
36 LLDRM-----TERKPRLEEWK-----WLVDMSNAICADYMTLSW
37 ERLKPNPGEEWKLFTEGLL--DMHSAGMSEIYHAITHGDI DIPQGILSRMNVFFSSDPL
38 IIYR--LRKDVPDEENIALISRDFKLACKIQQMFPKG-----VAL-----VD
39 PLLYVLGM-----IP-----EEIDSRFKIIEDPGA
40 IYWTDINCFQDGAPREQWSD--
41 TVEDVYSGRLAQRWHPKGHRILNRRIGVYFAWTSKRR
42 QGGGY-----
43 >GBKL01051391.1
44 -----
45 -----FFFLKKKMSETIRWKP-----PLPSIGRNG
46 PTVVVSKASTLTEKLKVYTLETGICVSFFEKR-----IPRSQLVRLNISWRETEPES
47 -----RRLSSLSPDQKTAVILSNIWPKM
48 SP-----HFIT-----LVRAAPVTAIRNWFQTLDGLLIPAIL-----FG
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 LDNEYGLIDRITK-SCLERMA-----QNYSLFLREIK-----EVRKRFRLFFFESD
4 QRF-----GTLRYPGPSYGLTSILE-----FWRRILTPLIGNQR-----
5 -EFLRAVVTIGQT--RA--TGLVDGATLQKSLDKFVRT-VTTPPEET----PC-----
6 ----EVDEVAVRLRG-----PLFARADPSRAKL-----SGGPA--SCFESV
7 RSQG-GQTG---AIV-----EICKRSINF SYGIDLVRTDIPEHRICTSA-----
8 -----
9 -----DLFSWAV--EELKI
10 NPET-----
11 -----SFAVRAVV--VNEPG--KGRVVTAASIAYLLV-MGVMG-----HIFQP-TI
12 WTERT-----SSGLKKTR-HLWTF-----LKEGLHTEGLAWEKLGTP-----
13 -----VLALST-----DLEECTDFGHP--RLSTAVMQVMAKF-----
14 -CPEV--LRPLVEFA-TTVYHSERPI--H-L-----
15 -PDGS-----IRTKRRGWLMGDPFTKFMLSAMQEYVYQRMLI---
16 RWG-----
17 -----PQFVTTVVGDD--LVVL-SSNKPLL--EDYLSQYKV-AGFKVSILD
18 TM IS--DQI-MYYCEEA---S-----Q--IPQKASKALT
19 VSL K-----RGE-----
20 GEFCYLDY-----
21 -----PRFRLL-----IPTHPDVDRHSYTD--TGRVELL-----NRE
22 FSWVVVKW-----KTGGEE-----AFERAV-----LLQKII
23 V-----P-RSSETSHPF--LPVWL--GGDG-----AYH-----PDPGYVF-----
24 -----EYVRSH-SWNVGEYAWRLQRYLEDTVEVAYTRTSRGPSVHRYHNIM
25 ALAEIDG-----LVPEEL-----RFPVVDEEREMWLAF-----KGTHILT
26 PPEAAFKFL KARYWLT
27 VLEGEIPE-----ETKISKVIIP-LGNVAPELTRK-SISE
28 LLDRM-----TERKPRL
29 EEWK-----WLVDMSNAICADYMTLSW
30 ERLKPNPGE
31 GEEWKLFTEGLL--DMHSAGMSEIYHAITHG
32 DIDIPQGILSRMN
33 VFFSSDPL IIYR-----LRKDVP
34 DEENIALISRDFKLACK-----IQ
35 QMFPRKGV-----ALVDPLLYVLGM
36 IPEEIDS
37 RFKIIEDPGA IYWT
38 DINC
39 FQDGAPREQWSD-----
40 TVEDVYSGRLAQRWHPKPGHRILNRRIGVYFAWTSKRR
41 QGGGY-----
42 >NC_032404.1-----
43 -----MTNSCNNIPINMLLKEYNLQVGPCVSIFTK-----RVPIWWLKR
44 SNIR-----FYHDKVPFIKNLQPDDR
45 VRY-----LLSSIWTPRDRP
46 LSFAFRAKLKE-----LSTKGSINTIESWWNTVQGCCLPIMVGTN----LW
47 TDQIPNWCNRITK-WAFESCS-----RNYSLFQSHIK-----AERKLMRQYFMETG
48 IST-----LEGYKSSRREVVT-YVS-----VSQEVLRN
49 NLSTVRE--REV
50 ---MIGLLTQTRA--TG--LAD--QRMINASLEKFLSV-VSH--PCA-----
51 ----HITLNQKVMDSMVPRISG--YIMGYTKL-----SSGPS--AVFSHP
52 RTRG-GQTA--AIA-----ELSQRYITYSYDLHTLEKTTVERTR-----
53 -----I
54 DNAR-----KVLDYCI--HRAIH
55 RRTE-----
56 -----VLKVRVST--VVEPG--KARVITVGTLEYVQI-LSVIA-----HIFKE-SA
57
58
59
60

1
2
3 NSQRT-----RSGMKASR-HLWNF-----LWKDLDPRNSIWIETISGE-----
4 ---GTIYGLST-----DLETATDYGNM--SMSGQIWDLIQA-ATRVPG--
5 -VPIK--LLHLGKELYQS---DRPL---F-----
6 -SAGK-----LIATKKRGWLMGDPMTKILFTLIQEVISRSVGR---
7 TYR-----
8 -----IAC--SIVGDD--LVVL-SRERSGL--KEYLRLLET-LDLKISWKDTC
9 IS--DQY-IYYCEEL-----ARIPQS-AGVSLR
10 SLAR-----KNK-----
11 DYIGYIDV-----
12 -----PKLRTL-----IPTRGESDAFSNTN--LGRFSLQ-----GKS
13 SEYTSRV-NTGYYP-----VNERAL-----LLQCMF-----I
14 A-EDTSCIHPH--IPIEM--GGGG--SLVNDPF-----YVQRSIS-----
15 -----LCKYQ-REIT----YRIRSILSH----NGWGFRFVHGDRPDRVTHKY-HR-
16 -----FGKIIQDPRVFPEGCIKPRNKAQSEL-----LASVRSSKIIPPVDAFMR
17 LQRMYYFRGLFSGTG-----EVDPDPE-YLKSLEIPDGITKPSI
18 GYTKLNPLTLDEVSTFL-----SLWVKRGLAFRDDIPFWINREKLDINHPLVIELFK
19 KPDMPWTTEDLIEESEKLSLRLAENPWEIPDTLTPALSLVLDQDAILINILERKIKKITN
20 IGRLVVVTLDVRLLIRMNRLRKRGKSLDRIILIDPF-----LNLL
21 GMAPDFNP-----DIYLEDQGGITYVE
22 D-----MYFQDGMPVPEVEEHMFSDPVYFKLCKPDLPYIAASVIHTKRFP
23 SPKYRACGANLLMS-----
24 >GEMJ0101040.1
25 -----ENILTDHWRP-----PLPLTGRFG
26 TRKIIENLNDRSNSYLLKQYFLSS---TSCTN-----IFQKRIPVFWLKWNNGIPFKH
27 -----TYVPYLR-----DLPKYYQSKFLIESLYPRMSPRNF
28 CN----RLQK-----LCYEGKLHIVEQWFNTAEGCVLPLFLPD----LW
29 KDVNLKLIDRVTK-WSLEWCA-----CNYAHFISILK-----SFKKAIRKYFALHG
30 NL-----DGFTHPEDRTMVP-YLD-----LVRSLLKDDTLERK--LQ-
31 ---VILTCTQTTRA---TG--LSN--HQMMMNSLKKFKAT-VTT--PPV-----
32 -----HIEMDQLLLSE--ILNK--RVLRFSEPLSRI-----SAGPS--AVQGYS
33 RKSG-GQTK---ALA-----ELSTKYIDYDYDLRTLTIKEERVQRR-----
34 -----I
35 SSAA-----DVLNYSI--HQALT
36 ERKR-----
37 -----VSTVTVTL--VNEPS--KARTITVATLESVTL-LNVAM-----HIFAP-TC
38 NSQPV-----RSGMRANR-HLWQL-----SWKDLHPSNTLWDKLDHK--HNDT
39 GYNPLWA-LSS-----DLESATDYGNP--SVAKQVWRKLIDI-LSVFPD--
40 -APIG--LLELCCEIHTG---PRDV--L-Y-----
41 --EGE-----P-FCVTQRGWLMGDPMTKVILTIQVYVL-KQFQ---
42 QLI---P-A-----
43 -----VG--SVVGDD--LIIL-SYSRAHL--EMYPSMLEA-LDFKVSMEDTF
44 VS--KRL-MFFCEEG---S-----I--VPQTVR-DTIVCR
45 VKRPTWKKG-----YKP-----
46 ISCGYIDY-----
47 -----PRIRLL-----IPVRPDQDRLSYTT--LGKFSLL-----GRE
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 VIWAHSI-----NEPLFP-----VMERAV-----VLQNIIL--
4 P-QDSDTLSPF--LPIEI--GGNG-S---YCPD-----PDFLK-----
5 -SAVEKRSRSA-SQCR----YRMCQLLNHE----YGFRYLR--SERRNKVLHKN-VV-
6 -----VARLVE-----DESVFPPDSVIRAE-----SDEHKTILS-----
7 -----SLKAYISPEEAFMKLEYQKYFTQVLSGTGDIPSNFE-YQRDHSLPTNLGGVDV
8 EPNYSKFLKIWH-----KPGFTFKNSWKFVVNPKKLGNDPLSVDLKFKPDL
9 QVKPDPGDLLGRTLEDFLN--RVQAWD--MYFPD-----HLL-ERLEL----VV
10 ESDS----LLVALLKARLQRGETAQLCLVS-NDKK-----LLT--RLVTIWD
11 QLRPGECP-----LEVVLVHPDVYLTMYSWQEYMT
12 PPIMVLKDPGS---MFWES--NNFNNACQYWEVCWEVDKYQPLPGVATFTTLFYLRPGR
13 RHPQDDPVQEITPGRVFPHMGIQDVETSGADTPQEHIPEYVPPHLRK
14 >NC_035120.1
15 -----
16 -----MPPLPLIGRLGPRMAVSSDLLT-----TLDGDEKISS
17 SVPL-----EVLLERYFTSF--GGCAN-----LFYKKVPISFLRKNSVRFH
18 -----EEVPYLSNLS-----RAMQFRYLRKSVPWPQMSPVF
19 YQ----RLYR-----AVSAGRLTTIESWFNTAEAMCLPFLLG-----VD
20 SHTEYSCVDHITR-WSLANCA-----NGYASFIKTIK-----TFKKCLRKHLFSGE
21 PI-----DTFRHGSSRTCSP-YLK----WAKVVLGSIPEQQT--ALR
22 AQ-AVALLTQTRA--SG--LAD--HRMMSMQSITKFKSV-VTT--PYEHIQMDEP----
23 -----VLSRMIHPRMRRCVLNNA--KLGTGPS-----AIRGFP
24 RSSG-GQNL--ALA-----KTCTMTIDYTYDLRSLSVTSTENRR-----
25 -----I
26 SNSR-----DVLDYCI--SQALN
27 NRQK-----
28 -----VTEVTLSV--VNEPS--KARTITVGDYALVQL-LNVAA-----HIFKD-VC
29 CTQPV-----RSGMRADR-HLFNF-----VWKDLHPQNTLWDDMGWSYETKGM
30 P-----IHALSS-----DLETATDYANP--SVGRQIWDCCLISG-LEIQYP--
31 -ESSPRALLELCRD LHVG---PRTV---Y-----
32 -YQKI-----FFCTKLRGWL MGDPMTKVILTLAQEYV----LFRS
33 NAG-----RGPTGRLVG-----
34 -----SIVGDD--LVIL-SRLRHHL--GWYLDDLRS-LDFRVSDDDTF
35 IS--SDF-MFYCEET---S-----R--VPQGPG-QSVVAR
36 TKY-----SHG-----
37 TSCGYIDT-----
38 -----PRIRLL-----IPTRPDEDRFSNTN--LGRFSLL-----GKE
39 YQWCLGN-----NSDLAP-----LFRRAI-----GYQNCLV--
40 P-QDADTQCPF--MPVEA--GGNG--SYTTDTA-----FWSKLVL-----
41 -----RRSKDP-RTTH----FRVNQLLSNE--YAYRWIRSDRPLRGLAHRYSVVR-QV-
42 -----FEDETR----FPEGSVVKAEEHEERM-----ILRSFKQFLSPETAFMKL
43 EQARWFNGLFDGRGVIPP NMQYDREILLPSVGQSGPETLDPPV-IRRFLERWRNP-GFT-
44 -----YRDDYRFLVRR-----DFLPNEDPLFIEVDPNEP
45 DKLSASVVPDDLIGKT FEE--AVNKILTADLYLDDNLWSRLPLLIESDSMVMARWRELL
46 RRDTYVETLVLISTDRKLLL RMIDLFTVMRPDG YIVEAILMHPALFLT GWTGYMDYDQY
47 N
48 GIPGRVVI-----QDPGSIFHQDLYLPQRRSWWEG--
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----TVMLE--PFRTQEREGVTTIYHGYMEVRPELDRKPLPRSNASRIW
4 RHRE-----
5 >GCRV01019573.1
6 -----
7 -----RHVGCAIMQSADRSWAP-----PLPSLGGKV
8 GPRI-----MVNPNLSSDKLKA-----YYLELSSSPSMWIRRVPYAT
9 -----LRAVGYSKAKHLVDLPWLQYQMKV
10 LL-----KTLYFKASRHFVAKLESACRAGRINHIKQWWITADAVVLPYLLEL-----KE
11 NDADYQVVDNLTR-FALESCA-----NNYSGFIGRLK-----KVKKLIRKSLA---
12 -----LDQEIPFMRDMYP-----YIKQYKASSPSEGG--IHA
13 MAVHCNLWCQTRG--AG--LAD--SKMVNKAETKFLDC-IQRPKVSI-----
14 -----KLDRQILKDTT--LCCR--AANGVHAHL-----SAGPK--ACFEST
15 REEG-GQTA-----FLSKLARTRVVTHSYNYETLERVPLRRR-----
16 -----V
17 GSPK-----DLLDYCI--DWSLN
18 EPVL-----
19 -----RKVVRLLHT--VAEPA--KARVITVAPFAYNRL-MGVCA-----HMIAP-TL
20 RSTPL-----RSGMTASR-HLWGF-----THDTLHPGDKLWEQGIGC-----
21 -----EA-LST-----DWEVATDNGNV--SVARQVWSEIIRH-LSKIDG--
22 -APLG--FLNLCKTLYCE---DRYV---L-----
23 -HKGK-----IHRKTRGWFMDMMTKVIITLVHDYC----A---
24 RVA--L-L-----
25 -----KCY--SLVGDD--FVAM--GYNENL--QAYLQEVAR-LGIVISDDDTY
26 IS--PRF-MYYCEEM---C-----L--VPHTTS-QLP-VV
27 QIA-----MGK-----
28 SKISYIDT-----
29 -----ARLRLM-----LPCKQENQIFSATQ--VGRYSLL-----GKE
30 TRWALST-----HREVAP-----IFERGV-----VLQHLLL--
31 P-RDQDTQCPF--TPIEL--GGDG--SYTSDSN-----FFKNVIE-----
32 -----KKSKNP-AETH----YRIASLMRGA----FGYRYIC--TERVDQVSHKY-RM-
33 -----YVPLFQELVKYLPEEAVIPIHDGNRTL-----TTLRARLLETQRTILRI
34 VKEAYYRELFKGPVVAKL-----PTLRLEAPPPMPSTGG-ATPFYSASRFL-EHWR
35 NPGFK-----FRDQDPYLVRT-----DIALTLDHMSLGWDFSR
36 GPPKRALEDWLSQNLSLFD---GNANEVLAKLLSNEDLP-----EGAKSRLNLFFE
37 SDGW----VLNDLRSRAEIPPGPIYLISNDIRLGAD-----VSRVMEGKQDPR
38 EVYCVRPILSIIGRMDEVEDLVPAP-----ALSLEDPGSVSFETL
39 VTLGEDF-----PEWAF--NPLLLVTKFSGVFRIVPEIYAD-----
40 -----
41 >NC_032457.1
42 -----
43 -----
44 -----MSGHFQRALSKLCRSYDYRT
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----
53 IQ----RWYW-----TVDGLVPLLLLTH-----PD
54 PDSQMETIDRLIK-FGFENCA-----NNYSHFISRMK-----RLKKLLRKEFAEG-
55 -----TDLTSGQLFVDLSTYRK-----VFYRYLDLTLERSATE--
56
57
58
59
60

1
2
3 ---CVMAWTQARA---AG--LAD--RRMIEKSLEKAEKT-LSTPRRVYLRGRWRE----
4 -----IHAHYQGFIPV--LEPEVPRGPLL DVVRRVANKVRGETAHISTGPH--ACLEIK
5 QAFG-GQAAGIAYLC-----KRDVLTVSYDPVTLERTFHPARPCASAQ-----
6 -----
7 -----DLLSWAI--SRALD
8 APVH-----
9 -----
10 -----TRMVRIVA--VSEPS--KARTVTVGALAYQII-LGVVS-----KIFQPALA
11 SGISRTGLEGRTRNLYEALNNDF-DPSNG-----LWGPLRDMDRTGKFP-----
12 -----VFALTS-----DLEEATDYGDL--AVARQILQALLMR-CRDIPG--
13 -FPLG--LAVLAKSLFLS--SRIIIIRPGY-----
14 -GHGL-----VHWFRKRNGWLMGDRMTKVVLTLAHEIGILSAGI--
15 QFA-----
16 -----RICGDD--VFAL-SKAPAQL--QRYHKVMTD-LGFKISEDDHF
17 IS--RRI-LFYCEEV-----S-----L--VPQDAK-DLPSVC
18 N-----RRS-----
19 -----EPSCYVDY-----
20 -----PRIRLL-----LPIKVETNALSYTD--TGRFHLL-----GKE
21 MRWVFQN--SSQQAD-----PFIRAS-----LLQHIA-----I
22 A-MPRDILSPF--LPQEL--AGDG--AF--PH-----SAEFLLS-----
23 ---VIERKSINYDECL----YRIHSLTHGK----WGFRYLR--ADNINEVVHKY-HQ-
24 -----MVPKLK-----VLESVLPPDAVVRAS-----EVLISSLKVKG-----
25 -----LETPEKT-----
26 -----FFRLYRSYYWY-----
27 -----KVLHGMKAP--VLFNDEDRTRVHG-----
28 -----GREFKLDIPPDV LVERFY-----TT
29 WRDSGFTF-----
30 -----
31 -----
32 -----
33 -----
34 -----
35 >GFUS01019776.1
36 -----
37 -----NHVRWSPTLPSYGGRLCPRV-----LIPANLPDF
38 ASKLEAYYLHREGSSIFVRRVDV---GWLRK-----HGHERLIAQGPFLADLPGT
39 -----LYKYLIASLGSLTMSAHF
40 QR----SLSQ-----LCLAGKMGAIKRWFQTTDGVVLPLLSV-----ED
41 PDQHYADIDRLTK-FSLENCA-----NNYSHFIGRLK-----RVRRLLRKDYA---
42 -----EMGTSAKSLTD-MTT-----YRGMYLKHAMPKNRD--QYR
43 DVQYLMCWSQTRA---TG--LAD--ERMIRKSLDKAEIT-LSTPRRPRRSIPRED---
44 -----WVQQSTKLPLGEVIDRA--LLDSVCERIYRRC--DGRTGHLSAGPN--ACIESP
45 RNKG-GQTG-----AISRLTRMKVLHSTYDFKTLEVTRHPDRR-----
46 -----V
47 GSVQ-----DLLSWCI--QESLE
48 HSTY-----
49 -----VKCVKLHA--VAEPS--KARTITVCTLPYLLI-VGSLA-----KLLQP-SI
50 ASDIT-----RGGLESSR-NFWNF-----LYQDLDPTADVWHRLRDE--KAE
51 DGDGKTYALSS-----DLEEATDYGDI--GVARQILDGILRS-ASRIPG--
52 -FPVG--LALLAKTLFLG---KRFC---F-----R
53
54
55
56
57
58
59

1
 2
 3 SENGK-----WRYFVKRSGWLMGDRLTKVILTLAHELAIASGL---
 4 HIA-----
 5 -----RICGDD--VTAF-SKSRSRL--ARYLVVIRE-YGFKVSEDDSY
 6 IS--SRL-AFYCEEA-----AIPPVRFLDLPDV
 7 CT-----RRK-----
 8 TVPCYVDY-----
 9 -----PRIRLL-----IPTSSEVKAYTYTD--VGRFAQL-----GKE
 10 MRWILQG-----EIHPARD-----VYVRAS-----MLQHLLV--
 11 P-QPADTLCPF--TPVEV--GGDG--AF--PH-----SENFFLE-----
 12 --VLAKKARNP-TEIR----FRLHQLHRGV----WGYRLVR--HQYLDNVIHKH-HM-
 13 -----LVPKIE-----ELRSQIRECVPDALIEGDEMLNSISVSGLETPERT
 14 FFRMWRAWYWSEIFR---GHDPPETILEISRECNLPRSEVTLD-ALGFYRVWRKGKGFTFQ
 15 NRVDYLVITKMVQPLDYMNLGNFGVRDLPTENPDA---DQGRFLNEENSSSFIEYIRT
 16 GRPMQDQVVRDSDLHRMVESD---SFLKWEFSRRCHA-----DIPQLILLVSTDVNL
 17 GIDF----QRLAATRRPDYEDAPLVFLIKPTDYLGY-----FM
 18 SHYESWAA-----DLTQHWDVIVD
 19 FGCVVYNQFTLDSPEFEWFEDPDVVHNPVRRDASLRVYYLSLKEGEGGSERCSLTQPA
 20 S
 21 QWRKGFFPSQYVREGLLGN-----
 22 >GFOW01018115.1
 23 -----
 24 -----SIMSQKRWSPTLPSFGGRLCPRV-----YLPEGYPSD
 25 YCSR-----LEQFYLERSSSSIFVKKVDVTWLAKGHQDQLFLDGRFRAPSLGDL
 26 -----EWTVLYKYLMSGISQTMSLF
 27 QR----KLSD-----LCRYGKIGMIKRWYQTVDGLVLPLLSV-----ED
 28 PDLHYAEIDKLIK-FGFENCA-----NNYSHFISRMRK-----SLKRLLRKDYA---
 29 -----LDTESSKVLSDVST-----YRKIYLRFPPAGRGVIPAD
 30 AIRYVLIWSQTRA---SG--LAD--DRMCHQSLVKCEKT-LSTPRRARRIAPPEE---
 31 -----RRYSYRGLPLGEPVVPKI--HLSRICERVARRLGETAHL--SAGPN--ACLEST
 32 RQEG-GQTG-----YISRCRSKSLHAEYDFRTLECKQCTPRPVRSQD---
 33 -----
 34 -----DLLSWAI--QEVL
 35 HPTY-----
 36 -----VRCVRLHS--VAEPS--KARTITVCTLPYLLI-VGVL-----KLLQP-AI
 37 ASDIT-----RAGLHASR-NLWKF-----LYEDLDPGADLWFTLRDQFKVRGQ
 38 ISNLPIALSS-----DLSEATDYGDL--GVARQILAGLLHS-CSEVPG--
 39 -FPVG--LGQLAKTLFIG---KRFC--LRC-----
 40 -ENGR-----YRFFVKRNGWLMGDRLTKVILTLSHEIA----L---
 41 RES--G-L-----
 42 -----YIA--RICGDD--VIAF-SRSRDDL--ERYHQVVTD-LGFKVSEEDHY
 43 IS--SKL-AFYCEEA---C-----IPPTTY-RELPQV
 44 AI-----RRK-----
 45 TGSIIYIDY-----
 46 -----PRIRLL-----IPTCVETHAHSYTD--IGRFSLL-----GRE
 47 TRWVYQN-GAPHVAA-----LFKRAS-----LLQHIH-----V
 48 P-RTADTMCPF--TPLQI--GGDD--SF-----PFSGRFL-----
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 -----EAVVSKKARNESETR-FRMSSLMTGK-----WGYRLVR--NANLDNVVHKY-HT-
4 -----LVPKLK-----ELRNLLPEHAILEGN-----QFLLNNSIKIDGVESPERT
5 FFRIWRAWYWHQILKGSNPPE---LVLGVRKEANLPQTKVNID-FEYFFHIWKDLGFTFE
6 NVPDYLVSTRDVHVLDYMNMGWWF-
7 GTPPVGKTPDIFDIGGRFLLNEENLESFINHVKD
8 NRELPQVVRSTIYKYVESD--SYIKWEFSKRDDL-DSLTRIIFLVSSDLKLGAELARMA
9 RSRY--SSNLADAPVVVVPCWVSYGMM-DIPIDWS--TNSMSCPDPIVIVDQGSLTFD
10 QMVNSPEF-----SFFDKPRVT
11 -----TRWYRSNPWLGTQCLDIVEDGTTPSEVWSLTQH-----
12 PSQWRPGHFPYVPLRT-----
13 >GFRX01320111.1
14 -----
15 -----FFKKKTSHMMTQNRWAPTLPSFGGRWAPRV-----AIEAEYSAS
16 YLSR-----LERYFLRRQSCLSMFSHRVEAHALRECFGKDILTGYMVTLKGVTPRL
17 -----EQLLPDELYRYLVKSFGKISFSKNF
18 LR-----LCYSKDLATIKRWQTVVDGLVYPLLSSV-----ED
19 PDLYYYAEIDHLIK-FGLENC-----NNYAHFISRMK-----SLKRLLRKDYAE--
20 -----MKDSAKALTDVST-YRK-----IMMKFFPKDEANPH---L
21 TVQYLLCWTQTRA--SG--LAD--RHMIKSLDKCEAT-LSEVRRIARISTPET----
22 -----REYSYRGLPGEPELPQD--ILNDIAWRVSDKVK--GETAHLHSAGPN--ACIEST
23 RQTG-GQTG--YIK-----HLARMKDSLHTRYNFCTLEPEPITPSPVRSQN-----
24 -----
25 -----DLLSWAI--QEALE
26 HPTY-----
27 -----VRCVRLHA--VPEPS--KARTITVCSLPYLLI--VGVMA-----KLLQP-AI
28 ASDIS-----RGGLHASR-NLWNF-----LFNDLDPSDLWGYLRDDGDPKVP
29 -----IMGLSS-----DLSEATDYGDI--GVARQIFSrvLEK-CSSVPC--
30 -FPTA--LGNLAKTLFIG---RRYC--LRY-----
31 -HDGK-----GYRFFVKRNGWLMGDRITKVILTLAHEIV----V---
32 RSS-----
33 -----GLFCARICGDD--VIAF-SRRREDL--ERYLSEVTR-LGFKVSWEDSY
34 IS--SRL-AFYCEEA-----CIPLQRFREIPS
35 CT-----RRG-----
36 DQSMYVDY-----
37 -----PRIRLL-----IPTRVETQAYSFTD--TGRFSLL-----GKE
38 SRWVHQNKLNQTK-----LFHRAS-----LLQHLH-----V
39 S-QSPDTLCPY--LPIEL--GGDG--AF--P-----PSHTFLM-----
40 -----EVVRAKARDLRETH--FRMLQLLRSK----WGFRFVR--SMNLNDVVHKY-HL-
41 -----LVPKIK-----ELRELIPPESIIEGN-----ELLLGSIKIAGVESPEKT
42 FFRIWRAWYWHISILQTGREPPKLELGVSRHFPREETEVEMTEEHFHQFFHHWKDSGFTF
43 R
44 DTPDYLVDTRYIHVKDYMNMNWFGIRPAWVSYRPTMEQVGGRMLLNETMESFIDH
45 IKD
46 GRPLNPIVGRTLHKYVEAD--SYILSECARNWSR-LNPPDMCVLVSGDKRLACELARIQ
47 RAYY-----VPERDIYDKVPWVMMVDPNHYLFG-----MM
48 DPYEDYAE-----AYSSRFEVVV
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 DFGAVAYSAYLTGGVDMFYENPRIDTKYLWNNPQLGVFEASLDEKSKEVLVSLTTQ
4 PR
5 DQVLKRRIPWNFGSGLRGATMSP-----
6 >GFJG01122116.1
7 -----DRPNNKLG
8 EAPGPNDHTGKRPRS DTQEMSDAPPVKRPKLA VRRLLYNVVESV-----IKTAKRTKI
9 SPPEPQLPDVAQHTEESQYKCDIS---VVENN-----VLREICTRLAPYEYKGIVR
10 -----EFRNHPAHR RWYF---MIGRRLTDRTNQFVQSLTFRDYST
11 LE---QIWR-----AVMHTVLLNPDFDPSRF-----DP
12 DL--IESVRSFKL-WCITYASSHK----NGIVKLGKALK----TLASWMQWFMLDNQ
13 ANTCPNRVEIVPGWNRSTKRPTFPWFSGHLS-HCR----DFNRVKSDRDLH-----
14 --LLAQMRTEFGRA---LP--CPP--KILVDEGLKATLDV-LATE-RKT-----PD
15 HVLECISSAVSYLNTS---WKIS--NMVRES-----HLSISVS--ASYDHT
16 QEKG-GSAA---TQR-----EILSPFLDHKV DQLLDGNVPIKG NRLFIDCFGNIMLHGF
17 SMGRSIRTCRLVDGLYATHATVYGREEISKNLGKPMFGR RTGNVLLASIMKVQNYGH
18 FE
19 PSPK-----GFLRTDD--NQINI
20 PYWD-----ESIIDEGLPS
21 FIEEKPFPCR FIA--LAEPG-FKVRPLTIGYTPMMNV-LKNMR-----FMMQS-VL
22 LHDPR L-----RVGLRSTN-KLWTV-----LKLLKQKKLE-----YG-----
23 --EGLVL-QSS-----DYKSATDYIPL--DVIRTIWQTICDG-LTK-AHP-
24 -FRVF--RSLIWH-----PRQI---W-ADIPD-----
25 -DQGK--YT-----RQEQQHRCGSFMGEPM SYMTLCILNLAV----E---
26 LLS---Q-Y-----RLRTLSLTLKPIEYDGLNV LHGRGI-----
27 -----NLDVIIVVGDD--VLA F-RRDMGII--HIFHEVATQ-CGFIFSEGKDG
28 YS--SRI-GIICEDH-----LLVDRR-----
29 -----ENP-----
30 VTIRFLDI-----
31 -----VKSRL L-----TNVARTHGDNRASVFGKGRML-----NMV
32 LDYLPNTMFKEHVWL AYR-----SVLERT-----IRLKLS-----D
33 L-----DLPWF--LPPSL--GGLG-----LQLSG-----
34 -----DIPDSG-----WKFICFLEAR MSSA-ASWQDQLALVWSFRHLSNRQPHG-
35 -----WEPEQG--T VEMA VECLKELKHAPKEQ-----SLDPNLEKVIW-----
36 -----STDH-LRRFLRISGAAAPFVG
37 SAVQLMH-----DHEYYRTEAEK-----LGLVPLADVID
38 QFERLTTFQS ALEGTAEHR-FKTIQQWIRRSKKY W-----PKNM MNVR
39 KPLW-----FTNLSDFERRMIRRF DGF NIIDL C-----SR
40 SVNYGPNL-----RIV
41 Y-----DVWSQRRRGMP RIGLGGNPLSLKT-----
42 -----
43 >NC_013111.1
44 -----
45 -----MSKRPLAACLSVEETSFYAKTEVSK-----LCGLVDRMG
46 KLPR-----TYAGDDLPIDDNLKPVAR SV
47 -----LAPILGTRLTRERTEYICSLSSQ
48 LVVGISDV TYI-----LLGSALSVKEIRWQSQLFAIL-----AI
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 AGKLEAAFKSHTG-FLLATCLG-----DRYDSPDPSSS-----NYGGIAPGRIF---
4 -----PPRFQRKIW-MRC-----VLRRRPQDVTLASS-----
5 ---LVNIKRVAPP--VP-----DSFVQESLNKNRER-LSKP-QSD--CPDP-----
6 ----ILDMLERQVVRTVDEIVYNAKKEGWNRKIARSA-----FPSQS--ASFARS
7 IASG-GQLG---EIV-----PENWSLRFPPTLLAMVERCQEVVPI-----
8 -----Y
9 GWHE-----SPETVLT--ENLRG
10 HIDA-----
11 -----CLEVRRSP--VLEP--FKVRTITMGPAEPYFK-ARRIQ-----GVLWD-LL
12 KHTRCTHLP----NRPIHESD-ISFFC-----SRRGDAVFPDEE-----
13 -----TFFVSG-----DYSAAATDCLSP--VLSTVAVDRLCDHLLSPENQV-
14 -LDPVHPWRALFHRLVVG---HKIM--E-G-----
15 -KRG-----EVEIAAQWTGQLMGSPLSFPILCIVNLAI----T---
16 RAS-----IEHTAKRQLYL-----
17 -----EECGILVNGDD--ILFK--LPLRGL--ARWDFLVTR-GGLHPSPGKNF
18 VS--KDY-AVLNSTI----Y-----RVGRTA-----
19 -----
20 ELIPTIKG-----
21 -----NLINGT-----ACRGIERPRDGSLYFSD-----ASK
22 YSWGTTIG---DRARELIK-----GFPNHI-----ADQLLS--RFLISMKPHLDLFP
23 R-ISWWAHPQY-----GGLG--L-----PVTRPGD-----
24 -----FLPVH-----RRVGAFLAC-----GGRRSQE-----FRMNMQ---
25 -----WLSNPVKQFNAFTSQYLMIDIARSLSV-----VKQIYNDTEP-----
26 -----E-FFLEREIMYSALRGV
27 EFPTNA-----RGNRELLAAWR-----HFYRQLE
28 RRALRTRITDGDLPERKAD--RKGLFLLSPEKLI-----AGPRFKYVYDWACKNQG
29 GNIWSSSSRGYVLFDKHRVLESSALPFGKDEVKSSVS-----THFT
30 PVNEYRQR-----
31 -----IFFWT-----
32 -----
33 >NC_011065.1
34 -----
35 -----MSNLADHEQVRTYSAYCSRRLCGLTD-----RIGKLPECN
36 LAPA-----FLKEL-----KLPGLPVPQTLYLKATRWLT
37 -----LSDRDEEREY
38 VP----VLIS-----HLITAVYDVVKFLLGRQIT-----RK
39 DIHYMVGVYTLLA-----MEGRLEKSFK-----EHTAWLVARAMGDV
40 APEVPPTDGISA-----GKIFSPRLNRIIW-MRC-----VLRRRPQDVSL-----
41 ---AHSLYQVKRV--AP--PLV--TKLVNKASEEMLRR-VTKEEDGEHHILRR-----
42 ----LEGEIERTVDELVKNARA--ERWTGNIRGQP-----FPAQS--ASYEHG
43 VREG-GQLH---SLL-----GTGHGLNIPHLIAMIQRGQYVYPVYGWD-----
44 -----E
45 SVGP-----NLKNVLE----MV
46 DCPE-----
47 -----FLPTRRSP--VLEP--FKVRTITMGPARHYFR-ARQIQ-----GVLWD-LL
48 KGSRST----HLPNRPIQESD-ISFFV-----QRAGDRIFPGE----EA-----
49
50
51
52
53
54
55
56
57
58
59
60

1
 2
 3 ----LY-VSG-----DYSAAATDNLRP--SLSMTCVNRLCDH-LSQSGSAL
 4 DAAHP--LRMLFQRV-LV---GHAI---M-----
 5 -DGKRGEES-----EAGRQSWGQLMGSPLSFVLCIINLAA---T---
 6 RWV-----LEQVSGRIVSL-----
 7 -----NQSGILVNGDD--ILFR--MPYRGY--ALWNRVVTA-AGLTPSPGKNF
 8 VT--SRY-AVLNSEV---Y-----D--TSSRTA-----
 9
 10 -----
 11 DPVPFLKI-----
 12 -----NLMYGT-----LARGCERRTSAELLYGDDLEAGGTL----GHR
 13 CRALIKG-FKPEQQ-----DYLITR-----FIRACQ-----PF
 14 L-SQAPGVSWF--LHPAY--GGIG--L-----P-----
 15 -----SVRRPRIEIH----L RVAAYLSC-----GGKDQAEAQVFHMWLKQPS-KE-
 16 -----FNRATL-----LRAMEVARELHVPT-----VQVSPEELKEAPE-----
 17 -----RFEIP-VKQELLSSALT LGTEF
 18 EVTD-----QGQSARLREWN-----RYYRLLV
 19 RRSLKTYVRGSKDTTGLGL--HAMGWEKAMRGPK-----FRT
 20 VFDW-----SARSGGYSMRLKTTSSYSRC-----PFS
 21 SSESILP-----
 22 -----SGGVR--VCYSFLNTQ-----
 23
 24 ----->NC_011068.1
 25
 26 -----
 27 -----
 28 -----MEPRQ-----ELLDPERVR
 29 EEAK-----RIVR--WLCGL-----VDRTGKLPGDYQGKILNTV
 30 -----SEICKRSKLP CGELAEALTKGRL
 31 TR----SLVD-----YSELLISNLVVGYFDVLEIYL-----GK
 32 QSVRLSDIREMACKYTFYAINR-----RLEDYIK-----FQTAWLQARAMRDV
 33 TP-----EPEAPS WLNEGFG-RCF--SALLNRKVHLRCVLRAR--PND
 34 VSLAASLYQVKRV--AP--PLP--DDQIEKNLEKSLDR-LTKD-EEP--AGVD---EP
 35 -----FLEDLKRECKRTVDELVQ--NARREGWNRKISR-----DCFPSQS--AAFENP
 36 ISKG-GQLG-----QLVKENNTPRLPVLLGMFEYKGRVTPV-----
 37 -----Y
 38 GWAD-----DGDTILS--DEELG
 39 REVP-----
 40 -----
 41 -----
 42 -----AALKCRRSP--VLEP--FKVRVITMGPAVQYYR-ARRVQ-----GCLWD-LL
 43 KHTRCT-----HLPNRPVEESD-IGFYV-----RRRGADLFRGE-----EV-----
 44 -----PY-VSG-----DYSAAATDNLHP--DLSLSVVDRVCDHLLSDDNRPL
 45 DPVSP--WRVLFHRV-LV---GHRI---Y-----
 46 -DGNSSRNT-----EVAAQSWGQLMGSPLSFVLCIVNLAV-----T---
 47 RYV-----LEKACGRIVTL-----
 48 -----EESGILVNGDD--ILFR--CPERTI--PFWTRMVVTI-AGLSPSPGKNF
 49 VS-----YRYCQ-----LNSELY--DMSG-
 50 -----SRA-----
 51 -----
 52 EYLPFIKA-----
 53 -----NLIYGT-----LARGCERKRAADLC--YGDTTTE----GGT
 54 FGHRARA--LIK-----GFGPD M-----QDRLMS-----RFLHSIKGF
 55
 56
 57
 58
 59
 60

1
2
3 L-EKIPEVSWF--IHPRY--GGLG-----LPLTR-----
4 ----PVTHNP-----YHLRIAAYLSCGGEQSQEARM-----MQ-
5 -----WLSAPT-----KSFNAATLLRI-----LEVARNCKVPF-----
6 -----RKVP-FALLHRAEAAGVDLEA
7 L-----FRKALLRSAPR-----LGVEY
8 PSNESGDMQRLGDWRRFFR---DVGKAARTRCRS-----GTADER
9 KGLF----LMSPDNAVKGPQYEYIFDWASHNMGGN----I-----WD
10 PSYKFRAD-----
11 -----PFSSS--ESDEPRAKEIGPNRGPE-----
12 -----
13 -----
14 -----
15 >GADD01004427.1
16 -----
17 -----SPKGPRVYPGRLGRAQM-----GAPAASAPG
18 GSPE-----ALIDFDLGPISWDLKRW-----LPDNPVVALWSVRGKANADD
19 -----QRRSFNLMTNMAIRRSSLEKFRK
20 LRHAERKTLSD-----LTSGLLDNLIASNPER-----IR
21 TARTDPDVRKILT-WAFSLGL-----YNVQHAISEWK-----KFCTLLKHRALQSE
22 TPWPEPLSDFPGFGRDACPDCPALWARLCPW-LKI-----IWSQGASSKSDMT-----
23 ---RICHLATSRN---LP--TAS--AAEREKTLLSHASV-LTSTPIP-----IE-----
24 -----RSSMLRMLSRRIGEMVKG--KRKGRRSSGHC-----SLAQS--ACIEQP
25 AGKG-GRAF---YIQ-----SEFQTWASRKQTETISETTWFGQPFRLKEGLPMWVTMCR
26 TED-----A
27 SNPT-----MAIGESRQDFNMNFDDFRHEDFLAGLDKATG--YQLLQ
28 WSLE-----EGIRRGCLLGDPFRSSGLA
29 ISKEKLPEIKADI--VGEPG-GKSrvMTIGEGWLTI-F-LQPAS-----HHLAG-LI
30 EHHEAA-----WAGFKRGW-QASEW-----AKRYRRKQHWT-----GE-----
31 --D--YV-LSS-----DLKTATDYCRH--DFSLSLLLGFLEG-SG--QG--
32 -DNPY--LRAAAELL-CS---PRYY--T-G-PV-----
33 -EGSI--GP-----TSRGILMGDPGSKIVLTLFNLAEEEEAFV---
34 RWA-----TPRMPIPTISVGMVTIADSATSW-----
35 -----KIRNLAVSGDD--HTAR--GPKSYL--KGITQAHIA-NGMEVSESSNF
36 LS--PIG-GFYCEEA---L-----L---FRGLDL--QKLRA
37 YSVFWDI-----
38 PYLEHMHV-----
39 -----DSLKLRC-----SPVMKGDATRDEKNPA--IGKASTF-----TGM
40 LRWFAGG-----WACTR-----PIFSAR-----FEQRFK-----T
41 F-LPE-GILRY--LPREL--GGVAAPCF-----HLTAGMLE-----
42 ---SSLTKLD-PSVR---SAINVLAGE---RGELAKL-LAKIPANTSMRG-IP-
43 -----ENLARD-----QVTQILETF-----PGVQGRIDE-----
44 -----SQ-LKAMADDQMD-QWWG
45 DAT-----ARDRARFARRN-----NLISVVVDALN
46 DIERPYLFREFLAPGQHAV--HEGPSQFKTRIYD-----
47 TPPW-----PVRVNRFRSLALSAKPDLDDHHHS-----AE
48 GVSDGKLV-----GMIQKGKGLPP
49 P-----PRMYI--PERVIFGDDLVTLQTPL-----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 >GFOL01002916.1
4 -----
5 -----
6 -----
7 -----
8 -----
9 -----
10 -----AISEWK-----KFCTLLKHLRQLQSE
11 TPWPEPLSDFPGFGRDACPDCPALWARLCPW-LKI-----IWSQGASSKSDMT-----
12 ---RICHLATSRN---LP--TAS--AAEREKTLLSHASV-LTSTPIPD----IE----
13 ----RSSMLRMLSRRIGEMVKG--KRKGRRSSGHC-----SLAQS--ACIEQP
14 AGKG-GRAF---YIQ-----SEFQTWASRKQTETISETTWFGQPFRLKEGLPMWVTMCR
15 TED-----A
16 SNPT-----MAIGESRQDFNMNFDDFRHEDFLAGLDKATG--YQLLQ
17 WSLE-----EGIRRGCLLGDPFRSSSGLA
18 ISKEKLPEIKADI--VGEPEG-GKSrvMTIGEGWLTI-LQPAS-----HHLAG-LI
19 EHHEAA-----WAGFKRGW-QASEW-----AKRYRRKQHWT----GE----
20 --D--YV-LSS-----DLKTATDYCRH--DFSLSLLLGFLEG-SG--QG--
21 -DNPY--LRAAAELL-CS---PRYY--T-G-PV-----
22 -EGSI--GP-----TSRGILMGDPGSKIVLTLFNLAEEEEEAFV---
23 RWA-----TPRMPIPTISVGMVTIADPATSW-----
24 -----KIRNLAVSGDD--HTAR--GPKSYL--KGITQAHIA-NGMEVSESSNF
25 LS--PIG-GFYCEEA---L-----L--FRGLLDL--QKLRA
26 YSVFWWDI-----
27 PYLEHMHV-----
28 -----DSLKLRL-----CLSPV-MKGDATRDEKNPA--IGKASTF----TGM
29 LRWFAGG-----WACTR-----PIFSAR-----FEQRFK-----T
30 F-LPE-GILRY--LPREL--GGVAAPCF---H-----LTAGMLE-----
31 ---SSLTKLD-PSVR----SAINVLAGE----RGEELAKL-LAKIPANTS MRG-IP-
32 -----ENLARD-----QVTQILETF-----PGVQGRIDE-----
33 -----SQ-LKAMADDQM GD-QWWG
34 DAT-----ARDRARFARRN-----NLISVV DALN
35 DIERPYLFREFLAPGQHAV---HEGPSQFKTRIYD-----
36 TTPW-----PVRVNRFRSLALSAKPDL DDHHHS-----AE
37 GVSDGKLV-----GMIQKGKGLPP
38 P-----PRMYI--PERVIFGDDLVTLQTPL-----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 >GFOL01000185.1
46 -----
47 -----SFNLMTNMA
48 -----IRRSSLEKFRKLRHAERKTL
49 -----
50 -----SDLTS-----GLLDNLIASNPER-----IR
51 TARTDPDVRKILT-WAFSLGL-----YNVQHAI SEWK-----KFCTLLKHLRQLQSE
52 TPWPEPLSDFPGFGRDACPDCPALWARLCPW-LKI-----IWSQGASSKSDMT-----
53 ---RICHLATSRN---LP--TAS--AAEREKTLLSHASV-LTSTPIPD----IE----
54 ----RSSMLRMLSRRIGEMVKG--KRKGRRSSGHC-----SLAQS--ACIEQP
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 AGKG-GRAF---YIQ----SEFQTWASRKQTETISETTWFGQPFRRLKEGLPMWVTMCR
4 TED-----A
5 SNPT-----MAIGESRQDFNMNFDDFRHEDFLAGLDKATG--YQLLQ
6 WSLE-----EGIRRGCLLGDPFRSSSGLA
7 ISKEKLPEIKADI--VGEPG-GKSrvMTIGEGWLTI-LQPAS-----HHLAG-LI
8 EHHEAA-----WAGFKRGW-QASEW-----AKRYRRKQHWT----GE-----
9 --D--YV-LSS-----DLKTATDYCRH--DFSLSSLGFLEG-SG--QG--
10 -DNPY--LRAAAELL-CS---PRYY--T-G-PV-----
11 -EGSI--GP-----TSRGILMGDPGSKIVLTLFNLAEEEAFAV---
12 RWA-----TPRMIPTISVGMVTIADPATSW-----
13 -----KIRNLAVSGDD--HTAR--GPKSYL--KGITQAHIA-NGMEVSESSNF
14 LS--PIG-GFYCEEA---L-----L--FRGLDL--QKLRA
15 YSVPFWDI-----
16 PYLEMHMV-----
17 -----DSLKLR-----CLSPVMKGDACTRDEKN-PA--IGKASTF----TGM
18 LRWFAGG----WACTR----PIFSAR----FEQRFK-----T
19 F-LPE-GILRY--LPREL--GGVA--A--PC-----FHLTAGMLE-----
20 ---SSLTKLD-PSVR----SAINVLAGE----RGEALKL-LAKIPANTSMRG-IP-
21 -----ENLARD-----QVTQILETF-----PGVQGRIDE-----
22 -----SQ-LKAMADDQMHD-QWWG
23 DAT-----ARDRARFARRN-----NLISVVDALN
24 DIERPYLFREFLAPGQHAV--HEGPSQFKTRIYD-----
25 TTPW-----PVRVNRFRSLALSAKPDLLDHHS-----AE
26 GVSDGKLV-----GMIQKGKGLPP
27 P-----PRMYI--PERVIFGDDLVTLQTPL-----
28 -----
29 >GFOL01000028.1
30 -----
31 -----SPKGPRVYPGRLGRAQM----GAPAASAPG
32 GSPE-----ALIDFDLGPISWDLKRW-----LPDNPVVALWSVRGKANADD
33 -----QRRSFNLMTNMAIRRSSLEKFRK
34 LRHAERKTLSD-----LTSGLLDNLIASNPER-----IR
35 TARTDPDVRKILT-WAFSLGL-----YNVQHAISEWK----KFCTLLKHRALQSE
36 TPWPEPLSDFPFGFRDACPDCPALWARLCPW-LKI----IWSQGASSKSDMT-----
37 ---RICHLATSRN--LP--TAS--AAEREKTLLSHASV-LTSTPIP-----IE----
38 -----RSSMLRMLSRRIGEMVKG--KRKGRRSSGHC-----SLAQS--ACIEQP
39 AGKG-GRAF---YIQ----SEFQTWASRKQTETISETTWFGQPFRRLKEGLPMWVTMCR
40 TED-----A
41 SNPT-----MAIGESRQDFNMNFDDFRHEDFLAGLDKATG--YQLLQ
42 WSLE-----EGIRRGCLLGDPFRSSSGLA
43 ISKEKLPEIKADI--VGEPG-GKSrvMTIGEGWLTI-LQPAS-----HHLAG-LI
44 EHHEAA-----WAGFKRGW-QASEW-----AKRYRRKQHWT----GE-----
45 --D--YV-LSS-----DLKTATDYCRH--DFSLSSLGFLEG-SG--QG--
46 -DNPY--LRAAAELL-CS---PRYY--T-G-PV-----
47 -EGSI--GP-----TSRGILMGDPGSKIVLTLFNLAEEEAFAV---
48 RWA-----TPRMIPTISVGMVTIADPATSW-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----KIRNLAVSGDD--HTAR--GPKSYL--KGITQAHIA-NGMEVSESSNF
4 LS--PIG-GFYCEEA---L-----L---FRGLDL--QKLRA
5 YSVPFWDI-----
6 PYLEHMHV-----
7 -----DSLKLR-----CLSPVMKGDATRDEKN-PA--IGKASTF----TGM
8 LRWFAGG----WACTR-----PIFSAR-----FEQRFK-----T
9 F-LPE-GILRY--LPREL--GGVA---A---PC-----FHLTAGMLE-----
10 ---SSLTKLD-PSVR----SAINVLAGE----RGEELAKL-LAKIPANTS MRG-IP-
11 -----ENLARD-----QVTQILETF-----PGVQGRIDE-----
12 -----SQ-LKAMADDQMGD-QWWG
13 DAT-----ARDRARFARRN-----NLISVVDALN
14 DIERPYLFREFLAPGQHAV---HEGPSQFKTRIYD-----
15 TTPW-----PVRVNRFRSLALSAKPDLDDHHHS-----AE
16 GVSDGKLV-----GMIQKGKGLPP
17 P-----PRMYI--PERVIFGDDLVTLQTPL-----
18 -----
19 ----->GFOL01002977.1
20 -----
21 -----
22 -----
23 -----PAASAPG
24 GSPE-----ALIDFDLGPISWDLKRW-----LPDNPVVALWSVRGKANADD
25 -----QRRSFNLMTNMAIRRSSLEKFRK
26 LRHAERKTLSD-----LTSGLLDNLIASNPER-----IR
27 TARTDPDVRKILT-WAFSLGL-----YNVQHAISEWK-----KFCTLLKHRALQSE
28 TPWPEPLSDFPFGFRDACPDCPALWARLCPW-LKI-----IWSQGASSKSDMT-----
29 ---RICHLATSRN---LP--TAS--AAEREKTLLSHASV-LTSTPIP-----IE-----
30 -----RSSMLRMLSRRIGEMVKG--KRKG RSSGH-----SLAQS--ACIEQP
31 AGKG-GRAF---YIQ-----SEFQTWASRKQTETISETTWFGQPFRLKEGLPMWVTMCR
32 TED-----A
33 SNPT-----MAIGESRQDFNMNFDDFRHEDFLAGLDKATG--YQLLQ
34 WSLE-----EGIRRGCLLGDPFRSSSGLA
35 ISKEKLPEIKADI--VGEPG-GKSrvMTIGEGWLTI-LQPAS-----HHLAG-LI
36 EHHEAA-----WAGFKRGW-QASEW-----AKRYRRKQHWT-----GE-----
37 --D--YV-LSS-----DLKTATDYCRH--DFSLSLLLGFLEG-SG--QG--
38 -DNPY--LRAAAELL-CS---PRYY--T-G-PV-----
39 -EGSI--GP-----TSRGILMGDPGSKIVLTLFNLAEEEEAFV---
40 RWA-----TPRMIPTISVGMVTIADPATSW-----
41 -----KIRNLAVSGDD--HTAR--GPKSYL--KGITQAHIA-NGMEVSESSNF
42 LS--PIG-GFYCEEA---L-----L---FRGLDL--QKLRA
43 YSVPFWDI-----
44 PYLEHMHV-----
45 -----DSLKLRLC-----SPVMKGDATRDEKNPA--IGKASTF----TGM
46 LRWFAGG----WACTR-----PIFSAR-----FEQRFK-----T
47 F-LPE-GILRY--LPREL--GGVAAPCF-----HLTAGMLE-----
48 ---SSLTKLD-PSVR----SAINVLAGE----RGEELAKL-LAKIPANTS MRG-IP-
49 -----ENLARD-----QVTQILETF-----PGVQGRIDE-----
50 -----SQ-LKAMADDQMGD-QWWG

1
2
3 DAT-----ARDRARFARRN-----NLISVVDALN
4 DIERPYLFREFLAPGQHAV---HEGPSQFKTRIYD-----
5 TTPW-----PVRVNRFRSLALSAKPDLDDHHHS-----AE
6 GVSDGKLV-----GMIQKGKGLPP
7 P-----PRMYI--PERVIFGDDLVTLQTPL-----
8 -----
9 -----
10 >IABI01004424.1
11 -----
12 -----LDPLPPCPPVAAVISTEGNGFVSRQQSE-----PLARSPKDV
13 GPPKGDLSSFVPTYGGFNAGIVDRPGGLATHL-----AWRNPLSEKILRRSRNGKPY
14 -----GVVAYSNGVK-----PSLREAILRRHPEICTYPDVVIRR
15 LSEDG--HRFN-----RIRGAYLSIEDAIVCA-----VP
16 EILSDDPVFPLLRRWVLTNGAVH-----EGVDQVIKMWK-----SMIVSWRWSSLRSR
17 TEQPLAVPGSIFFN---EKGHFVGFPPEESI-LSS-----VMEDIFITGITSKH--DGT
18 ---RVCHLC SARG---MP--PAG--QKKSEDSLREHHET-LTGQ-RDD---ISQE----
20 -----RLDLVKKLARRIGRKVAT--VLSKEKGDFRSDG-----HVSATTS--ASVEDT
21 VKSG-GRAN---TLA-----TQFIPWAKKVATEPLEGNIDILGNNA VRERPGHARWRTTCR
22 EVPVPFERDCDCAGEDGVRPNSSNGLPAVHSLEANREECPNLGLFSSIWRMPE----F
23 DMPT-----NLDNPLTGVDESIG--KQMLQ
24 WAHH-----QGCSAGYIEGPLVGSREEFNTSKVVK-
25 -----PIPVVRSV--VREPG-GKARVVS KNEAWFTTF-LQPFG-----HHSVG-LL
26 SRLPST-----RAGLSAAA-QGYEW-----AKSASKSTGIDGASG-----
27 -----MLFLSS-----DLKTATDFCDH--RVSKSLLEGFYEG-VGTNP---
28 -DNPY--IRLCIEIL-CS---PRIL-----
29 -EDGS-----ITSRGIMMGEPGCKAVLTFLNLVA----E---
30 EEA---F-----TAHCQNYPEILD LAVVPLRKSW-----
31 -----RVP--ACAGDD--HIAY--GPETYI--RRITRNHHS-NGMQVSWPSNF
32 VS--PIA-AIYCEEM---I-----LKLPHLRFEVGYL
33 PKADYRIT-----
34 VHIDAMKV-----
35 -----RLLSPY-----RKETSGGPNDSDKTNPAVGKSSSF-----TKK
36 LDWQPKG-WELARK-----PFLERF-----RSRMMQ-----Y
37 L-PHIEAGYQY--LPRYL--GGLG-----VPSVIPS-----
38 -----SLRTLSDLSP----RHVA AIEAIMDGRGTPAIRRCL--SRFASNDEARGIDL-
39 -----SNTAAE-----DENIALMKRAIQGF-----DTPEEGVLT F E-----
40 -----A-VREKLGLSFDQAENMR
41 -----RRDFNRKAGRE-----LQLISITDALR
42 EIERPYITVELLKGKASEP-----KSRFDV-----
43 -RPW-----KVRYAKLQDDLET CGVPKNPNNVLV-----AS
44 TLDEVFNA-----VITGNDEYI-
45 -----PKERISRVISAFGV EVVR-----
46 -----
47 >GEWH01000774.1
48 -----
49 -----KFNL DSTPCARVAA ALDDARYGEA-----QATRYSIPD
50 PDATGVELVTRIPWRNPLTEKSEVSFGGKR-----RQRRKPLKRPSYHSMIKPSQ
51
52
53
54
55
56
57
58
59
60

1
 2
 3 RFRILSVTPIIKELSRRQLNLSE
 4 DG----DRFN-----RVRGCFLGIEDAIYSSC-----PE
 5 LRDDDPPIWSKLLK-WTFATGSSRH----SGVDTVIKSWK-----SFATWVKWAALRSE
 6 TEEPAV-----IRGFPFFDENGFTGFPADSQFSELIEPILRTGITSKGEGT--
 7 ---RICHLITSRG---LP--PAG--QSKADASLDEHFKV-LTSDSGIS----DE----
 8 ----RVEEVRIAKRIGKLVRGCIVKEGEDRFRSDGHV-----SATSS--ACVEST
 9 VKQG-GRAA---SMS-----ANLLTFLNEIPKTETKGVDILNNPFIETGDLPRWRTTCR
 10 DEV-----N
 11 ELSRDCDCFETCLNKWECKNNTWSSSYWSIPDFESDLENMYDNPLMGVDHAIG--
 12 PQMVQ
 13 WAYL-----EGRYRGVFTGPRVGDKQQDFLT
 14 CTVSRPPQIIRSV--VAEPG-GKARVVTKGEDWLTVF-LQPFS-----HHTIG-LL
 15 RRLPSA-----RAGLSAAA-MGYEW-----TKALSRKEGLD---KE----
 16 --DNLYF-LSS-----DLKTATDYCDH--RISKALLEGFYSG-IGW-YK--
 17 -DHSY--IKLAIDLL-CS---SREC-----
 18 -EGEF-----TKRGILMGDPGTKAVLTLFNLVAEEESFQ---
 19 DWK-----RKNPEREVKW-----
 20 -----RCP--AFAGDD--HVAL--GPKEYV--SQITVNHAR-NGMKVGVWESNF
 21 VS--KIG-AFYCEEL-----LAHVNG--TKFEV
 22 GYLPKAEY-----EST-----
 23 IHVDAMKI-----
 24 -----RLMSPY-----RKETSAAPNDSDKANPAIGKASSF----YKK
 25 LSWFQDG---WKSACA-----PLLQRF-----RLRMMD-----Y
 26 L-PAENQVYQY--LPYYL--GGLDMPMLLEGNP-----SVSDWTE-----
 27 ----LIPQLD-LAHQ----SAIQMVIDGKADSWLLRCLSRFASNDYARGIDLSN---
 28 ----RYAEDS-----IRGLLQIKKISTQE-----TESTYEQVVE-----
 29 -----ELV-KTGISKTSIEDMR---
 30 -----KRDFGRKAQDQ-----LKLISMSSALN
 31 LIEKPYIVTNLLRGEVDSQKQKYDLVPWKTRYNR-----FIKDLESVGVSAE
 32 GYEEY-----KTGMTKDDIFNRFVSGKDIYIPER-----RIT
 33 SILTAYKV-----
 34 -----GVWSTPSLLTETKLS-----
 35 -----
 36 >GEWH01001802.1
 37 -----
 38 -----PILRTGITS
 39 KGEGR-----ICHILITSRGLPPAGQSKADA
 40 -----
 41 -----SLDEHFKVLTSDS-----GI
 42 SDERVEEVRIAK-RIGKLVRG-----CIVKEGEDRFRSDGHVSATSSACVESTV---
 43 -----KQGGRGAAAMSAN-LLT-----FLNEIPKTETKGVD-----
 44 -----ILNNPFIETG--DLPRWRTT--CRDEVNELSRDCDC-----
 45 -----FETCLNKWECKNNTWSSSYWSIPDFESDLE-----NMYDN--PLMGVD
 46 HAIG-PQM-----QWAYLEGRYRGVFTGPRVGDK-----
 47 -----
 48 -----QQDFL

1
2
3 TCTV-----
4 ---SRPPQIIRSV--VAEPG-GKARVVTKGEDWLTVF-LQPFS-----HHTIG-LL
5 RRLPSA-----RAGLSAAA-MGYEW-----TKALSRKEGLDKEDN-----
6 ----LYFLSS-----DLKTATDYCDH--RISKALLEGFYSG-IGW----
7 -YKDHSYIKLAIDL-CS---SREC-----
8 --EGE-----FTKRGILMGDPGTAKAVLTLFNLVAEEESFQ---
9 DWK-----RKNPEREVKW-----
10 -----RCP--AFAGDD--HVAL--GPKEYV--SQITVNHAR-NGMKVGWESNF
11 VS--KIG-AFYCEEL-----LAHVNG-TKFEVG
12 YL-----PKA-----
13 EYESTIHV-----
14 -----DAMKIRLM-----SPYRKETSAAPNDSDKANPAIGKASSF-----YKK
15 LSWFQDG---WKSACA-----PLLQRF-----RLRMMD-----Y
16 L-PAENQVYQY--LPYYL--GGLDMPPLLEGNPE-----SVSDWTE-----
17 ---LIPQLD-LAHQ----SAIQMVIDGK----ADSWLLRCLSRFASNDYARG-ID-
18 -----LSNRYA-----EDSIRGLLQIKKIS-----TQETESYTY-----
19 -----EQV-VEELVKTGISKTSIED
20 MR-----KRDFGRKAQDQ-----LKLISMSSALNLIE
21 KPYIVTNLLRGEGVDSQKQ---KYDLVPWKTRYNR-----FIKDLESVGVAE
22 GYEEY-----KTGMTKDDIFNRFVSGKDIYIPER-----RIT
23 SILTAYKV-----
24 -----GVWSTPSLLTETKLS-----
25 -----
26 >KX373304.1
27 -----
28 -----
29 -----
30 -----
31 -----
32 -----MPTIRNELNPRV-----SFFAEKHGK
33 GTFK-----CPLSLRHYLGDSRHLRKCL
34 -----LHPRAGLLKEV
35 LG----KSWK-----TFKCFKTSSIVRAHELYLCYLDICVCL-----DP
36 EILSIESNRSIGV-KVFNTAC-----FSTSTATSLYK-----EFVENVLILLK---
37 -----CKEGERWTRQNPF-KRM-----MTWKIFQDAIASVQLSKGK
38 SLERVLSVLTRN---FP--EPD--GKSIQKKKVDFIGI-ISQQPPKG--WKSEE----
39 -----YRSHFLESIEEIANECES---TTVSDCHI-----SVTAA--GSLKKT
40 VREG-GKFA---EMI-----EEVKGFLAETPSDEELYEFANLKWTCNTNEPRWKTFGII
41 GEINPMSAMT-----L
42 FTEGVDFLNEVPNQYLGPIDFSG GSLPPGIPEEDLFITDLRPFRMGLGAQFG--NQLLL
43 YCCV-----F
44 YKKDELPMIRASP--VLEGG-DKVRWITMASWRDLVI-QQAAA-----TIFRS-LM
45 ESHKEM-----KPIFSRAN-LAWVY-----LNKAREIGPSD-----
46 -----IC-YVS-----DYSSATDTVDR--EFAEFILTNFIKRFRNRLSDP-
47 -LLNF--LELGIRNA-VS---PKVV--I-F-----
48 -PTGE-----RITSSRGVFMGEPM SKVILTLIMFTI----G---
49 KAA-----KSIYKLRFPKSMEKLT FW-----
50 -----APG-----DD--LVAT--GPTEYI--DIYSELAKI-LGQILNHSKVF
51 KS--RTV-FKLCEQW---F-----WVPGLKSSVG--TWAIT
52 TDPGKY-----RES-----
53
54
55
56
57
58
59
60

1
2
3 AWVDTVKL-----
4 -----KLLGAM-----SLSNHSHFEERNEVIGKAKAL-----SKI
5 LRWLPSD-TY-----PLEYKK-----LLRTWF-----LCRFEPKLPRTD
6 S-KTF---AFLMLPGHL--GGFD---LLLDDQ-----EIREAYQ-----
7 -----KVSSFSRILFN----NYYDDFIVG----SYLRNMLRNRSYRGFALSED---
8 -----YNLTVE-----KLVKYIDLLEEA-----PYEEYKDSKVE-----
9 -----GLS-FRETDRALFKAGFLTE
10 NRIRDIV-----SRPLAFGSIWG-----RTLKHSPYNT
11 TPISVRLAKLWKNLTDTE---RWKNISLKGEISE-----
12 -----DDFVVLCENKGPKVLKVYRRNDSL-----ED
13 EVLIGLPS-----MKVSLGILPHDG
14 Y-----ISW-----
15 -----
16 -----
17 -----
18 >KU935604.1
19 -----
20 -----MPTIRNELNPRV-----SFFAEKHGK
21 GTFK-----CPLSLRHYLGDSRHLRKCL
22 -----LHPRAGLLKEV
23 -----
24 LG----KSWK-----TFKCFKTSSIVRAHELYCYLDICVCL-----DP
25 EILSIESNRSIGV-KVFNTAC-----FSTSTATSLYK-----EFVENVLILLK---
26 -----CKEGEERWTRQNPF-KRM-----MTWKIFQDAIASVQLSKGK
27 SLERVLSVLTTDN---FP--EPD--GKSIQKKVDFIGI-ISQQPPKG--WKSEE---
28 -----YRSHFLESIEEIANECES---TTVSDCHI-----SVTAA--GSLKKT
29 VREG-GKFA---EMI-----EEVKGFLAETPSNEELYEFANLKWTCTNEPRWKTGII
30 GEINPMSAMT-----L
31 FTEGVDFLNEVPNQYLGPIFGSGSLPPGIPEEDLFITDLRPFRMGLGAQFG--NQLL
32 YCCV-----F
33 YKKDELPMIRASP--VLEGG-DKVRWITMASWRDLVI-QQAAA-----TIFRS-LM
34 ESHKEM-----KPIFSRAN-LAWVY-----LNKAREIGPSD-----
35 -----IC-YVS-----DYSSATDTVDR--EFAEFILTNFIKRFRNRLSDP-
36 -LLNF--LELGIRNA-VS---PKVV--I-F-----
37 -PTGE-----RITSSRGVFMGEPMISKVILTLIMFTI---G---
38 KAA-----KSIYKLRFPKSMEKLTW-----
39 -----APG-----DD--LVAT--GPTEYI--DIYSELAKI-LGQILNHSKVF
40 KS--RTV-FKLCEQW---F-----WVPGLKSSVG--TWAIT
41 TDPGKY-----RES-----
42 AWVDTVKL-----
43 -----KLLGAM-----SLSNHSHFEERNEVIGKAKAL-----SKI
44 LRWLPSD-TY-----PLEYKK-----LLRTWF-----LCRFEPLPRTD
45 S-KTF---AFLMLPGHL--GGFD---LLLDDQ-----EIREAYQ-----
46 -----KVSSFSRILFN----NYYDDFIVG----SYLRNMLRNRSYRGFALSED---
47 -----YNLTVE-----KLVKYYTDLLEEA-----PYEEYKDSKVE-----
48 -----GLS-FRETDRALFKAGFLTE
49 NRIRDIV-----SRPLAFGSIWG-----RTLKHSPYNT
50 TPISVRLAKLWKNLTDTE---RWKNISLKGEISE-----
51 -----DDFVVLCENKGPKVLKVYRRNDSL-----ED
52
53
54
55
56
57
58
59
60

1
2
3 EVLIGLPS-----MKVSLGILPHDG
4 Y-----ISW-----
5 -----
6 >KY628364.1
7 -----
8 -----MPTIRNELNPRV-----SFFAEKHGK
9 GTFK-----CPLSLRHYLGDSRHLRKCL
10 -----LHPRAGLLKEV
11 -----
12 LG-----KSWK-----TFKCFKTSSIVRAHELYLCYLDICVCL-----DP
13 EILSIESNRSGIV-KVFNTAC-----FSTSTATSLYK-----EFVENVLILLK---
14 -----CKEGEERWTRQNP-FKR-----MMTWKIFQDAIASVQLSKGK
15 SLERVLSVLTRN--FP--EPD--GKSIQKKKVDFIGI-ISQQPPKG--WKSEE----
16 -----YRSHFLESIEEIANECES---TTVSDCHI-----SVTAA--GSLKKT
17 VREG-GKFA---EMI-----EEVKGFLAETPSNEELYEFANLK-WTCNTNEPRWKTFGI
18 IGEINPMSAMT-----L
19 FTEGVDFLNEVPNQYLGPIDFSGGSLPPGIPEEDLFITDLRPFRMGLGAQFG--NQLLL
20 YCCV-----F
21 YKKDELPMIRASP--VLEGG-DKVRWITMASWRDLVI-QQAAA-----TIFRS-LM
22 ESHKEM-----KPIFSRAN-LAWVY-----LNKAREIGPS-----
23 -----DICYVS-----DYSSATDTVDR--EFAEFILTNAFKRFRNRLSDP-
24 -LLNF--LELGIRNA-VS---PKVV--I-F-----
25 -PTGE-----RITSSRGVFMGEPMISKVILTLIMFTI----G---
26 KAA-----KSIYKLRFPKSMEKLTFW-----
27 -----APG-----DD--LVAT--GPTEYI--DIYSELAKI-LGQILNHSKVF
28 KS--RTV-FKLCEQW---F-----W--VPGLKS--SVG-T
29 WAITDPGKY-----RES-----
30 AWVDTVKL-----
31 -----KLLGAM-----SLSNHSHFEERNEV--IGKAKAL-----SKI
32 LRWLPSD-----TYPLEYKK-----LLRTWF----LCRFEPKLPRTD
33 S-KTF--AFLM--LPGHL--GGFD--LLLDDQ-----EIREAYQ-----
34 -----KVSSFSRILFN----NYYDDFIVG-----SYLRNML--RNRSYRGFALS-ED-
35 -----YNLTVE-----KLVKYY-----TDLLEEAPY-----
36 -----EEYKDSKVEG-----LS-FRETDRALFKA-GFLT
37 ENRIR-----DIV-----SRPLAFGSIWG-----RTLKHSP
38 YNTTPISVRLAKLWKNLTG--DTERWKNISLKG-----IS
39 EDDF-----VVLCKENKGPKVLKVYRRNDSLE-----DE
40 VLIGLPSM-----KVSLGILPHDG
41 Y-----ISW-----
42 -----
43 >GFLH01017038.1
44 -----
45 -----SIGTLMHGSQTAHLGTRVGSIHWT-----VLPQPHDVG
46 LKLR-----TSSKPFERIS-----FKGEQISSKKVFGEKGRSGL
47 -----RAYLCTTIGIGVKHAEVIVSRKTSEIRR
48 MF----RQWE-----SIIDNF-LLNPKFELRF-----
49 -----QPAWKKVLRAFLGVAP-----FSVQCTTRMYK-----EFVEFLKQEYFEP-
50
51
52
53
54
55
56
57
58
59
60

1
 2
 3 LEGEKISLPKDNL-FKP-----LVEVVRSVLNKDR--KTD
 4 YV-HFAHICTTRN--LA--PGD--AKEQDNSIREFIEL-TTTEFEVS----KL----
 5 ----DLSNLSLAMDRIVSICLN--SYESGGYLGAHM-----SISES--AEYNVP
 6 QSKG-GKAA---AIL----AEAGEILNRIADKDQKVETPFRS-YTHRQGYPLWMCIAD
 7 VGVKDP-----V
 8 EWLS-----PKEMTIVDRPPPREGLDKTTG--EQLFV
 9 AAWI-----A
 10 FNESRSIKVRAVS--VPEPG-GKSRRVTVSQWWLPVI-QQPLS-----HLLSG-VL
 11 SKHPSA-----YSVFRRQD-QAWQS-----LKLASRIKHTQMSGFAW-----
 12 -----LHS-----DFTSATDAVPH--KVAYCIMERLIMR-----
 13 -CFPN--FKPVLRIL--G---EREV--E-----
 14 -VKKD-----TFFTKRGIFMGEFPSKLVVGSLAI----E---
 15 ELC-----YSIEKSISLRRTFTKPSNW-----
 16 -----RFM--HVGGDD--HRCY--GPVSYL--KRLTRLCKQ-FGYIVSPTKHI
 17 MT--RRA-GMYTEKC-----IYYHDR--VINMK
 18 VQQINDNV-----EKS-----
 19 VVWVDISKL-----
 20 -----RLLSPF-----TKPLDGRDDRNIAIGKAKTL----SRS
 21 IQWFPMG--TELE-----TIARLA----IERYKL-----RFSTMIPSS
 22 N-RKTMTAII--LPDSV--GGLG---L-----VYQPEVE-----
 23 -----DFRNLP-EIFW----WALNVIHAGHP---KNYKVRSLQNTWQNFSGYEQLK-
 24 -----YLSDFK-----SQIEDYPNMVGVT-----FNALKEEFNEH-----
 25 -----FNGD-NTEFVDFLREK-NIIQ
 26 -----IHEFMKLIERP-----HKFY
 27 KLMTRSRIKHFNLDNRTR--IARTWEALENLKS--EITDPIPLIKEVEMEVIQGAAKLA
 28 RHDW--FIKLDEPSEAAFISTELSEKQLDDEEYFTM-----LD
 29 GTISMRTI-----
 30 -----PFSEH--ILYNAPELKLHNMKIQN-----
 31 ----->HACV01002003.1
 32 -----
 33 -----ILNIETGIPRDPAGVQDCSEGSSLRLGSP-----SESVRAPLG
 34 PVPSTKSRQWGRGLLVPTLSPDSVLQNPVRPIYEVKWGGVSLSGGITAYGSRKWRQRLR
 35 K
 36 -----YLPTVWGVSVRTAGILSGRSRSYLIQ
 37 VF----RNLD-----SLIDNMLFDRTLFD-----QN
 38 NAESRSVITRFVR-KVLTVSA-----FSIGDCIAKWK----DFVNHFVTVSSDAT
 39 -----VSSTVSLKGNCFK-FLL----NEGKFKRLVNDKPS--TRE
 40 DYMDIAHVISTRN--LA--NGG--VKAQQAAVSQFRKT-TSTEFKVS----PE----
 41 -----NLKKRLVASRLGRIVTS--IRGSNKLRGHC-----SLNGS--GTLDVP
 42 VNQG-GRAT--DALL-----DLKKFLSHIPPSGESRTYPWGEVWF--PSGIPRWCAFGK
 43 LENFTD-----A
 44 EWMS-----LANREIIDRGALRNGCDSTTG--RMVMA
 45 VAYE-----CYRPF
 46 EESGTPIPIRQAA--VSEPG-GKVRMVTGPWWLAVL-QQPLC-----HGLRE-IL
 47 AYHPSA-----HSCLMRCD-QAWQS-----LHVLERLGISKLEDGMAV-----
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 -----LSS-----DLKEATDAIPH--SVGKVLLSGFLEG-LG-----
4 -REDH---DWITNF-IG---ERTV---F-----
5 CEDGE-----IFTLRRGVMMGEPLSKICLVLGLAM----E---
6 ELA---F-----SEHEGLSLHRRDTPRCHW-----
7 -----RAF--HLGGDD--HLAV--GPVTYL--RSITSNHKL-FGSVISPVKHR
8 MS--SIY-VVYTEKV-----LRIKGAVLNLKAS
9 QVDGNP-----GNC-----
10 TFIDSVKV-----
11 -----RLLSPF-----SKATDSQNDKNIAIGKVRGI-----AAT
12 LKYFPSE-----SIKRTV---FDRVAY-----RFNHYIGSE
13 R-HRTIRAVMS--LPSEL--GGLG---F-----AFDDR-----
14 -----YLGRLP-PIFN---RAVRTVINGGG---VGFRVRRILGQIFANRCPRGLVSN
15 ALMVDDYVEQMI-----EQFLDYPGSVDMV-----SSGDLWARLDE-----
16 -----NGLMS-FREFLEKASHE-GWIS
17 -----LHDIPRIAEP-----FLMRKILE
18 GSEVPEYFRTASNAQRIA-----CWNSLEAIDDVLD-----PTGGE
19 LSPL-----EIREASRRAKHTVFVNLNEMTCAA-----LVD
20 PTSEAYDP-----
21 -----GDPWISADFIQLSMRELLTYGEPSLTVP-----
22 -----
23 >GFBP01062652.1
24 -----
25 -----
26 -----
27 -----
28 -----LTLRRYMPPTQLTAPPGRFEGGV-----HSPGLLVPN
29 LTPD-----HVVVN---PVRPI-----WEHRGFGVCFASLGSKQKWR
30 -----TSFARTLRVELD-----MGRRASETLANRGLSYIKR
31 VL-----KNLD-----SLIDNLILFSPSMFDR-----ST
32 PGVIYDTLKRILR-VTLKVSTY-----SVEDQISQWK-----DFVNHFVTLVSKSQ
33 VEKQICL-----RHNCYKWMLEQDP-FKS-----LLSKIDMTTKDW-----
34 -----SIAHLISTRN---FA--PGG--ESAKLKAIRAFKET-TTS--GFS---VSEE-----
35 -----NLRKLKLATLR---VTGI--SLHLSEEKRLSGG-----HVSLNSS--GTLDVS
36 VAKG-GRAA---DAL-----LDLKQFLDEEPVQGETRVYPWGE-LYYPPGLRRWRAFGP
37 LQD-----D
38 DGPF-----MSKSRYDLIDREAMLNGCGEYTG--VMIFT
39 VAYE-----CYRVYLNNSGE-
40 -----LIPIRQAT--VSEPG-GKARIVTGPWWLAVL-QQPLC-----HGLRE-LI
41 GYHPSA-----HSCLLRAD-QAWQS-----LHWAKCSDNCIPEGEAV-----
42 -----LSS-----DLKEATDAIPH--SVGQVLIRGCLEA-IG-----
43 -ATKW---LWMADL-IG---PRLV---L-----
44 -AEDK-----DVFTLRRGVMMGEPLSKICLILLGLTM----E---
45 EIA---F-S-----EYSSLSSLKRKFTPAQGW-----
46 -----RAF--HLGGDD--HLAK--GPMRYL--KSITSYHRL-FGSIVSPYKHR
47 IS--VRA-VTYTEKI---L-----V--FENRIL--NMPVE
48 QVNSNI-----DKS-----
49 IFIDSIKV-----
50 -----RLLSPF-----TKATDTMNDKNVAIGKCIGI-----AGT
51 LQWMHDR-----SLKRTL-----IDRIHY-----RFRDFIAGP
52
53
54
55
56
57
58
59
60

1
2
3 H-HRTLSAIQS--LPTTL--GGLG---L-----SVDTK-----
4 ----YLENLPEPYNRAIREIKGGLAYKARMVMNSTFVNKNLRGIKSDDYVTELITQ-
5 -----WVEYPS-----SVNAMNFKEAIAVV-----DPEGNNNSFRHN-----
6 -----LWLKKRN-IVSMNDIPQIAESAYV
7 -----MRRLLEQSQSGKYFRTERLRDRVSKTWEGLTKLDL
8 PESGDVLSEDELKTGTRLA---KQALFIDLTTMGT-----JAVA
9 DSTN-----PDYLPDDPWLAEEFIDIPFGQSIK-----YG
10 QPNMGIDL-----
11 -----DHPFP--EEKDLAC-----
12 -----
13 -----
14 >JN400242.1
15 -----
16 -----
17 -----RS-----TLPQSWERG
18 D-----RYHSDLLKRDLMIRLTRDLG
19 -----LSGRHADRLVNRPAIHLKR
20 IE----EFIS-----GLVDSFWLGDPRTFAL-----GS
21 PD--FKTFRTIVR-KIFAVGT-----SNLGMLMANWK-----EWTNWFTHSVCETG
22 LREEAKLS-----NHNMFHRLARLSV-VQK-----ALSGGVPLMELGS-----
23 ---IVAHLTSTRQ--MP--YMG--LPTEVKAKADFIEI-IST--PFH---VEPS-----
24 ----HHRQMVQCAARMGRLCMS--LRNGVPVSDRSSHF-----SATSS--GELDHS
25 LTRG-GQAQ---ALK-----DAIDRWLIPASGSTYMEDTPFGV-AEHREGVPLWKTLFV
26 DQE-----TQLELIFSDFGDS-----L
27 DWIK-----DVPGRVYGLDDYTG--RQILY
28 VAWK-----EMED-
29 -----IPHIRAST--VPELG-NKARIVTLSAFWLNVL-QAPLS-----HIMKE-VL
30 KYHPSC-----FASFTRGE-QAWHA-----ASGLGRLNPRSVAG-----
31 -----YSVLSS-----DLKNATNAQHI--ALTRDMLRAFIYS-SG-----
32 -LVVSKAYVNLVLDT-IC---PRLV-E-----
33 -LDGE-----FTLSCRGIMMGEAIAKPSMTLLNLVV---E---
34 GLA---F-----LKYENKLVLLNTDKAAPSRRW-----
35 -----RCY--HVGGDD--HLAV--GPDQYL--DLITHNHEL-SGSIISPDKHA
36 KS--RKM-VKYCERV----L-----L--VQNLQY--NTP-D
37 RKPF-----EHG-----
38 LIVDSIKV-----
39 -----RLLEKG-----QSTLIAKDNKN-VA--VGKSQQL-----VKS
40 LDWLPKQ---LYSKGFIR-----SIQHLF-----IKRMGS-----LLPNRNR-D
41 E-EAFHSVC---LPKIL--GGYG--LG-----LAEDLLT-----
42 -----HLSKSK-PEIQ----QTVSAVLYGIETRRARNILSRL--NTTISDRSIQG-TK-
43 -----WYEDEM-----VQQFREY-----PDMVGAIDG-----
44 -----
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----REMRR-----
51 --RF-----
52 -----
53 -----
54 -----PAST-----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 >JN400243.1
4 -----
5 RS-----TLPQSWERG
6 D-----RYHSDLKRDLMIRLTRDLG
7 -----LSGRHADRLVNRPAIHLKR
8 IE-----EFIS-----GLVDSFWLGDPRTF-----AL
9 GSPDFKTFRTIVR-KIFAVGT-----SNLGMLMANWK-----EWTNWFTHSVCETE
10 LREEAKLS-----NHNMFHRLARLSV-VQK-----ALSGGVPLMELGS-----
11 ---IVAHLTSTRQ---MP--YMG--LPTEVKAKADFIEI-ISTPFHVE----PS----
12 ----HHRQMVQCAARMGRLCMS--LRNGVPVSDRSSHF-----SATSS--GELDHS
13 LTRG-GQAQ---ALK-----DAIDRWLIPASGSTYMEDTPFGV-AEHREGVPLWKTLFV
14 DQETQLELIFSDFGDS-----L
15 DWIK-----DVPGRVYGLDDYTG--RQILY
16 VAWK-----
17 -EMEDIPHIRAST--VPELG-NKARIVTLSAFWLNVL-QAPLS-----HIMKE-VL
18 KYHPSC-----FASFTRGD-QAWHA-----ASGLGRLNPRSVAG-----
19 ----YSVLSS-----DLKNATNAQHI--ALTRDMLRAFIYS-SG-----
20 -LVVSKAYVNLVLDT-IC---PRLV-E-----
21 -LDGE-----FTLSCRGIMMGEAIAKPSMTLLNLVV----E---
22 ELA-----FLKYENKLVLLNTDKAAPSRRW-----
23 -----RCY--HVGGDD--HLAV--GPDQYL--DLITHNHEL-SGSIISPDKHA
24 KS--RKM-VKYCERV---L-----L--VQLNQY--NTPDR
25 KPF-----EHG-----
26 LIVDSIKV-----
27 -----RLLEKG-----QSTLIAKDNKNVA--VGKSQQL-----VKS
28 LDWLPKQ---LYSKGFIR-----SIQHLF-----IKRMGS-----LLPNRNR-D
29 E-EAFHSVC---LPKIL--GGYGLGLA-----EDLLT-----
30 ----HLSKSK-PEIQ----QTVSAVLYGIETRRARNILSRL--NTTISDRSIQG-TK-
31 ----WYEDEM-----VQQFREY-----PDMVGAIDG-----
32 -----
33 -----
34 -----REMRR-----
35 --RF-----
36 -----
37 -----
38 -----
39 -----PAST-----
40 -----
41 -----
42 -----
43 -----
44 -----
45 >JN400241.1
46 -----
47 -----MQEDLILVQPDTDWVGGVNRSWRS-----LLPVNGVE
48 YILP-----RG--SLVGL-----KLHGRSTLPQSWERGDRYHS
49 -----DLLKRDLMIRL-----TRDLGLSGRHADRLVNRPAIHLKR
50 IE-----EFIS-----GLVDSFWLGDPRTFAL-----GS
51 PD--FKTFRTIVR-KIFAVGT-----SNLGMLMTNWK-----EWTNWFTHSVCETE
52 LREEAKLS-----NHNMFHRLARLSV-VQK-----ALSGDVPLMELGS-----
53 ---IVAHLTSTRQ---MP--YMG--LPTEVKAKADFIEI-IST--PFH---VEPS----
54 ----HHRQMVQCAARMGRLCMS--LRNGVPVSDRSSHF-----SATSS--GELDHS
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 LTRG-GQAQ---ALK----DAIDKWLI~~P~~ASGSTYMEDTPFGV-AEHREGVPLWKTLFV
4 DQE-----TQLELIFSDFGDS-----L
5 DWIK-----DVPGRVYGLDDYTG--RQILY
6 VAWK-----EMED-
7 -----IPHIRAST--VPELG-NKARI~~V~~LSAFWLNVL-QAPLS-----HIMKE-VL
8 KYHPSC-----FASFTRGD-QAWHA-----ASGLGRLNPRSVAG-----
9 -----YSVLSS-----DLKNATNAQHI--ALTRDMLRAFIYS-SG-----
10 -LVVSEAYVNLVLD~~T~~-IC---PRLV-E-----
11 -LDGE-----FTLSCRGIMMGEAI~~A~~KPSLTLLNLVV---E---
12 ELA---F-L-----KYENKLVLLNTDKAAPSRRW-----
13 -----RCY--HVGGDD--HLAV--GPDQYL--DLITHNHEL-SGSIISPDKG
14 KS--RKM-VKYCERV---L-----L--VQNLQY--NTPDR
15 KPF-----EHG-----
16 LIVDSIKV-----
17 -----RLLEKG-----QSTLIAKDNKNVA--VGKSQQL-----VKS
18 LDWLPKQ---LYSKGFIR-----SIQHLF-----IKRMGS-----LLPNRNR-D
19 E-EAFHSVC---LPKIL--GGYGLGLA-----EDLLT-----
20 -----HLSKSK-PEIQ-----QTVSAVLYGIETRRARNILSRL--NTTISDRSIQG-TK-
21 -----WYEDEM-----VQQFREYPDMVGAI-----DGREMRRRFPA-----
22 -----ST-YQESKAHAASA-GWLP
23 -----IDQLARRVTRG-----HLFGELMSGAKRQNVFK
24 TRNWIQE~~F~~TRLRKLEEEV--GYLPLPEYADLEV~~K~~TLRNALGA-----MT
25 TDLF-----IDSQVTSFDKGVGDDSFDFDGE-----LL
26 RTYEAHQP-----NLVVGLQFIG
27 LDSTC-----NAFTR--RVRRRY-----
28 -----
29 >KU295726.1
30 -----
31 -----MQEDLILVQPDTDWVGEVNRSWRS-----LLPVNGVE
32 YILP-----RG--SLVGL-----KLHGRSTLPQSWERGDRYHS
33 -----DLLKRDLMIRL-----TRDLGLSGRHADRLVNRPAIHLKR
34 IE----EFIS-----GLVDSFWLGD~~P~~RTFVL-----GS
35 PD--FKTFRTIVR-KIFAVGT-----SNLGMMLMTNW~~K~~-----EW~~T~~NWFTHSVCETE
36 LREEAKLS-----NHNVFHRLARLSV-VQK-----ALSGGVPLMELGS-----
37 ---IVAHLTSTRQ---MP--YMG--LPTEVKAKADFIELIST--PFH--VEPS---
38 -----HHRQM~~V~~QCAARMGRLCMS--LRNGVPVSDRSSHF-----SATSS--GELDHS
39 LTRG-GQAQ---ALK----DAIDRWLI~~P~~ASGSTYMEDTPFGV-AEHREGVPLWKTLFV
40 DQE-----TQLELIFSDFGDS-----L
41 DWIK-----DVPGRVYGLDDYTG--RQILY
42 VAWK-----EMED-
43 -----IPHIRAST--VPELG-NKARI~~V~~LSAFWLNVL-QAPLS-----HIMKE-VL
44 KYHPSC-----FASFTRGD-QAWHA-----ASGLGRLNPRSVAG-----
45 -----YSVLSS-----DLKNATNAQHI--ALTRDMLRAFIYS-SG-----
46 -LVVSEAYVNLVLD~~T~~-IC---PRLV-E-----
47 -LDGE-----FTLSCRGIMMGEAI~~A~~KPSLTLLNLVV---E---
48 ELA---F-L-----KYENKLVLLNTDKAAPSRRW-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----RCY--HVGGDD--HLAV--GPDQYL--DLITHNHEL-SGSIISPDKHG
4 KS--RKM-VKYCERV----L-----L--VQNLQY--NTPDR
5 KPF-----EHG-----
6 LIVDSIKV-----
7 -----RLLEKG-----QSTLIAKDNKNVA--VGKSQQL-----VKS
8 LDWLPKQ---LYSKGFIR----SIQHLF-----IKRMGS-----LLPNRNR-D
9 E-EAFHSVC---LPKIL--GGYGLGLA-----EDLLT-----
10 ----HLSKSK-PEIQ----QTVSAVLYGIETRRARNILSRL--NTTISDRSIQG-TK-
11 ----WYEDEM----VQQFREYPDMVGAI-----DGREMRRRFPA-----
12 -----ST-YQESKAHAASA-GWLP
13 -----IDQLARRVTRG-----HLFGELMSGGAKRQNVFK
14 TRNWIQEFTRLRKREEEV--GYLPLPEYADLEVKTLRNALGA-----MT
15 TDLF-----IDSQRVTSFDKGVGDDSFDFDGE-----LL
16 RTYEAHQP-----NLVVGLQFIG
17 LDSTS-----NAFTR--RVGRGY-----
18 ----->KU295722.1
19 -----
20 -----MQEDLILVQPDTDWVGGVNRSWRS-----LLPVNGVE
21 YILP-----RGSLVGLKLHGRSTLPQSWERGDRYHSDLKRDLMIRLTRDLG
22 -----LSGRHADRLVNRPAIHLKR
23 IE---EFIS-----GLVDSFWLGDPRTF-----VL
24 GSPDFKTFRTIVR-KIFAVGT----SNLGMLMTNWK-----EWTNWFTHSVCETE
25 LREEAKLS-----NHNMFHRLARLSV-VQK-----ALSGGVPLMELGS-----
26 ---IVAHLTSTRQ---MP--YMG--LPTEVKAKADFIEI-ISTPFHVE----PS-----
27 ----HHRQMVKCAARMGRLCMS--LRNGVPVSDRSSHF-----SATSS--GELDHS
28 LTRG-GQAQ---ALK----DAIDRWLIPASGSTYMEDTPFGV-AEHREGVPLWKTLFV
29 DQETQLELIFSDFGDS-----L
30 DWIK-----DVPGRVYGLDDYTG--RQILY
31 VAWK-----
32 -EMEDIPHIRAST--VPELG-NKARIVTLSAFWLNVL-QAPLS-----HIMKE-VL
33 KYHPSC-----FASFTRGD-QAWHA-----ASGLGRLNPRSVAG-----
34 ----YSVLSS-----DLKNATNAQHI--ALTRDMLRAFIYS-SG-----
35 -LVVSEAYVNLVLDT-IC---PRLV-E-----
36 -LDGE-----FTLSCRGIMMGEAIAKPSLTLNLVV----E--
37 ELA---F-L-----KYENKLVLLNTDKAAPSRRW-----
38 -----RCY--HVGGDD--HLAV--GPDQYL--DLITHNHEL-SGSIISPDKHG
39 KS--RKM-VKYCERV----L-----L--VQNLQY--NTPDR
40 KPF-----EHG-----
41 LIVDSIKV-----
42 -----RLLEKG-----QSTLIAKDNKNVA--VGKSQQL-----VKS
43 LDWLPKQ---LYSKGFIR----SIQHLF-----IKRMGS-----LLPNRNR-D
44 E-EAFHSVC---LPKIL--GGYGLGLA-----EDLLT-----
45 ----HLSKSK-PEIQ----QTVSAVLYGIETRRARNILSRL--NTTISDRSIQG-TK-
46 ----WYEDEM----VQQFREYPDMVGAI-----DGREMRRRFPA-----
47 -----ST-YQESKAHAASA-GWLP
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----IDQLARRVTRG-----HLFGELMSGGAKRQNVFK
4 TRNWIQEFTRLRKRLEEEV--GYLPLPEYADLEVKTLRNALGA-----MT
5 TDLF-----IDSQVTSFDKGVGDDSFDFDGE-----LL
6 RTYEAHQP-----NLVVGLQFIG
7 LDSTC-----NAFTR--RVRGRY-----
8 -----
9 ----->KU295727.1
10 -----
11 -----
12 -----MQEDLILVQPDTDWVGGVNRSWRS-----LLPV LNGVE
13 YILP-----RG---SLVGL-----KLHGRSTLPQSWERGDRYHS
14 -----DLLKRDLMIRL-----TRDLGLSGRHADRLVNRPAIHLKR
15 IE-----EFIS-----GLVDSFWLGDPRTFVL-----GS
16 PD--FKTFRTIVR-KIFAVGT-----SNLGMLMANWK-----EWTNWFTHSVCETE
17 LREEAKLS-----NHNMFHRLARLSV-VQK-----ALSGGVPLMELGS-----
18 ---IVAHLTSTRQ---MP--YMG--LPTEVKAKADFIEI-IST--PFH---VEPS-----
19 -----HHRQMVQCAARMGRLCMS--LRNGVPVSDRSS-----HFSATSS--GELDHS
20 LTRG-GQAQ---ALK-----DAIDRWLIPASGSTYMEDTPFGV-AEHREGVPLWKTLFV
21 DQE-----TQLELIFSDFGDS-----L
22 DWIK-----DVPGRVYGLDDYTG--RQILY
23 VAWK-----EMED-
24 -----IPHIRAST--VPELG-NKARIVTLSAFWLNVL-QAPLS-----HIMKE-VL
25 KYHPSC-----FASFTRGD-QAWHA-----ASGLGRLNPRSVAG-----
26 -----YSVLSS-----DLKNATNAQHI--ALTRDMLRAFIYS-SG-----
27 -LVVSEAYVNVLVLDT-IC---PRLV-E-----
28 -LDGE-----FTLSCRGIMMGEAIAKPSLTLNLVV-----E---
29 ELA---F-L-----KYENKLVLLNTDKAAPSRRW-----
30 -----RCY--HVGGDD--HLAV--GPDQYL--DLITHNHEL-SGSIISPDKHG
31 KS--RKM-VKYCERV---L-----L--VQNLQY--NTPDR
32 KPF-----EHG-----
33 LIVDSIKV-----
34 -----RLLEKG-----QSTLIAKDNKNVA--VGKSSQL-----VKS
35 LDWLPKQ---LYSKGFIR-----SIQHLF-----IKRMGS-----LLPNRNR-D
36 E-EAFHSVC---LPKIL--GGYGLGLA-----EDLLT-----
37 -----HLSKSK-PEIQ-----QTVSAVLYGIETRRARNILSRL--NTTISDRSIQG-TK-
38 -----WYEDEM-----VQQFREYPDMVGAI-----DGREMRRRFPA-----
39 -----ST-YQESKAHAASA-GWLP-----
40 -----IDQLARRVTRG-----HLFGELMSGGAKRQNVFK
41 TRNWIQEFTRLRKRLEEEV--GYLPLPEYADLEV-----KTLRNALGAMT
42 TDLF-----IDSQVTSFDKGVGDDSFDFDGE-----LL
43 RTYEAHQP-----NLVVGLQFIG
44 LDSTC-----NAFTR--RVRGRY-----
45 -----
46 ----->HAFC01092190.1
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----MRELSYKLIRVTIPRGATGKSVSEAYRA-----ILPCLLPTD
55 QEDV-----NYTLPPLGIFKL-----KFKGQSPLPQLWEKADSVYL
56
57
58
59
60

1
2
3 SEREKKIFRFHL-----KTGLGLGKRVTEVFLSRPRSHYLR
4 TF-----GLVDAFWAADPKFF-----LV
5 GSKDYKIFKSIVR-KIIKVGC-----TDLGRLVDDWK-----DWGTRIFHQVARTK
6 TIGSLPEAK-----PENIFRILSELSW-IKS-----ILAKDVKDLDSDLG--N--
7 ---LLASIRSSRQ---FP--YMG--AKTEKKSLERFKEV-VST--PFI---VDPE----
8 ----TSWQMVTARRIGGICRA---IRTRPIPDMAAHF-----SVSSS--GEYTYT
9 INKG-AFSQ---AVF-----DGIVPYLMRIIDTSSEETPFGLV--HYQSGKEVWRYLF
10 DEEISQGEFLDNR-----K
11 GFPK-----CQLPRLGLDETFG--KQVLW
12 AAWK-----ASLEFQGK-
13 ----PIPVRAEI--VPEMG-NKARVVTVCPWWLQIL-QAPLA-----HACIA-MM
14 KWHPSV-----YSSFHKMD-QAWCA-----SVGLHKVKDPDLTKY-----
15 ----WL-LSS-----DLKDATNAQSI--DLTKAMLKAFLLG-SNLLSGRE
16 AYVNY--CLELLGSTRLG---SREV--H-F-----
17 -KDKT-----VVLSKSGIMMGEPPLAKVGLTLLNLSV----E---
18 ELA---L-----LTHIGRLDLLLKPDPTPKLKW-----
19 -----RFC--HIGGDD--HLAW--GPEPYL--EKITELHYK-VGSHLSPGEHG
20 HS--RIF-VRYTEKV---L-----NVMNMGKSPMDAQVNKI
21 DAFKR-----
22 IMVDGIKI-----
23 RLLESG-----QSTMISKDNKNVA--IGKAGQL-----VRS
24 IQWLEPK-IGFTESRIDSIR----DLFCLR----LNRFLP-----SQHRNWRCYQ
25 A-TK-----LPRVLL--GGFGLGIT-----PSEQYQA-----
26 -----YVDSP-SPIR----WVVKWKAGEPIVDELQILSFL--NSNVSTRGIASVEA-
27 -----YRKEIL-----DQLQTYPNLVNAK-----SWHDICELPV-----
28 -----EQDAP-SSMIVDKAAEAGYLT-
29 -----FEKFAEMSTRG-----AIFQ
30 NLLLATEEVKIFQTRPYIQ---TMRRVTHQLQLRG-----VDQFGDTPIE
31 QADF-----EEMLKEIGRPIFFNIDQVTTCDIG-----EYD
32 PLKPEEET-----
33 -----WNFQDTTFRELYTKGTPSLIVGLSFLGLRY-----
34 ----->NC_035126.1
35 -----
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 MVS-----GLVDALWCADEFVF-----VL
47 GSIDLKTLKALVR-KIFCVGT-----FNLGLLVDDYK-----AWANRLFHKAASTQ
48 TIGQMEPLR-----CNNIFKMLDSIEY-IKA-----IQDSGSDCSFINLQ-----
49 ---KLSHLVSSRQ---LP--YMG--LKTEIKSLKGFKDV-VTSDYRPT----LE----
50 ----TLDLLTKAARRIGGICRS--IRKDRPIPDAEAHI-----SMTCS--GEATTS
51 IIDG-GQAC---AVR-----EGFERIMGRIAESDLLEDTPFGV-ASHREGYALWRTIFR
52 EEPVEGGSLYDPLY-----Q
53 GYPK-----DQPGRYLGLDRVLG--KQLLY
54 VAWK-----ESILNP-
55
56
57
58
59
60

1
2
3 ----TIVLRASI--VPEMG-NKARVVTMSPYWVQLL-QAPLA-----HLTIA-GM
4 RLHPSV-----FSSFARQD-QAWEA-----AKGISRLNIDNWEDL-----
5 ----HV-LSS-----DLKDATNAQQQ--LLTRTMLRAFMNG-YG-----
6 -MIHRNTYSELALST-IV---ERLI--T-F-----
7 -EDQT-----SVWATTGIMMGEPIAKPSLTLLNLAI----E---
8 ELA---F-----LTYTNRLDLLSDLPSPYRAW-----
9 -----RFI--HIGGDD--HLAI--GPKKYL--ELITNYHLL-AGSHIDPGKHG
10 YS--QVA-VKYTERI-----LSVKNFRFKAPFD
11 YDNY-----PKS-----
12 MIVDSVKV-----
13 -----RLLERG-----QSTMQKKDNKNVA--IGKSRRMM----SGC
14 LDWLPSD--PAYW-----PSDKKI----SIRNLF----INRMGPLLPSKS
15 L-HPKCYHSVQ--LPSIL--GGFN---L-----GLKSEIY-----
16 ----QAYTLAPPIR----WVLNKAAGVDVTaelkilsll--NRNISDRGTDPLQE-
17 -----YRASII-----DQLYSYPQLIGAI-----TWKEILSKFPP-----
18 -----QNENV-RWSIYAAREAH-----
19 -----YLSIEEFAEWSTRGNTFVQLIMGRGKNNVFNTRPY
20 VRTFARVWTELEALQTDLYGDSLGEGEFNKTLRG-----LSRCLYFDTSQSTTID
21 IGYY-----KPDDPDREEFEFVDAITLEAFDRM-----LP
22 TLVVGKKF-----
23 -----IGFRN-----
24 -----
25 >HACV01003050.1
26 -----
27 -----
28 -----
29 -----
30 -----AECMNFCGKLIQAQHLRDCVGKRRNGSP-----WDKLLPVLG
31 NPNGQQS--NTLLY---YHMPSG--SLLSQ-----TFKGRSLRKQKFEK-ESKYF
32 -----SPRTKYGLFVLL-----TRDIGVTQIAKILVDRPVSHFKR
33 IQ-----EFIN-----GLVDSLWLANEQVFLE-----CS
34 QE--LLIVKKTIR-KLFCVGT-----FNLKDLVDQWK-----EWGNYLFLHTVARTE
35 TIGPLKAVA-----KNNIFRTLNGIPY-ISR-----MYHTN--EPDMLLL-----
36 --QHLSHLVSSRQ--MP--YMG--FKTELDSSRKKFKEV-LLS-SYE---PPKD-----
37 -----FLQKISMAARRIGGICRT---LRP-K-LNSGE-----THISVTSS--GTITHS
38 IEQG-GQAA---AVM-----EAMKRILLVVPEESFEEDTPFGT-AVHKKGIPWKLYR
39 SSN-----DDMSSYEFLEPYNL-----I
40 KEMS-----G---RFRGLDRVHG--AQIMY
41 VAWRELEATPILRAEVVPEMGNKARHVTLSAYWLNVLQSPLCHELLIDS MKYHWRELEA
42 T-
43 -----PI-LRAEV--VPEMG-NKARHVTLSAYWLNVL-QSPLC-----HLLID-SM
44 KYHPSV-----FSSFNRQD-QAFE-----VKGMCLKELQLLPGHAV-----
45 -----LSS-----DLKDATNAQQW--SVTKAILKGYISG-ANLS----
46 -FSDR--YINLVLDL-IG---PRLV--K-F-----
47 -CDGT-----TVLSQVGIMMGEAIKPSLTLLNLSI----E---
48 ELA---F-----LEYNNATELLTSNDPAPYKQW-----
49 -----RFL--HIGGDD--HLAI--GPVCYL--NQITHNHML-AGSHIDPGKHG
50 FS--QIC-VRYTERL-----INIRNFQYKQPFH
51 RDDYSKS-----
52
53
54
55
56
57
58
59
60

1
2
3 IIVDSVKV-----
4 -----RLLERG-----LSTEMKRDNKNA-AIGKSQTL-----GGT
5 LEWLPID---DRFWT-----ETKKES-----IRALFV---ERMGELLPRKAV-N
6 P-RAYAAIH---LPQYF--GGYG-----
7 -----LGFKRE-----
8 -----WGKYLR-----DSPAPHRGLVM-----
9 -----
10 -----
11 -----LAHLGVDRVHH-----
12 -----LKVFRKLNSNLS-----
13 -----
14 -----
15 -----
16 -----
17 -----
18 >GACM01002912.1
19 -----
20 -----NFTGNLIQAHLLSECAGKDRKRSPWDR-----LLPVLGPNP
21 KVDVNRQFYLMPSGSLLSETYKGSSLMRQKFH-----KRERFFNPRQKAGLRIILTR
22 -----DIGLSQRVAVQLVNRPIVHYKR
23 IE----EFIS-----GFVDSLWLAEEQIF-----LF
24 DGKWIKLIRKLIR-KIFCVGT-----YNLGSVLVDQWK-----EWGNNLFHTLAEST
25 TIGVLRKPA-----ENNIFRRLLDTIDY-ISA-----VYRGDKSMKLMQ-----
26 ---HVSHLISSRQ---MP--YMG--SSTEKAWDKFQSV-LTSDFVPS---EK---
27 ----TIVQLQMAARRVGGICRS--IRSRRIPDGVAHI-----SVTSS--GEYSYP
28 IAKG-AQAA---AVK-----AAMERILTVVPLEDQEEDTPFGL-VRHHKGIPIWKTLFR
29 TEPLDTE-----L
30 PFLA-----PYALIKEQEGRFAGLDRVVTG--KQIMY
31 VAWK-----EYRP-
32 -----LPVLRAEV--VPEMG-NKARLVTISDYWLNIL-QSPLS-----HVLID-AM
33 KFHPSV-----FSSFHRQD-QAFEA-----TKGLCSLKRKSLLKGEAV-----
34 -----LSS-----DLQDATNAQQW--SVTIAMLRGFIQG-FGLS---
35 -FRPE--YVELVLST-IG---PRLV--C-F-----
36 -RDEI-----SILSKVGIMMGEAIAKPSLTLLNLSI-----E---
37 ELA---F-L-----RHCRAEELLYSVEPAPYRDW-----
38 -----RYI--HIGGDD--HLVK--GPIPYL--NLITQIHL--AGSHIDPGKHG
39 FS--RIC-VKYTERL---L-----N--LSNLEH--GCP-F
40 DPSDY-----SRS-----
41 TIVDSVKV-----
42 -----RLLERG-----QSTLLKKDNKNVA--IGKSQL-----GGC
43 LEWLPID--DRFYTETKK-----ASIRAL-----FIERMG-----SLLPRKAV-N
44 P-RAFAAIH---LPTIV--GGYG-----LGMSS-----
45 -----ELQKFLEASPE-----PHKGLLMKAFCGVNVKEDLKIF--RVLNTNTSDRG-VE-
46 -----NIQQFQ-----QKII---DQLSEY-----PQMVNAMDW-----
47 -----WELKRKFDPDS-----GN-SKRTIALAADA-GILS
48 -----FEEFAKRATRG-----NLFQALLMGRKDLKVFN
49 TSPFVRTYKNIVWDEAESR---GLLTWSEGFLNNDEIAIAIK-----NI
50 APQW-----YFDINQETAMDTGHWDOPENPETET-----WD
51
52
53
54
55
56
57
58
59
60

1
2
3 FKDSTYID-----KYTEGFPSFD
4 V-----GFKVL--GLRH-----
5 -----
6 >GFCJ01029165.1
7 -----
8 -----
9 -----
10 -----SHLISTRHFPTGGKSVE
11 -----
12 -----RQ
13 AVSDFVEVTTTLK-----NVDSDFLKKVR-----SASTIIGARL-----
14 -----KENSRPNLQTHFS-----
15 -----LSASGSLNKT--IS--EGGR-GQEVKAEAIENLKFI-----PIE-----
16 -----DKEIKILG-----WTVN--DIAGRPRWQT-----WFRDP-----
17 -----DWEN-----PEPDVPFGELLKDLLGYPLYRQ-----
18 -----
19 -----
20 -----GFDEVLG--KQIMA
21 LAAN-----TFLEE
22 VSAGRPSVRVIT--VPEPG-GKARIVTGPWWLYIL-EQPFS-----HAMRS-LL
23 VHHPSA-----IAGLCRAD-QAWQY-----LRLLSNAKKYLGKSN-----
24 -----YC-LSS-----DLKSATDAIPF--EICEAVFEGFFDG-IGQ-----
25 -SPTM--SPVIAKLLTCG---RECI---L-----
26 -KSGF-----SYTTVRGIFMGEPMMAKPILTQLVNLVY----E---
27 EIALRDF-----LKVKQKPISVPW-----
28 -----RCF--AVAGDD--HIAV--GPLPYL--KMITENHLK-GGSIIISPDKHG
29 LS--RVL-VRYCEKL-----IDLTRLKTDWTME
30 SINKDTLGY-----RNG-----
31 PFVDCVKV-----
32 -----RLISPA-----SKGNESFNEKNTAIGKGLTL-----GAT
33 LRWLKD-HFPTKWV-----NMVRDR-----FFNRMG-----TLLPSRES--
34 -----GVFWHLLLPPAL--GGLG--LWIEDDVPLNNNRLPQPSRALVR-----
35 -----DIDANH-PDLK-----KNLKLFKELTSNSTFRGFKLGD--EPFEQCIDTAKEIL-
36 -----QSTKCT-----LRDLETEFPVLKGL-----NPRGQIAFLRT-----
37 -----KE-YVTEEEELTKSIKRPF
38 -----FNQILAGTSKR-----KPFN
39 TEALKHRYAKIWESLYKIEPNFELDEVVVRKALKY-----
40 -REY-----QDLYSIQERELIPFRNSEVFVNA-----IE
41 EITLGLPA-----
42 -----FSVEWSFAGVASDTQDQSTKNQPT-----
43 -----
44 >GFIK01043228.1
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----RPFFCFNMSSFIACDLRAPAGRKHGYRL-----YIPRLFENG
51 DYIL-----PAKYDSV LNGCFL-----WKRKNLNTNVGRGLVIQVLCR
52 -----DFGFSKRFAEVFASNISEIRR
53 IE-----QTIH-----AVLDCLLLYDTKLF-----I
54 ENRYYSLLKHIRV-RIMSVGT-----FSIGLVISYWK-----DFTNDIWMTVSEAE
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 YQGEISID-----KRNFFYEIERVLC-LKS-----LKGEKLDKVQLQ-----
4 ---RFAHLTSTRG---LP--CGD--KQSEKRAILKFKEH-VTTPTQVS-----DY-----
5 -----VVEDYGRAAGLIARKCAR--RFEFGVRAAHI-----SLATA--GSLDTT
6 VSEG-GRGE---EIR-----RAIFPMLTFIPKESSIVKTPIFD-LNEVAGVERWRTWCR
7 PEPY-----M
8 DFPD-----KAFGTPNGETLGGCEVFIQGFDEAIG--EQILA
9 CAYF-----AYQQDKGDQV-
10 -----EILTRVLT--IEEPG-GKSIVTTAPWYVYVL-EQPVA-----HIMRL-VV
11 SNHPSA-----ESGLTRSD-QAWAY-----LHLLERAKDHLQDDFKG-----
12 -----LSS-----DLVASTDLIPK--PILEVLYTEFNKG-AH-----
13 -LVGP--LFEISLDL-LR---RQRF--CLF-----
14 -RDGS-----NCLTARGIFMGEPLTKALLTLYNLCT----E---
15 EIA-----IRKYLSYDFIQPVKVPW-----
16 -----RCF--AVAGDD--HIAV--GPVEYL--RLITQTHLA-AEAQLSAGKHA
17 VS--SVA-LRYCEKM----L-----D---LRKVKS--DWKVK
18 T-----INS-----
19 SLENYLNS-----
20 -----PFVDSV-----KVRLLSPCSKTNESFNDRNATAIGKGKSL-----ART
21 LSWLNST--LFPTKWV-----SMVRDR-----FFSRMG-----PLLPSRSS--
22 -----GVYWHLLLPTL--GGLG---L---GC-----P-EDIPD-----
23 -----LLGKLP-TLSK----SFLSEIAEGKITDLYRAFRSLLSNSTYRGYEMATTEIE-
24 -----LAKSLL----FDSKWLFPCPKSMREL-----RESYPEIEE-----
25 -----KP-VQNQLSYLMKQ-GWMR
26 -----YDRLEDTFLRG-----FLFKEILTGNVVVKAFN
27 SQKIKKRYSDLWDRFYDGD---CVLTEEQLKKAWK-----FRDFFDFYQVKNE
28 VKAW-----IRKKLVTCTALQELTYGLPDLTIR-----PS
29 LIAMRYHD-----EELRDEDKVLT
30 EALKASDI-----TTWED--AITHEVLKDLKAKLDSESQVEPKEALTNIVSESCDKKE
31 G-----
32 >KX883482.1
33 -----
34 -----MQSTNLNTSKLIQTGCPRAPAGKLGRCYRI-----FIPHYENG
35 DYLLPG-----RNHSLLKGVVVRNRMVKPML
36 -----LISKLTMGFGKRFATLVSRLPISHIER
37 AE----QSVL-----GVIDSLLL-YDSELFITDTG-----
38 ---YKLLKFIVK-KTFKVAI-----FN CALVTKLWK-----EFSFYVWAHVTETV
39 -----TREKPVRKRENFF-FSL-----VSEQDIEVLKLTLN---KK
40 TLTRLAHLTSSRQ--FP--PGE--RQQEIKSLKDFESH-VTEPYIGN---PV-----
41 -----FLRRIRLASRV--VGRKARRLGMKPLTDSHL-----SLAAA--GSFYTS
42 VKDG-GRAE---ELL-----NMLFKYLAYVPKESCEIKTPFY-T-LKDVEGVERWRTWCR
43 PAVY-----D
44 HFPN-----VSFGNLLPETLMGFETYYQGFDEILG--MQILA
45 CSYL-----EQEEH
46 LKCNCNQNPVRVLT--ITEPG-SKARI VTTGPCWL-YTL-QQPVS-----HVLRG-FL
47 GQHPSA-----AAGLTMSD-QAWQF-----LFLLEKAKSHF---GD-----
48 ---DFSC-LSS-----DLTAATDVIP--CILKELYEGFLEG-LRI-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -EGPF--LNIVGKMI-QL---PRLC--S-V-----
4 -EKIN-----SYFLNSRGIFMGEPIAKVLLTLLNLSC----E---
5 EIA---I-R-----DYLGVDFRASVQVRW-----
6 -----RCF--SVAGDD--HIAV--GPVGYL--DKITETHIK-AGSKISDLKHG
7 HS--RIA-VRYCEKILDVRNF-----K--GSWSKY--TIN-D
8 STEAY-----LAS-----
9 PFVDSIKV-----
10 -----RLLSPA-----AK----NVLSFNEKN-TA--FGKGKSL-----GRT
11 LQWLKD-----CF-DSKWI-SLVRDW----FFKRMS--S---LLPDRSS-G
12 V-YWH---LL--LPEHL--GGLG---L---GT-----E-RDFED-----
13 -----LIVRLP-SPSR----TLLKSIEDGN--PN-MGHIRLF--SGFTSNNTYRG-YK-
14 -----IPETEK-----EIAALFFQGILEEE-----KSDKFENLV-K-----
15 -----LYRYED-----LI-PNTAIRLLRTQ-GWVR
16 -----KDYLEDKILRP-----FLFKEILTGRAKVKAFTN
17 EHLKSRYARLWDLTYRGE--TISPQLQVCFKP-PAYHLYYK--MGEKLEL----PF
18 RGRV-----LKVN-----LLE-EGLLGMP-----DL-SIP-----WE
19 VIGDISGP-----L--AP-----KRGNHARSRV-
20 -----IKSPR--RRNS-----
21 -----
22 -----
23 >GEHH01003501.1
24 -----
25 -----EV-----GEQDRGPVK
26 SPLR-----MERTINKSKDVSLAPAKAVP
27 -----
28 -----MLWA-----KIRSVLEDVGCGKGKGLP-----
29 -----WRV-----PDKGFLAEKLK----RVAEGERKETGSGR
30 RA-----QNRELTAASGFL-FRK----VIPGTVGKTEAEVD-----
31 -----YLRSLGSDHR--HK--CGS--NCRWKLDVRRKVEE-LF---PIG-----
32 -----WDSEYKEYVER-----AYVGRS--ACLEVN
33 KKSG-GFAR---FVR---DMEWDD-----
34 -----
35 -----WVEGYWG--EGDIF
36 ESWR-----REVELGKEER-
37 -----KVSV--LLEDG--KTRTVTVSSGKQAWL--QPFH-----QMMYD-RL
38 ARQKWL-----VRGEVTRKQ-LKRMY-----EGSGLM-----
39 -----VSG-----DYEAAATDNFCP--EHSRWLMVEVISSR-CR-----
40 -YVPVDLLQLMENRF-DG---GFLV--G-----
41 -KTGR-----VERRKGQMMGDYLSFPFLCLVNYLT-----
42 AWR-----SLGENESL-----
43 -----MINGDD--IAFR--APEHEI--EKWEKGVAE-AGLSLSKGKTL
44 KS--SWM-FSLNSTF----W-----VEGKTGMTKLRAV
45 -----
46 RAKTFLKE-----
47 -----SER
48 GMWARFD-----EVRKGWSG--
49 -----KDRQI---IG-----RSFLE-----
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----WRAHQK-----KKWDDVGGDMIWLVRKTKGKGSE-
4 -----WKRTVP-----AKVRSTIIDL-----YERCALEGVRG-----
5 -----GG-KNFCEDDPRSLYDLRP
6 LSYDE-----EGEPGGRFEWP-----GLYKGHRSEN
7 RQKKVIYFKPIEVVVRLKG---MKRSCPPIRVRRVK-----NKE
8 RMGW-----VKKENPEHGIPAVISGEES-----N
9 ECSGEQGP-----
10 -----
11 -----
12 -----
13 >HAAN01020927.1
14 -----
15 -----EKE-----GEQDQGPVK
16 SPLR-----MERTKNKSEDVSLAPAKAVP
17 -----
18 -----MLWA-----KIRSSLRK-----
19 -----VGCEGSLPS-WL-----PSKGWLAENLK-----KVAEKERESTGSGV
20 RR-----QNRELTAASGFL-FRK-----VIPGTVTKAQAEDA-----
21 ---YLRLGEKHR---HK--CGS--DCRWKSDVRRK---VEDLFPIG-----
22 ---WDSRYEEYVER-----AYVGKS--ACLEVN
23 AKKG-GIRS-----WLAEQEFEEFWQGAWGEGE-----
24 -----
25 -----EFEEYAE--LCRLG
26 GEER-----
27 -----SVKV--LLDDG--KTRTVTVSSGKQLWL--QPLH-----QLLYD-HL
28 AREKWL-----VRGEVTRKQLE-----RMAKGH-----
29 -----GVMVSG-----DYEATDNFCP--EHSRWLMELISTR-CRY-----
30 -VPDG--VLKLAESRFVG---GFLI--G-----
31 -KAGS-----AMRCKGQMMGDYLSFPFLCLVNYLT-----
32 AWR-----ALGENRTL-----
33 -----MINGDD--IAFR--APNDRI--RRWEELVEE-GGLVLSKGKTL
34 KS--DWM-FSLNSMF---W-----VQSGSRMKKLRVV
35 -----
36 RAKTFLKE-----
37 -----SER
38 GMWARLD-----EVTKGWSG--
39 -----RSKEI-----MCRSFLE-----
40 -----WRAHQK-----KNWTPVEGGIVELVRRTGAKGGV-
41 -----WKRSVP-----AKVRSKIVDM-----YERCALEGERG-----
42 -----G-GRVFEEGDPRSLYEMR
43 PLEYWE-----EREPPGGRFDWP-----GLYHGYQ
44 KGKMKEKKVIYFKPLEIVV---KLKGFKRGRPIRV-----QKVKKRR
45 DMGW-----IKIENPEHGIPAVINGEESNECGD-----RQ
46 GSVHAEGS-----
47 -----SVPPWAAV-----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 >KT598235.1
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3
4 -----MGSENVV
5 GSAS-----RNGRTLPLFASMTDTSYYK
6
7 -----NNRYSSVLYLP-----KG
8 TPRPCRDFMARTT-IVVQRAT-----SSVRRIHPGVY-----
9 -----PTFSPSRIHTCAD-LLK-----QLKQFLATELDGASQEA--
10 -----QMAFQSVKKL---LP--ASC--KCLKHGMGLDLRVR-LSRPPPSL---PVG----
11 -----YLSFARKVSREI--FTKG--WDHRWGSKVS-----TFSPSLGSCIGKS
12 RKHG-GQLS---ELS-----VAGQQAWQESLNPLPG-----
13
14
15
16
17
18 -----SLEGELLL--VDSSG--KPRPLTRFVSEAATL--RPLH-----GLLYD-TL
19 SKQPWL-----LRGEITAEKLRGAGF-----DRAREEPL-----
20 -----TSG-----DYKSATDNLSI--EVAETILDVAWSSAKY-----
21 -VPAS--VFRYALAA-QR---PSLS---Y-----
22 -EDDE-----GLISTFVPTRGQMMGSYLCPLLCLQNYIA----F---
23 RYA-----EYVSGVEG-----
24 -----TPV--LINGDD--ILFQ--SELSFS--KAWMGIIVGD-LGLEVEPTKTS
25 VS--TEY-GSLNSTL-----LRWGPQ--GL---
26
27
28 AVVKTIRM-----
29 -----GMLREV-----SHPSNLGTTALQFARVGPR--NTWLLNF-----EEF
30 LSWHVST--IVKWRCVASDM---GFTGRL-----ALRAWS-RFRGGRLLWRDDVL-H
31 Q-MKIDRLPSAH-CPHNIVMGSEEFVTV-----PEESLTK-----
32 -----ELKRDAAVWMASRK--WELGREYTAFKQGKVVSERANA-----TRIPNLL-QD-
33 -----WRTSAR-----ELKEASSAVMRR-----KEYYWNHKVLC-----
34 -----HGDFASRGLL-----
35 -----VRDRPSVSTWR-----
36 -----WV--GLRRTWWRDRVAR-----EGTRVP
37 KVLW-----EALHPPLSPFSRDDVVGDLLQKTGGI-----FS
38 SAYLALKQ-----
39 -----AGVES--MVGQFFALVGP-----
40
41
42
43 >GFHZ01053830.1
44
45
46
47 -----SFSPRRCRAYRARTTELVRR
48
49 -----ALAD-----WAYIFGYPTPALGA-----LS
50 DGGSCLQFSASVK-GLLASCP-----SSVEEVMAWQ-----SVKKLL-----
51 -----PDSCPCTMNEPLL RG-----VVSGFRRPAVS-----
52 -----LP-SGY--LSFVRREVRLF-----PKA-----
53 -----WDSGLYESNVL-----SCSPGLSATTESS
54 RTEG-GCLS-----GWTGSHAYFLECCLRETAFEV-----
55
56
57
58
59
60

1
2
3
4
5
6
7 RDSASLLV--VQSAG--KPRPLTKFSDSTLVL--KPLH-----QTIYD-RL
8 RRSRWL-----SVGDSSETL-ARAGF-----VKVEGE-----
9 -----VL-TSG-----DYKSATDGLSI--EVAEVILETLLLNAVS-----
10 -VPRS--VCAAALDI-LR---PTLF---H-KGLCP-----
11 -EGLE-----PRVGQMMGSFLSFPLLCLQNRIA----F---
12 LFA--L-----RRCGLTGPQSEK-----
13 -----VPC--IINGDD--ILFR--SAPALS--ERWMALVSE-LGLEVERTKTS
14 VA--AEY-GSLNSTL-----LRWRGT--NL---
15
16 RVAPTLRF-----
17 -----GRLRGS-----GYVTS-----SRE
18 FRMFVEG-----LKGGRR-----FRAAVV-----F
19 F-RRHLGSLRS--VRLSL--LELG---F-----RGRLA-----
20 -----ARMSELFSLVPPGPPSEVVPLP---PVGHNVQLT-AD-
21 -----VATFVD-----ADSLGPELTAAANAR-----ETASWKFSVSY-----
22 -----SG-DRERAHAVRYCM-ALSR
23 VRPPEP-----FVFSRERPLWR-----
24 -----LLT--PESPAASRSRVLS-----
25 -----W-----FLSPLPAEPPKVAL-----MNR-----LL
26 ACQHELF-----PP-----
27 -----PSYES--SVAF-----
28
29
30
31
32 >GEWO01002579.1
33
34
35
36
37
38
39 -----YSRSKG-----GKFSEFRDRV-----
40 -----DEFTFTPINEI-VGH-----
41 -----MP-KID-----LID-----
42 -----VTGNVFLRAED--WVYG-----TTIRDC--AYPDLL
43 RDEG-ASADTRIGLF-----GLLWAVLEVLPDKDPIETCTND-----
44 -----M
45 AWP-----GLGSGIS--PQIFL
46 DEGT-----
47 -----VRAKVVS--LPEPG-YKARVITIVPLAISII-GSVAR-----HLLDD-MV
48 QSDEMT-----KLGLLSKDKL-YSFLR-----KMNGGKISGTADVSDSCK-----
49 -----FAYA-ESV-----DLTTATDTPPL--SSVRSSLQGTIAT-YLY-DKH-
50 -AGAD--FCQFAASL-STL---PRTF--V-L-----
51 -----PGGK--VT-----NPNDHRSGIMMGEGLSGTFLNVMSSAVRFLLRP---
52 TID---V-FSDYDGTTAVDADDFIGCNRDRIQEFLDGRERPMFEPD-----
53 -----PTS--TQSGDD--VLVFDNLAPGMA--RRYLVLLYR-IFGLLPSGTTF
54
55
56
57
58
59
60

1
2 FS--SEFYGTFCEEV----I-----V---RTHDSR--GW---
3
4 -----
5

6 TYVDTTKS-----
7 -----KLFQTT-----SVEGPEKLLGRIGAI----ASS
8 LKYSGDE--ALL-----LRAAER----VDAMIS-----GHPTLSD-R
9 I-LRY-DLPLG--LPECL--GGI-----PH-----PNGSLSD-----
10 ----FVACLD-----LVSLEILTGASTVPDAVLLEIE-FSRFFDDLPSDQ-IH-
11 -----LVEDTV-----RNVLEVIGGW-----ESPSFVDSTIP-----
12 -----FQV-FKRENFLPQGDLKYEV
13 -----WVEVAEAVSRR-----HGLISIRSAI
14 EKISSLRASIFLSGGVVP---PVNPLIRRRNRRQ-----KL
15 LTRL-----RQHSQSIGAVDNSTPRTVVVKRLRT-----CE
16 NRYFRFQD-----ISDIFELHG
17 Y-----PSFVV--PFSLS-----
18 -----
19 -----
20 >GEWL01001678.1
21 -----
22 -----LSDEF-----VPLGVWQVR
23 PSRQ-----CGSDLLPSCMAFEGDVL---
24 -----
25 -----
26 -----QSWR-----AVEAALLLAGASAPHCSEVHELLLR-----TF
27 VRGGIRAFALTLKGWGKAVRGL-----LSLEPLDTHLV-----GYTRLPKNCVCALR
28 REGMGES-----DIRDFLDFNGKLGGIPP-----FLLFSEFNKVLLPGRRRLSD
29 IQRIASLTTLGRS---AP--PAT--QEVLRGAVDAFVKN-VTT-----
30 -----HARTFDLFTIG---WEFG--QYAKGKSPPPRSI-----GLNTH--SSYDYS
31 RSLG-GKFA---EFR-----DRVIDDFLFWPIEEVVERKPNHDLIDITGNVFIRREDWS
32 AGTTLRDCA-----Y
33 PLLE-----RDEGSSADSRSRIGLLGLLWAVLEVLK
34 DPNI-----VAETVMVQPGGLGSRDPPQ
35 IKLGGFVPAKVVS--LPEPG-FKARVITVVPLAVSII-GSAAR-----HLLDN-MV
36 WADDMS-----RIGLLSKN-KLYDFMR-----VLFNGRIEGNATDGVSCN-----
37 --HLPSV-ESV-----DLTTATDTPSL--QAVGSLLSGATNG-YL-----
38 -EDSI--AFEFSQMA-VSLATSSREF--T-L-----
39 -PGGR--KL-----SYLLHESGIMMGEGLSGTFLNMSSLVRFLLPE---
40 AVI---AFWDYKGNSVAEADVFIGLNSGWIQLFLDGQGRA-----
41 -----RFEPPDPHSTQSGDDVILFND--LRPGEA--RRYLVLLYRIFGLKPSETTFY
42 SS--GTF-GTFCEEV---C-----V---KTKSSN--GW---
43 -----
44 -----
45 -----
46 -----
47 TFVDTVKA-----
48 -----RLFQVT-----ST--EGPEKIL----SRI
49 SAISDAL--KYNRDDTLI-----LRAAEA----VDAMIS-----THPTLSD-R
50 I-FRY-GLPLG--LPECL--GGI-----PH-----P-IGLASDFVACLDEIS
51 LEILTGASTVP-DAVL----LEIEFSQYFD--DLPDAQVELI--RKLIQNLLGTI-DG-
52 -----WERPSG-----LVTAEPFQVYTRER-----
53 -----IL-PQGDLQYSTWV-GIAT
54 -----AISERHGLSSV-----RTVV
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 EKISAQVRASVFLSGGSVP---PVNPLIRRRLRRD-----RL
4 LKKL-----REHTTSIGGVDLSTPKT-----V---VER-----LR
5 SCEDRYIR-----F-----SDVQDVFELYG
6 L-----PSFVI--PFSLS-----
7 -----
8 ----->GEZJ01006036.1
9 -----
10 -----
11 -----AGGQDNVPQTGGPTRPARNYLLSSVK-----SLASVGRKV
12 LG-----KPRTLNLLDGVLAQMTSGRH
13 -----KKPLSKPKKKHLRE
14 LR-----RNWS-----ATRGFLELTGVLKGK-----IF
15 APNNRKAIREQVS-FLEVLE-----SGLQVTAK-----RWSSLLRSKSI---
16 -----YQGTSDSPCTGIG-----TRGKTPMA-----
17 ---LFLASTYSRA---CVVDPPT--KDEVQTMVLAAIER-ISTREPLP---EPD----
18 ----VLRHLKGFIADL--FRSQ--RMKPRRAPVLP-----VPAPT--SCYERS
19 GSEG-GSII---ALT-----RFGRQDLSEAVHEHIARLFAESQRRAPEERHDYNWTDL
20 DEPLY-----E
21 GIPR-----DPQEMIRCLDRQGAMDLMC
22 PPPP-----PRLLTPETFKKAV
23 KDSSRPEERRAKALPIVQQD-GKIRVATLHMSGVVWS-ARAMS-----AFLLP-RL
24 RNLTFS-----KDMLRNRT-VQLES-----HRRSDDL-----
25 -----LV-YSA-----DLSKSTDPISI--ELSRFVLQEITRH-TG-----
26 -KPTW--WNDALEAV-IT---NHQI--E-Y-----
27 -TADS--VK-----KTVMSEC GALMGLPGWTVLSVLNAFA-----A---
28 EAA-----GARQGDY-----
29 -----AVCGDD--LIGL--WPTEVC--DRYEGNLAS-LRLKANKEKSF
30 RS--THR-GVFCERL---V-----LRSDNR--ASARS
31 CV-----
32 RLGEASGA-----
33 -----RAIDGG-----KGRVVV-----DAL
34 REVAS-----SSEAHR-----IIRQTA-----YRTIMRMA
35 L-SQT-----AKGSMKEGGGG-----IGQAD-----
36 -----AVTLIS--YLL-----HGPTQLYKVETDSRIAAMKLLRKAPNLSTGVPVE-DV-
37 -----LTAGRA-----QIELNRRLETA-----ERGPRPKYLPS-----
38 -----QS-VRRRIHSRRKDAKKAI
39 ADYGG-----VINAVKGLLAS-----
40 -----GCYVRENNRIRRDLTH--AVRGRAYTRALRI-----A
41 RSSW-----RTKVCPhAAKEAFEVSFPEHQPRV-----TS
42 DLEPMP-----
43 -----QAWDS--RLAD-----
44 -----
45 ----->GDKJ01026489.1
46 -----
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----STWGGMVAKTGLGSTSRK-----TAASGSISV
53 CPPR-----STQQKPSGKVAKKVSHNLSQ
54 -----LTTLGSLLVWQDLVKSNKAL
55 -----
56
57
58
59
60

1
2
3 TRVQRALHLYK-----AAADGKGNNLKVFREKKRRSRTMK-----LR
4 FKAIETAFLPILHAWLLV-----NGGASCCSRPN-----NIGKWMMSFTSY---
5 -----MMAMLTDSKYLKSITHQ-----WREESIRHSNTRVA-----
6 -KRLFQASTLGRA--VP-FPHT--MLQEDRAYQGAIDA-WTQGPAFV---SKQ----
7 ----HLDQIAEFISK--LPIR--RSKEKPGLHDL-----MLQNK--SCIENT
8 QAHG-GIAH---YLYTQEVEENKKRRDARLFKAGDLIQGQATTIRTKILMESFSRGEELD
9 SRLL-----E
10 GQRA-----TMSQFKRLQQEELE
11 DIRR-----QYRYQDLLDKARARFE
12 NTCHATQPMNPVI--VRELG--KLRCVTVHPADEVIT-ARTIT-----TKLLR-VL
13 RGCVST-----RDTLRGQC-PTVRT-----ETGR-----
14 ----PVLYSA-----DLSSATDHVDH--RLAQFIGEKLLDK-MG-----
15 -LESS---KPLLKSI-FS---PKEL--R-F-----
16 -TKKE-----YKKHNGTLTSRGVHMGLGPSWVILCLINMYA---A---
17 HNA-----GYNKDEV-----
18 -----QICGDD--LIAL--GSERAAC-QRYQQTMKE-LGMVVNETKSF
19 VG--KR--GVFCERF-----LEIKVD-----
20 -----RIG-----
21 CYASALDV-----
22 -----GHIAEA-----SAVRTTSDPD--KSRLAAV-----ERL
23 YKERHLP-----TLARKV-----ADQLAP-----RGVGK--
24 -----GPIPM--GGNG-----RGAPT-----
25 -----TP-----QLVHFMKCGPIRLVQTQFSPKL--ELNASRVPIKDIP-
26 -----WSGLRT-----EMMRHDAQMS-----YFKNESKPLTI-----
27 -----IR-AEHFSRLARGY-----
28 -----VRSTATVHLKS-----QSGTYLSLLPRAA
29 RQKVKLLLKRFTDSSNTIN--SNLSLPKATRAAL-----AKALTGPT
30 AERF-----VNLSTASNLLGLHHCPDADIDLA-----LM
31 RSRAKMTN-----
32 -----DGIRL--GPKEAELRRITNLDLL-----
33 -----
34 >KX883566.1
35 -----
36 -----MPLQ-----SMPSQKVNF
37 GSLS-----AVAASLRRLAREGRTVGSSQ
38 -----YATFARLAEKLRGLSPRQR
39 RR-----CLSTASWVAGGLSVSTTD-----RS
40 KGSSLGHRLAVVK-FAIKLWNII----SDVEQFVVDYK-----RVCNSAAYAAVRGY
41 -----TSIVAAGKCRDYL-HRA-----FGPCGIIPHKGRRS-----
42 -----LFALASTTRC--LP--PASLTPSRVKIELEALKSR-LTSVPPP---SGA----
43 -----LLSELVSFSHD-----LLRDAGPPVQVEVTGN-----ALGPPGKTSLDRA
44 QLYG-----SYAFKGVIMSKTLLMMSKSAIPPDVPFKE-----
45 -----
46 -----DLEENVQ-----
47 -----
48 -----RVSL--LAEAG-DKVRAITVPSSLAVQLAGRSIN-----RALIR-CL
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 RR-----SDVFGPGF-GKRPG-----TLMSGANLGDK-----
4 -----HI-VSC-----DLTAASDYLNQ--DVAIAVLADYLRSH-----
5 --SGI--WYPWASAL-LG---PQRI--R-S-----
6 -RDDT--YT-----TSRGVLMGTPLAWPILSICHAFAC--
7 YKA--L-G-----
8 -----NTYQFLC--RIVGDD--GLLI--CNASQY--RAYVSVMNR-LGFVINEHKTF
9 QG--KQC-GILAGRL---Y-----V---LKTRRA--KFQ--
10 -----DLN-----
11 SPTKTMAW-----
12 -----RMVDCQ-----APYVPELLAARGAQRFGRSDAP-----TAP
13 IRVPLGS-----QRENKA-----ALRAFK-----ILNPRTFTDA
14 M-LRA-GRP---TPLTT-----V-----QRRGLLV-----
15 ----LKNAAPSSGAR-----SLWKSSL--AKRGHYASLMT-SY-
16 -----LRQLMD-----INLPACSKAPVNGS-----YLARMSGTVSQ-----
17 -----L-VTARLNPKDS-RMIL
18 RPSYV-----LSRLGKRLDKK-----MH
19 TPKQSSVLRLLRGLEAEMV--PPSHVSVMEKLLR-----
20 -----SMVGPDDGARPISGVIGDFSQLGI-----TS
21 PTELYPQF-----
22 -----LHWIE--EGAP-----
23 ----->GGIC01436417.1
24 -----KNQRIS
25 NKKK-----IKKNNISKSKIYKKIVSPEV
26 -----KNYIDKLGLKTYKIED
27 ME----KIKD-----KEVEVLKDYVDEITHQHIKELKDER-----KR
28 SRNECVKDIRLFV-YLVRSILA-----SGPNQLIKGLK-----EKLTSVRRYCMGDI
29 EEEQEIVDKHFSQVLARDQEKRHLSKKERQQR-RLDVRSDRQKILGLYPQWLELPGIRKFT
30 RHQHQIQMTFISRS---LP--EGT--RNVKDEAIEGHKKM-VSSGGKMD---EE----
31 -----IRRDLVRWVRE--YCRG--KVSDIQLLGLAGKW-----SATPS--SCLEAT
32 RKEG-GHLMGSYQLYNMIS--PCRHTLFNQSRSEDIERMNWERHGGDRFCVSVSLVATSL
33 MIHESEPLKEFVTKYLHEAKQIPSLPNVIQERHNLKIAPGFSFRSKTDCKTFVKQQKRRRL
34 GIKD-----KVWDEKLGRMVND
35 DRRI-----LSPFLPSGDQSEVHSTLGGILSTEEFSDYAKS
36 CFDEGGLRSRVSA--IAERG-WKSRRVVTNHSELEIQS-GNIIR-----NLLNS-IL
37 SKDKRF-----LGSRRTHFD-DQEVL-----DAMGRNSVGG-----HP-----
38 -----LV-LSE-----DMTAATDSIYM--DVFAWIWNTLCST-IG---L-
39 -KFSS--IHIAGLIM-CA---PRLY--E-Y-KDL-----
40 -----TKVRTNRGIPMGTPISWPLLNLANMFC---W---
41 ERA-----NIDRAFEII-----
42 -----PPV--TILGDD--LLGV--ADLTYI--NKYSIDLKR-IGGISNPRKRV
43 LS--VFY-GLIGERI---Y-----Q---FDVEDN--EKSLV
44 IKHKETQLEAEQTNYQVHAEIAGKSSIKTLLSEDGKRKDLVELGSSDYRISSSQARYTE
45 SLKMMLED-----QDCHTGRRSKIVYELEIVKYIESDLIKEKDRCIEEK
46 LRFVVKKDTEDRQKEIENQIINICKIIYKCYGKGRSFSGII---GPFENY-----KGI
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 VEWLFGS--IVIMIRGVT-----YSISTR-----RKDCVTSKDGIIIRSRRPLWVR-Q
4 T-NGKLEKKAYEEGPVNLDPGFDG----VELHLNQYQKKFKTLLRQWHS-----
5 -----PRRNRLGLVISK----RLKCTSLKGIVSGTNKLGATDINKGIKPRWTCMGSIEH-
6 -----FVLES-----GDSNEVRQILNRIRSI-----FHPDIIRELRGHRLEPFI
7 PESMGGGGLVYKNKLDSTRWPCCSVFKKYINKILFGQNVKDYL-
8 EKINTSLVISEKQDW
9 ELASEMGRADFEDVKIVERLLHPSIEGITIGNHQD---LVTENILQYQGDFIRQMGPDP
10 FCKKVSKLVRGKTKQAMKK---EYESWASMRGFKGSIEEFIRFRDIYLSRFMAITIRQEE
11 LSTY-----LRRTKGGFPLIVPKEHTHLNY-----ITTKTNTSQTDQ
12 PRGQAETV-----LNADVTVVITDGLLETIV
13 PSE-----PQRRQ--LIAAAGRSMHLTHILPGLYFYTNPAVWFKMLESTP---
14 -----
15 ----->GCIB01004342.1
16 -----
17 -----THVNFFRVS
18 HDMV-----IVSKLMRHGAACAAAGRSIAK
19 -----
20 -----ALNTASKIMT-----RE
21 FCPSYELFAPLKV-----RTCSDVGDAWG-----EWSKECLRSL---
22 -----PRRKDRFRMAVKGTKR-----IFDVPCRRCDEASK---
23 -----LLAIENWDRI--TS--VKT--SPVDETRLELLRKH-----
24 -----VRELSSGW-----WRKIESCRKDTCPVLPSC-----SYVPDRQGCLETK
25 SIWG-GTLA-----
26 -----
27 -----TRDDECLGDHSLVR
28 RGT-----
29 -----VKTKG--KFRVVTMQSARVKRV-LRPLH-----NALYD-SL
30 SKDAWL-----VRGDVTRDDF-RAVAE-----TMGKNE-----
31 -----YL-ISG-----DYEQATDNIYQ--DAVKA VVDVLSEG-EM-----
32 -----L-EE--ERECLLGS-FH--NLRW--I-S-C-----
33 -----DFGA--KQ-----HEIVRGSMGMNLMSFPILCLLNKAC---H---
34 DIA-----SDLVYGEGV-----
35 -----RVG--RFNGDD--CLFV--GGDSMY--KKWREVTSW-FGLVVNEKKTG
36 RS--RKT-CELNSQI-----YRVSTN-----
37 -----
38 NLIPKIVL-----
39 -----NFLRPL-----SNEPGECLSAVL--TGISSCR-----PAV
40 QQWVVNV-----VMRYEI-----SLKGFA-----LSNLSRYW
41 I-RVLKKKKWF-----RTLWWDGPA-----
42 -----TTHTSIHARADKPLARGHLVLE--RDLPQVVGPPP-HP-
43 -----SRVEEA-----TIACAE-----VSRLWTAKWQG-----
44 -----VSGLVVVKPK-----
45 -----VDRPTFRARYE-----
46 -----KKRDIPLPKTKFI-----GV
47 IVRW-----GWLWPSEILGFSTDNVD-----YG
48 LLDTYPSS-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----PYLMLISTYR-----
4 -----
5 >NC_028476.1
6 -----
7 -----
8 -----MPTSNDMEAGCKATRSIRES
9 -----
10 -----
11 -----LLRATQLI-----SR
12 EFNLPPVPGPFNG-----TSCMELRSQWD-----TYAKWSCLQFV---
13 -----SKKEERRKT-LVQ-----LIKSGCRLFDGDCR-----
14 -----TCD--PKLKELEKRKWVEQ-MSTARPRVAYYKGLTRST--
15 -----WLGRLKDSARE--LITG--WGRNLEGCRMEDY-----IPDQS--GCFEAK
16 SIEG-GTLS--VSQ-----
17 -----
18 -----DFASF-----
19 -----
20 -----
21 -----PNEVRLGT--AKTKG--KIRVVTMQGANTKKV-LRPVH-----SALYD-YL
22 AGFGWL-----VRGDVTAADFEAIIE-----DKQGDE-----
23 -----KF-ISG-----DYEQATNHINI--DSVQAIISVIAEE-----
24 -PLLSDEEREVLIRS-FR---DVTY----Y-----
25 KNNGS-----FLCKVRNGSMMGNLVSFPLLCILNKCC-----Y---
26 DMS-----REIESEENGVPYCP-----
27 -----RVG--RFNGDD--CAFC--GTDRFF--EIWRETTSI-FGLVVQEKKTG
28 IS--SRW-IELNSES-----FDSLKH--RFVQK
29 -----
30 NFLSYLRV-----
31 -----TRDTPGDLLSEIV--SGTKGFK-----GST
32 KMWLINH--VLRYEIAIR-----GVCAST-----IPRKLF-----LLLIKRVW
33 F-RKTLSNPL---PPFPPT--TGVD-----RSIRQ-----
34 -----TVVSPPPLPSFI----PLIDDIETLVRKSHVKKWVGVSTIDGRRARCSNEM-NP-
35 -----DHDHFR-----PMVHGSLFFPFISSS-----YSPTPLSSLR-----
36 -----GSLLRQRKVAFDLSS-----
37 KPSA-----SVRFSRKHSWG-----
38 -----FTMSELCKDTLTTFF--GETWLMPPKSLCG-----ST
39 LPYY-----HPHLKIKCDFLTTVTPTYYPPPVS-----LTTLTKTLEHAA
40 TIYNLERT-----SSSILRNSRVRIA
41 KADS-----KLWSF--VPDNYIGSRLVV-----
42 -----
43 >LC150604.1
44 -----
45 -----MASQFCDSIAR-----RCGEVIRIT
46 LHRS-----NDHTTSPPHGGPRLEGNLET
47 -----FPPAKRDKH
48 AK-----RVMQ-----AFYAAIYIAMG-----YH
49 CHLRDRSLTRFFEKLYIRAWH-----EGLDRVCPDIK-----KQVTLLRSALI---
50 -----AETGVSPDVLGKL-LPR-----SWANLAHPRK-----
51
52
53
54
55
56
57
58
59
60

1
2
3 ---LFQLTCLGRA---LP--RPT--QKVISSKREALDG-FTSPASVT---TDIN----
4 -----VLHELVPMALNLRRYSAKADTTIFRGSSSTEQSTLGP----VFGTS--ATLFNS
5 RKEG-GRAG---IIH-----DACRSRPRQVGAIVALKRQAAIRKTAAPL-----
6 -----
7 -----RFDIIRSLKSEPLIG
8 EDSR-----RVLNDVRH
9 RFPNNLGESEIAV--VPELG-FKARVVTKSHPITVVA-GHQLR-----RRLYP-TL
10 YQSRVF-----SRALGDKP-KALRF-----ARRSG-----
11 -----AVFFSA-----DLSRATDGLSH----HTGVFCRN-AE-----
12 -IDPD---IIFRNM-----S-----
13 -VDGH-----ALKRGIFMGLPMSWTILSYIHRAV---C---
14 DSV-----DPIQNY-----
15 -----YLKGDD--LVAH--WTARQI--SLYRERVAS-VGMPLNESKTF
16 VG--PRK-AVFCEGY---Y-----E-----
17 -----
18 SSKASFNK-----
19 -----KMRTE-----IWLRQPSISLRRFPDS-----EAA
20 ALRLSEG--LDEV-----DYNLRS-----QYNRLV-----YCFYPNYIR--
21 I-SRKIRLNPF--LPPQL--GGSG---L-----LIPDLNS-----
22 -----QVHSVFDR---MRLNSIFGGV---KRFNASIPVGGSSSNMKQVN---
23 -----
24 -----IL-LSRVRMVPFRGS-----
25 -----GLTC SHFD KW L-----EFA
26 TSAAAFRDAKIGIFPKDPS---DHTFFAEAKRISR-----DRRR
27 GVDF-----PIDYRTVRSVCRSLEPLLSTVPLR-----
28 ---HTCE-----
29 -----
30 -----
31 -----
32 -----
33 -----
34 -----
35 >NC_033257.1
36 -----
37 -----
38 -----
39 -----
40 -----M
41 VR----RAYD-----LKRPLPYSQVEAAKPC-----A
42 HQ--VKTWKEFTN-FLLDLVA-----HGKSALVRRPG-----STPFTRSFTSF---
43 -----IVSRCAWSRGRGK-WAS-----
44 -----IAFSLGQLKRC---FP--ALP--DEMRREALDKHQEL-LGSA-AEP---FVEF---PI
45 -----VDRIIQQQYFPV--NWARK--TSVKVYN-----AIPNT--ACIE--
46 -----G-GSCR--DYL-----ASTGPKPA-----
47 -----RGC-----T
48 PIPS-----SGQ---ILAKRRY--ESIFD
49 KLYM-----QRLQP-
50 -----LPTLKAVA--VPDAG--KFRIITIDSVSAKCL--QPVQ-----RALLS-HL
51 SRFPEFEFC----TSTFQGELPPKWGK-----MAKGEKL-----
52 -----LSV-----DYSNATDGVDG--SFMSELMDRIIDRSGS-----
53 -----FELLALREQAVQETTLG---RRI-----
54
55
56
57
58
59
60

1
2
3 -SNDL-----CSFWQRRGTLMGSLVSFPLLCLWNASI----I---
4 REA--G-M-----
5 -----KKF--IVCGDD--ALAY--ATNTQK--KRWADYSRK-LGLVQSPGKTF
6 HS--RKF-GTFC SKV-----VGTEHL-----
7 -----NAG-----
8 RILAPMSG-----
9 -----ENYNRT-----
10 -----PKQFRH-----IYKRYC-----F
11 Y-PHGGEFRSLY--GPREI--GGLG-----GEVD-----
12 -----DIDVPS-RRAL-----RKAQAHLAR-----QGTKYQL-----GMLDMQD-
13 -----ENHAGI-----YTTAWSACDPT-----
14 -----
15 -----PQRLTPRLRKG-----
16 -----GSLDYRTKPIR-----
17 -----W-----FTR-----WD
18 RKSRSRRF-----
19 -----
20 -----
21 -----
22 -----
23 -----
24 >GBSG01013692.1
25 -----
26 -----RMVRHTNKRLT-----SFVKSPPCG
27 AG-----SDSRKNDRTLHRGRRKSVEQ
28 -----
29 -----QVWT-----ACYSSLVHS-----GW
30 DNTNMAWF--LHD-WVSKTLYS-----RGVRFTVGEKL-----KLSHGVRFA S L-----
31 -----RATGTVPDAPRI-PKK-----VYSAFLYLSVREKR--N--
32 -----AFAFSRLSRG--LP--CPP--KEDEIKALRDAKEMSVSYHPTSG-----
33 -----KAIEAIMWH----ILQE--RGGKTRLPD-----RLPNTLSSCFERS
34 TTQG-GVDG---HLR-----HLGF GELINIQCQYMGSCSSADMIRAKYSSFATDTLGQF
35 CLKSI-----C
36 DVAD-----NFDSTKD--DEALR
37 CLGV-----LKLRSDRP-
38 -----DPRCKSHC--LRTPG-YKCRVIGVPDALTFVE-GTWVR-----WS--S
39 QLLPRK-----HFDP SKNT-CPAVK-----GRTASDK-----
40 -----RY-CSV-----DLSKATDGLSH--DAVRAVIHALAVS-GR--IR-
41 -----KEDL--SMSLRSLG-VG--DTEM--I-W-SW RIPKE-----
42 -----PVEE--LR-----CVRGSPMGTPLSFIVLSWINSCA-----T--
43 ESF-----
44 -----THS--SIHGDD--AVGT-MYN TAE L--EEYKEFVRD-IGATVN VS KTY
45 IS--PTS-FTMCERM----Y-----APEKTK-----
46 -----
47 GMFTAFCP-----
48 -----PS--CPIPGGV-----VPV
49 PPQAHIP-----NLYLRR-----AERVQR-----TIHPWVSKLP
50 I-LR-----LPTSV--GGFG--Y-----TG YGL-----
51 -----KVS KA-----TRCRLATACS-----RNIADVT--KEILKQSTFRG-EG-
52
53
54
55
56
57
58
59
60

1
2
3 -----LFPSA-----VPTPKNSSKYYKYL-----NLIKD KR FST-----
4 -----VG-SEEFEDVV LAS-DFVT
5 -----FREKVILQSYL-----AIGGKLGRGSS
6 ERKPERTRARALFKGKAPC--NIAPLTVKRGIGA-----
7 -----LSKLSDRVKNQKVRVRTDVACVIR-----
8 -GRTTHSK-----
9 -----
10 -----
11 -----
12 >GBSG01012161.1
13 -----
14 -----EYFKIPDFNDALSR-----GRATVLQFA
15 HSRT-----YPNSVPIDVPSVKGHSALSI
16 -----WVKPIARKDIHR
17 AI-----SVCQ-----AVVSGFVVVYPKFACLHLKR-----QA
18 DVQACVPLVRILR-FILNLYSR-----GAEYKVSEYLK-----GMCNWLRYMACPSN
19 LRI-----SPPPRMPAGPFT-VLGKLDTSIFRGQLCSMKISTR--V--
20 ---LLCMSRMGRA---LP--PGG--KRVMAETIQKHKDI-VCDKDIRS---PPIS---
21 -----LVNAYGRFCKN---LRG--RRSLLTEPRVSFPL-----SSGVDGVLQTHIP
22 RRVG-GAEY-----VVLGALGCSSRGSGFSNAKELENRPLTML-----
23 -----
24 -----TLNEFERTLSDIIK
25 GEVI-----RR
26 LDQPVTRIARMVA--IPDRGGFKCRGITVSDSVIQQH-CTPLR-----RALYRRVL
27 SRLEC-----SRVLSGDYSHSTYS-----SRVASRRVGGRSTP-----
28 -----YVLLSA-----DLTAATDYFPF--EFQEVLNSNMFYH-----
29 -KEKSDTQVKAWMWL-SR---PHTI--M-Y-----PKV
30 DEDGF-----PVQGVTPLVQTRGTLMGTVPSWFHLNCLNSFMIR---A---
31 SLA---C-LSE---LVKRGRRWNILLTESAYDADREMLARELQ--SILPTITY-----
32 -FPRSE-----S--VICGDD--LQAV--LPIGAS--LIYEILIGM-YGGRLSVGKHY
33 V-----F-----R--LPREPA--AG---
34 -----PQV-----
35 PFTVFAET-----
36 -----FRTY-----SSSRAKNVSAP-----PL--RSVTTL M-----S
37 ADWKDIG---DRFRDIEL-S---CSSPAQ-----RRSLLSLAFLYNKKTLTNIRR-T
38 L-PGAPALPLF--VPRTW--GGMG---A---PD-----V-KGDPF-----
39 ---RTGVLN-PQLK---RMIVNL-----RSVINDP--VRFIFEVLKIR-WS-
40 -----WYAGNK-----
41 -----
42 -----SRVGLWV-----
43 -----NRIHD-----
44 -----
45 -----
46 -----
47 -----E-----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 >GBPT01016202.1
55 -----

1
2
3 -----GVMSKKFLEFLIPYVNSDALSRRGDF-----LCFHWQRPT
4 APHK-----ARGLLATGASLFVKPERPAE
5 -----IEFPFLRCALIPDFPNSVV
6 RD---RCAK-----TINGLLVTLVALWPTCQFGLPYT-----LS
7 SAIALGPWIELVA-WTMNAFSQ-----NGEKFAERLK-----GMAAWARYLASNNQ
8 LYPPPKN-----LTRFPLAVKGRFS-PSK-----LFGAIASKATSLR--S--
9 ---LLRLSRVSRA---LP--PAG--TEIQAEALASHLDL-CQDEPHVP----PA----
10 ----IAGLFTKAAAQ---ICAG--RLEGIEDARI-----ALSNS--ACFENT
11 RING-GAMS---YLR----SMLWEYHPHAFYSRSSDYGDVIRQSRPNEGYPDEVVETR
12 A-----V
13 FHKT-----CVHLAFKSTYG--QPRVY
14 QDTV-----PGDLLGFDE
15 EFPAETPNHRVVA--IPDRGGFKARVVTAGPAALQAL-AHNVR-----RVVYAQVL
16 AKPTR-----HALREGGV-ELFFA-----DIRSQRFFAKHWEPYFGP-----
17 ----TVLLSS-----DLTAATDRFPH--SLTLATNMGMESC-LSEDQRI-
18 -SANW---CAWRSL-SG---PQLL--V-Y-----PSY
19 ETKGENNEI-----VDFSVVSSCGNLMGTAPSWFLLNMFNLTIFE---A---
20 AWS---I-WLN---PTFRSGFTGRQL-FSLTPREELLPIWEELKKDVLAEAARDSFEKF
21 RIPREYGGLRLRDSY--ILVGDD--LGAV--CPYAVA--LLYEILIEA-CYGELSQGKHF
22 VQ-----PYREGC---F-----L--LVAEEF--AT---
23 -----VRS-----
24 GVLVHERI-----
25 -----EHLRG-----ASTQTSFDPR-----DKK
26 CLWAQIG--STLASAVAH-----IDDLGR-----ANSL-----MASMSLGTWRAT
27 L-LKK-GLPIY--IPEAL--GGLG--W--PH-----P-KGLEA-----
28 ----GINRTA-IRCI----EHYHLL-----RGFRSDP--VGFSIEIAKTR-AC-
29 -----WHTSIE-----VNHLGPLLDMRGH-----LNGIPVARDRD-----
30 -----GRL-EEQDEERPIGLLSIDS
31 PSSMRFSDFLSLLCMEGICFRYL--VREPLDHA VRK-----AGLSMLPIREESALVSGA
32 PSPLLNMQVDRDAREFSI---PLASSRYRRRLREL SLVNAGYDVQLLSPRTGYVEM-MT
33 QDMW-----LLEALHILPCAVIADLFPHVCH-----LC
34 HVTGQFVF-----
35 -----DAAIG--GRHA-----
36 -----
37 >GECQ01195387.1
38 -----
39 -----TMTSIK-----ATLSLPIVD
40 SVGL-----SRGGS--YLGFY-----NQRPATSQPGTSVLVLTPAS
41 -----FVVTRPMPGYPHLRIPKIGNGCV
42 RR---RLAT-----ILNGIGAAFVALWPSAQYGLQPN-----WA
43 NKEAMRAIVSLYK-WTISSLG-----SGEAYVAVRLK-----GLCAWARYLAAEKQ
44 LYPPPAPL-----ERPFADMEGRFQ-PSL-----LFRGSIAAMAESRS--C--
45 ---LLRLSRMSRA---LP--PGD--SDASLAAIADHKAL-SLSQPTVA----ME----
46 ----LLPSFRHVAAR--VSTG--RMAGIENPMI-----SLSSS--ACYQAS
47 RKQG-GSAN---YLR----QRLWNFFPDTWCMDERRPAVAPSSTFGRTVQAITGMPW
48 E-----V

1
2
3 PLVQ-----ERFHTTCEWDSADG--SDLPL
4 KRAT-----QPWASGHS
5 WDNPRHPTIVSA--VPDRGGFKTRVVTAGPAGLNTM-AQAVER-----TIVYSRVL
6 MSTPTK-----WGLINDGV-LKWFK-----QLPDLHYVDPRELGD-----
7 -----WVLLSS-----DLKSATDRFPF--SVVEAINDGFEESLTGEQRQ--
8 -CPNW---IAWRSL-SG---PQLL--H-YGPRVSHIHPPFD-----
9 -DDGV-----RVCGSCITTSCGNLMGTVPSWTHLNIYNLTAIR---L---
10 AWS---L-WTW---PWFRRAPYL---ERISDRSDLEPKMQNLRELILEALRRPACHPF
11 RFPSKW---KMNELT--ALVGDD--LGAL--CPWAFA--IVYECIIHM-TNGRLSVGKHY
12 VQPYREGSYLLIAEE-----FAILRS-----
13 -----
14 GRMTHLHL-----
15 -----DHLRG-----AQTVLSLTHD-----ANV
16 PAWASAG-QLLQNSLSVPT-----PEEPVA-----RYRALV---SFAHVSLASLRQR
17 L-MRA-GLPVY--LPTYV--GGLG--W--PH-----P-RGVTA-----
18 -----GVERTS-----RRALLHYHCL-----RGFRDDP--VRFAVEISKTA-RM-
19 -----WASVTE-----RNHFAPIVDALQSY-----LRRIPVSRIH-----
20 -----GQL-YAATSECPIGLHNPEL
21 EHKQIR-----LTELLNALAVQ-----GIGVRYLVTEPFE
22 EWLANNPVTSGELAAMAQQ--SLELPAGQVLE-----FDKREPTISSAARA
23 YHRR-----VDRLLSLRTGGGLQLLPRTSYGDL-----FE
24 QDREYLD-----LFVAVTPELEEQYPFLF
25 K-----PRGGLVLSHNPLALSSAVEMDSPSLPNE-----
26 -----
27 >GAMU01000156.1
28 -----
29 -----RLADVGSP-
30 -----
31 -----
32 -----
33 -----NFSLSFSSCLERSRKD-----
34 -----GGAGHYLREVL-----
35 -----WTPMIVG-----
36 -----TP-----
37 -----LWSSLRMFTDR-----YLPPP--GSLT-P
38 GMEGIGHDV-----IALRHTMYDHLATEVYYHHW-----
39 -----
40 -----EKGVS
41 -----
42 QPRR-----
43 -----
44 -----
45 -----
46 -----
47 -----RLDHVVHC--IPDRGGFKRVVVTAGPAVHQSL-AHAVR-----KVLYRTVL
48 PRLPTR-----HALRSGGL-VRWFR-----GLVLESSDITIRLGD-----
49 -----WVALSS-----DLKSATDLFPF--YLVEAINDGLADS-LSTTQRA-
50 -SPGW---VSWRSL-SG---PQTL---WTSEVRKTSLVKGVP-----
51 -PEDN-----SDWGHCCCLVDGTSSGNLMGTAPS WVLLNIFNYSLFR---L---
52 AWA---V-WST---PLYRRRLWGLLGRAIPREEMKTRGTWDIVKQKALRYIGDPGFYPF
53 RYDGPQ---RLNRLT--VLIGDD--LGAL--CPFAVA--VLYEIFIEC-AGGRTPGKHY
54 V-----MPWEDGS-----Y-----I-LIAEEV--AL---
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
 2
 3 -----IRD-----
 4 GRANFLQV-----
 5 -----EALRGL-----AGGAIVA---DTRSRK
 6 VPWSQIG--ENISTTLRTAR---PSVSTA----LCSYAH-----MQLSHWR-K
 7 L-LQQNGLPIY--LPTFV--GGLG---W---PH-----P-RGLDY-----
 8 ----GIQRTS-VLAL----RAYQVL-----RGYRQKP--LRFLTEIAKIR-ST-
 9 -----WLVGQS-----CADFSRLYDLILSK-----LEVLQESRDDH-----
 10 -----GMLAP-LDDLHPIGLPSPGIKS
 11 IP-----VQKFLDDIALE-----GASCQYLCs
 12 PCREAPPATPRIIPVEPE---AWGRWLPTVRLRPWDLIELPEPPKVIKKTVRS---AS
 13 KAYY-----ARARALVSLAGRIRWLQNRKDIPSF-----LRE-----DE
 14 ELLGSYHI-----TVNDLVREIFPHLS
 15 EV-----SDWDE--PTPWPy-----
 16 -----
 17 >GAMC01013668.1
 18 -----
 19 -----KVMTTNIRe-----SVISTPVFN
 20 SEGL-----RGGS--IVRFF-----IERSTTYPsrcSDKPVCPGS
 21 -----YVGPyRGNLPSFKVPPHHMRMSE
 22 IM-----RLAQ-----ILNAQVAAFVLVWPACQYGLTAN-----YA
 23 NVQAILPLVNLIR-WTCSSYIH-----NGEQFVAPRLK-----LFCAWARYKSGPQL
 24 HHPPEGI-----SRFPFCDVNGVPD-FGK-----VFIGSIADKSSRPK--T--
 25 ---LFRFSRMGRA---FP--PGD--INVQRQALADHREL-MTSEFSTP----QV----
 26 ----VLFNFYKVIKT--VTKD--CLEGIENPGI-----SLSNS--SCWERE
 27 RRRG-GAAE---EIT-----STLLEYYPLVPYLSRLEGSTFESYKEACQGLGYDETFIG
 28 MSTGP-----E
 29 GLPH-----EVSDVTGLAFDAIV
 30 QKEF-----SGNVsRR-
 31 -----LATHKVVP--VPDRGGFKVRVITAGQARIQCL-AQNVR-----KVIYRLVL
 32 PKTPSK-----WSLVENGV-LHFLR-----QLRRPAGPQADTLGA-----
 33 -----WVAMSC-----DMKSATDRFPH--DVVETINDAIESN-LSPPQAE-
 34 -CANW---VSWRML-SG---PQLL--R-Y-----
 35 -PGES-----AENQ---LTRTTCGNLMGTAPS梧HNLNFNLTLIR---T---
 36 SWS--L-WAS---ERWRRQLMRFFKESSWNTGDIASVDKTGLRDKILIILADKRFNPf
 37 VFPKIW--KFNELT--CIVGDD--LGAS--CPFAVA--VIYEIILEI-CNGRTSFGKHY
 38 VQ-----LWQDGA---F-----V---LLAEDW--GI---
 39 -----VKE-----
 40 QSLEVQPC-----
 41 -----ESLRGI-----VAATHQYDSR-----QAM
 42 PAWHQIG--TTLSASV-----QRVRPA-----ARRAVC--SLGHTALYDWRS-Y
 43 L-LRV-GLPVY--LPTSV--GGLG---W---PH-----P-RGLEY-----
 44 -----ALDRLS-----IKTKLAYQAL-----RGFRQEP--KNFVFQILKLR-SS-
 45 -----WLSSEr-----NSAYTRLLSTLKSy-----FNGFSVAKTAE-----
 46 -----GQLAPLD-ETNHIIGLPT-RWAI
 47 SSQESR-----CRR LTDVIDKL-----VLN
 48 GLPGYGMCENiSLTFDKPY---REASLAkRFSRE-----
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 -----RDRLVRLRQGHMKYVKKGSELAQD-----IRS-----DE
4 ELLDSLYI-----FIGDEFENTFPNLS
5 S-----VVLDQ--PTLD-----
6
7 -----
8 >GGBS01024784.1
9 -----
10 -----INRE-----QILAVPVFN
11 SESL-----RGGS--FVRFF-----LERSATYPSRCSDRPVCPEA
12 -----FLGRYRGSLPHFIVPTHKLSKQK
13 LA-----RLSQ-----ILNAQVAAFVLVWPACQYGLTVN-----YP
14 NAQAVQPIVELVR-WTCSSYIH-----NGEQYVASRLK-----LLAAWARYKAGPQL
15 HAPPGKL-----LRFPFCDSDGVPD-FRK-----LFVGSISDKATRPM--T--
16 ---LFRFSRMSRA--FP--PGD--KIIQEQLADHKAL-MTS--SFT---TQPQ----
17 ----LLRAFKTVIQT--VSKD--CLSGVSEPGV-----SLSNS--SSWERA
18 RKRG-GAAE---EIT-----SSLRLHYPLVQYLTLEGTSFEDYEAACRGVGFDESGIG
19 MSTGP-----E
20 GLPH-----EVSDILAMAHSACI
21 RDEF-----PDQVERR-
22 -----LAKHNVIP--VPDRGGFKVRVITAGQARVQCL-AQNVR-----KIIYRNIL
23 PKTPSR-----WSLIEDGV-LQFLR-----QLRKPTGPQADAL--GA-----
24 -----WVVLSC-----DMKSATDRFPH--DVVEAMNDSLESN-LAPGDSR-
25 -SANW---VCWRML-SG---PQQL--R-Y-PEM-----
26 -KEDD-----PLLLTTCGNLMGTAPS WVHLNLNFNLTIR---T---
27 AWS---L-WAS---ERWRRMILLPLVNKG--QHCDITMLDKQALKDKISVILQHDRFNPF
28 IFPKIW---KFNEMT--CIVGDD--LGAV--CPFGVA--CIYEIILEM-CNGMTSFGKHY
29 VQ-----PWQDGA---F-----M---LLAEDW--GI---
30 -----VKD-----
31 DYLLVQPC-----
32 -----ESLRGI-----VAVTHQYDSR-----QAM
33 PAWHQIG--TTLTASV-----NRVRKS-----ARRAVC--SLGHTALRDWRS-Y
34 L-LRV-GLPVY--LPSSV--GGLG--W--PH-----P-RGVEY-----
35 -----ALDRLS-----IKTKLAYSAL---RGFRQEP--KNFVFKVLELR-SS-
36 -----WLSTER-----NSAYTCLLNTLKSY-----FFGFSVAKTPE-----
37 -----GQLAPLD-GDRYIVGLYPT-RWAI
38 MYQESR-----CRRLIDVIDKL-----VLH
39 GLPVGYMCENISLTDFDPY--REASLAAKRFTRK-----
40 -----RDELLNLRKGHLLNVKSGSELAKD-----IQE-----DE
41 LLLSSLYI-----YIGDQFEQTFPHLS
42 A-----VVLDS--VSDD-----
43
44 -----
45 >GEEA01024908.1
46 -----
47 -----RHTNMPTINREQILAVP-----VFNSESLRG
48 GSFVRFF-----LERSATYPSRCSDRPVCPEA
49 -----FLGRYRGSLPHFIVPTHKLSRQK
50 LA-----RLSQ-----ILNAQVAAFVLVWPACQYGLTVN-----YP
51
52
53
54
55
56
57
58
59
60

1
2
3 NAQAVQPIVELVR-WTCSSYIH-----NGEQYVASRLK-----LLAAWARYKAGPQL
4 HAPPGKL-----LRFPFCDSDGVPD-FRK-----LFVGSISDKATRPM-----
5 --TLFRFSRMSRA---FP--PGD--KIIQEQLADHKAL-MTSSFTTQ----PQ----
6 ----LLRAFKTVIKT---VSKD--CLSGVSEPGV-----SLSNS--SSWERA
7 RKRG-GAAE---EIT-----SSLRLHYPLVQYLTLEGTSFEDYEAACRGVGFDESGIG
8 MSTGP-----E
9 GLPH-----EVSDILAMAHSACI
10 RDEF-----PD
11 QVERRLAHNVIP--VPDRGGFKVRVITAGQARVQCL-AQNVR-----KIIYRNIL
12 PKTPSR-----WSLIEDGV-LQFLR-----QLRKPTGSQADALGA-----
13 ----WVVLSC-----DMKSATDRFPH--DVVEAMNDSLESN-LAPGDTR-
14 -SANW---VCWRML-SG---PQLL--R-Y-----
15 -PEMK-----EDDP---LLLTTCGNLMGTAPS梧HNLNFNLTLIR---T---
16 AWS---L-WAS---ERWRRMILLPLVNKG--QHCDITALDKQALKDKISVILQHDRFNPF
17 IFPKIW---KFNEMT--CIVGDD--LGAV--CPFGVA--CIYEMILEM-CNGMTSFGKHY
18 VQ-----PWQDGA---F-----M---LLAEDW-----
19
20
21
22
23 GIVKDDHL-----
24 -----LVQPCESLRGIVAVT--HQYDSR-----QAM
25 PAWHQIG-TTLTASV-----NRVRKS---ARRAVC---SLGHTALHDWRS-Y
26 L-LRV-GLPVY--LPSSV--GGLG---W-PH-----P-RGVEY-----
27 ----ALDRLS-----IKTKLAYSAL---RGFRQEP--KNFVFVKLELR-SS-
28 -----WLSTER-----NSAYTCLLNTLKSY-----FFGFSVAKTPE-----
29 -----GQLAP-LDGDKHIVGLYPTRWA
30 IMYQESR-----CRRLIDVIDKL-----VLH
31 GLPGYGMCEISLTDFDKPY--REASLAAKRFTRK-----RDELLNLRKGHLLN
32 VKSG-----SELAKDIQEDELLSSLYIYIGDQ-----FE
33 QTFPHLSA-----
34 -----VVLDs--VSDD-----
35
36
37
38 >GDAR01017830.1
39
40 -----CTPVFNGEALRAGNSVTFFQR-----RALNFPARQ
41 SGPQSALDY-----CIRTVEEVPTSHQSGLLPRC
42 -----RIPAAFPFLGNTKKEKAI
43 RN-----RVAI-----CCNAAVAAWITVWPKSQWNLALN-----AA
44 NSEAMCGLRDLVA-WIVRTYSG-----AGEKFVAERLK-----AFCSWLRYLAAPNQ
45 LHTPPGPV-----TRFPFSDSEGTIV-PSK-----LFRGSIARCATTsk-----
46 --SLLRFSRMARA---LP--PGG--PWAAEEALSAHRTA-MQLVPEKV---EP----
47 ----LALIFTSIVSKVTPSALL--KKADFEKAAF-----TLSSS--SCFEVS
48 RGDG-GAAE---VIR-----TLVRDTFPDLSAVITDDGDVFRPTLFDDSGELRQGGTI
49 FASGGERFDYPH-----S
50 GLPD-----EHDIYLNKVVHRALV
51 ESVF-----GAPRASELYSQR
52 EYDRVVPKHRVIC--VPDRGAFKVRVITAAPAAIQLS-AHNVR-----KAIYHHVL
53 PQLPTK-----WALHEGGM-VDFLK-----QLRKPELKEDLGP-----
54
55
56
57
58
59
60

1
 2
 3 ----WVALSC-----DMKTATDLFPH--YLIEAINDGIEGN-LPHSAKR-
 4 -CANW---QAFRSL-SG---RQKL--A-Y-----
 5 --GGS-----PEAP--EDWVENTCGNLMGTAPSWCLLNLFNFALFRATAWA---
 6 LWA---Q-----KPIRKKGFRDLMDSFAEIPRDSQKWSFLRDTLRQLPEHSRSKR
 7 LFWYPKEVPKFNELV--ALCGDD--LTAC--CPYGVA--CLYELLLEL-ANGRASPGKHY
 8 VQPWVPGSLLLLAEE---F-----AIVQDN-----
 9
 10 -----
 11 GEVHHVRA-----
 12 -----ESVRSF-----AAVTLYSDTS-----DKR
 13 VEWAQIG--SILSTAV-----EKARPS-----VRRGLC-----SLAHLGVSGWRAV
 14 L-LRS-GLPVY--LPTQL--GGLG---W---PH-----P-RGFY-----
 15 -----ALERTS-----VRSLQAYSVV-----RGFKSDP--VQFCFKLTKMH-SL-
 16 -----WCRPGI-----ASAYSELTHTLDFL-----FRDIPVSRDQF-----
 17 -----GAI-GPLSAEFPPLGTQ---
 18 -----VAGLRSAPLFD-----CI
 19 DAVIKRCISFQMASTRSL--VEGEAAEPKRVST-----AAKQ
 20 YRRF-----VSDLRLRTGDVLAPNRR-----
 21
 22 -----
 23 -----
 24 -----
 25 -----
 26 >KX883540.1
 27 -----
 28 -----MALT-----YELCMPVN
 29 GDAL-----RVGS--IVNFF-----CRRSCTYPARQSDGPCKCPKD
 30 -----FISREILKSSLPHFRVPGPACSHKL
 31 RR----RLAQ-----TLNAFAAAFISVWPSSQFGLCPT-----PP
 32 NLTAMWALCKFLH-WVCKVYVG-----NGEPWLAPVLR-----GIATWARYHASTTH
 33 LHLPEGI-----FRTPLSSADGSID-FGK-----LFTGSISRKATSPS--C--
 34 -----LLRFSRFARA--LP--PGD--DWTCQAQGLVDHQKA-MRDNPPTVP----EQ----
 35 -----LTKLFKDVTATRI--SCQP--LLKDVGTPGF-----TLSMS--SCLERT
 36 RSQG-GAAE--EVK-----EAVWEIFPGLRRIHEGVTEEDTELSRVPCLDRNPLYSD
 37 VEYYYYSFY-----E
 38 GLPE-----HVPRVER--LTYLA
 39 CCEK-----VFPPC
 40 RLERRVPSHRIIP--VPDRGGFKIRVITAGEALTQSL-AHQVR-----KVVYRKVL
 41 PTLPTV-----WGIREDGV-RNFIE-----QLRMPSSPDPRRLGR-----
 42 -----WVALSC-----DMKAATDRFPH--PLIEAINDGLESN-LSESQRR-
 43 -CPAW----LAWRSL-SG---PQRL--L-Y-----
 44 -PGGD-----VVLSSCGNLMGTAPSWALLNMNYNYALFR---L---
 45 AWS--I-WST---GWSRRKFPELAAREVLPQRQSLDSWEPLKSRLVRLSDSRFHPW
 46 NFPKGL---KFNELA--ALVGDD--LAAA--CPFGVA--MLYELILEL-CNGKPSAGKHY
 47 V-----MPWIDGS---Y-----L--LLAEEF--GF---
 48 -----VKG-----
 49 DRLVHLHA-----
 50 -----EFLRGI-----VTGTNCVDTR-----SRN
 51 PEWAQIG--LSLAASI-----AGCRNE-----TRRAL-----SFSHMLADLRK-T
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 L-LRF-GLPIY--LPASV--GGLG---W---PH-----P-RGLEY-----
4 -----ALERTS-AKVL----RAYSVV-----RGFRQDP--VQFILIIARLR-SS-
5 -----WMRPEL-----RTPFSNLLSLLSGY-----FRSLKVSRDEQ-----
6 -----GRL-APCTPESPIGLPSLDG
7 -----TSKNFFDCIDS-----IGVH
8 GVTVAFVAGEELPLLEQKE---RFRTIGASRRAYG-----QA
9 ISDL-----LRRREGHVPPPLLPRRSDYAAL-----IAE-----DE
10 KLVRDCHL-----YLGESLEAQL-
11 -----PFLQG--VLGKSPEEALEDIGEAE-----
12 -----
13 -----
14 -----
15 >KX883626.1
16 -----
17 -----ME-----YILCTPVVN
18 GDGL-----RAGS--FLRFF-----IRRSPRAYPARLS DGPVGLSS
19 -----FVSDVSGQQLPQIKIPPMESP
20 RR-----RVAS-----ICNAFLAAWIEVWPKAQYGLEAN-----LA
21 NILPFRSLAHLIK-WVIRTYLG-----NGETYLRGQLK-----LFAAWLRF
22 FLAGDQL
23 HLPPSPL-----RRFPFSSEDGRLL-PTL-----LFTGSISAKAGSPA-----
24 --ALLAFSRFARA--LP--AGN--KQAEVEALHEHFKL-MQS-TPE---VVPP-----
25 -----LAKHFREVASK--VGKD--SLSDIEKPGM-----TLSFS--SSLERS
26 RAEG-GAAE--EIR-----EVTWDYFPGLRRMHFDVDMQTDVYVSSFEIDPKTDDL
27 VV
28 DGT-----PYDYGFTFG-----E
29 GLPG-----ECERVVHMAAACI
30 YKAF-----GTEHLRGKIPLSLPRV-
31 -----IPKHRVIA--VPDRGGLKVRVITAGPACLQLS-AHNVR-----KVIYRSV
32 L
33 SKTPTQ-----WGIIEHGP-QRFLE-----QLKLPADRDPELGP-----
34 -----WVALSC-----DMRSATDNFPH--YLVEAINDGLEQN-LPEDIRS-
35 -SPNW---VAWRSL-SG---PQTL---FYKDILN-----
36 -EDGE-----SSEITSTCGNLMTAPS
37 WALLNIFNLSIFR---L---
38 AWS---L-WSSRDFRRRCRRQFGQESDSA
39 RYPKGY--SFNELT--CLVGDD--LAAA--CPLMVA--ALYETLLGL-AGGSVSTGKH
40 VQPWEDGNFLLVAEE-----FGRIEG-----
41 -----
42 QGIRRLHF-----
43 -----EHLRAF-----TQISSSVDT-----SKK
44 ETWAQLG--SALDSAV-----RSCRST-----HSRALC---SFGHIASSGLRQR
45 M-LKL-GLPVY--LPTSV--GGLG---W---PH-----P-RGLLV-----
46 -----GLSRTS-ALAL--RAYRVL-----RGFRSDP--VKFSCEIARLR-AT-
47 -----WYLSEI-----SHKYLTLEIVRGY-----CTFPVSRDKDGKVAP-----
48 -----LTEECP-IGFPNEDFPDSVNLFD
49 FIEEVV-----ISGISAGFISGEETFSAGHISQRAHKAFRIVYTVG
50 SGTSQTSTRS RASHTIQAA--ARSFRRACRLLDLRQNIQGAMSYS
51 DYMNCLAEDEEL
52 LARY-----HVLKNCDRVRVSLPSFATPLQE-----DD
53 QMSPPQEH-----
54 -----LGRSCV
55 -----LVH-----
56
57
58
59
60

1
2
3 >KX883471.1
4
5 -----KSAPNGKDTRLKPAKPAVIKIDLSHP-----AGSRVDPKV
6 ADKNKPRLEFASLDSSLRQVSLLAR-----SSLGVQRISTIPLSPDPSKS
7 -----AAQLFRRFAARGKQNIAK
8 FQ-----RLWS-----TVLNAVSLAFKIDQFDLRT-----PQ
9 GRIAYTDLQCAF'R-HLLARFTT-----VGYQDSVRVVA-----FWALQARQMAL---
10 -----NSAVATSGHRPN-LVK-----FLRGRLDIRRASQR--T--
11 ---KFLFSGVGRA---LP--FAKIDAIRIDAEIKAHRAR-VTVPGTTD----PL----
12 ---ILGKYAEWLQKKLNLLDS--RADLPSMKLNIAPL-----TAARS--SAVGAP
13 SSKG-GRYG---FYK-----SLFEQGPQFQERILEHA-----
14
15 -----IKEVE
16 DCPT-----
17 -----PLPVKLTV--IPDKG-DKARTVSTTLPGIQVL-GEQLN-----QSMMV-LL
18 RRLAPS-----AIPLSENP-DVSRR-----LNNHG-----
19 -----TM-LST-----DLSAASDYIDH--TVLKVTTRLVARK-LG-----
20 -----WQFNSKLVRL-TG--PFMF--TID-----
21 -QEGV-----KDSYLGRLRGAPMGLPLVWPTLSLMNMFC----A---
22 EYR-----ISPNSK-----
23 -----RSY--AVLGDD--HVAR--WPKADI--DQYFRNMSD-IGLKINKAKTF
24 LS--KRD-AVFAEDL----Y-----S--LVNGKV--KR---
25
26 TVVPKLSA-----
27 -----VMSTRQQ----GTD
28 PQIFAII--ASLQSALA-----NAKNFK----ERKQIK--KIFFSQWKKLTR-Q
29 I-RRA-QVPLQ--LPTQL--GGLGIAPF-----
30 -----KLSKLT-RKQR----SRLAYLLS-----PEYASKS--ARSDQRSYFRSLDQ-
31 -----FKNANR----FAHAIRNKDKLVKE-----AARQVIEQISD-----
32 -----EEERT-SRSFNATGCAVEDCAP
33 QLPE-----GGVPAEAVLKE-----LLALKLTRLTQ
34 EGKVQATDRSVSSGTVLRR--LTRVWERLPLCLK-----PV
35 GARG-----YTKLLSLCVQSSQILSTQSVAKLI-----TT
36 NLPLYYAV-----GVFYKDGGMPIQC
37 S-----PNSND--AYDSATAG-----
38
39 >NC_033474.1
40
41 -----MTVTSVPVVAQVRIPTLRLAA-----AVLQRSPVL
42 RKGS-----LRESLLRLRSLSLKAEERLD
43 -----PIERKF
44 RE-----ELLR-----WVDAALYMSLPMYNRTISGS-----HQ
45 VEQYLLALWKLLH-YVYK-----NSIPDVVKTLK-----DLFFRYLDAGLQSD
46 REF-----RRRRFTYLGKRHP-LKG-----FFAIAVKGLTRES-----
47 ---ARTIGLAGRC--IPSLPLS--KERRLEEYLNKER-LTTPMPTI----PACRGCG
48 HLPPEHPDKRLAAGPGCQRLSRDH-IVESVIKFAEVFARPVKHLPPMSYPADNMVYIPGP
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 LHHG-ETRP---WLM-----GVTKPAVEAISDAFDSFDEWKEAVNSPKLPYQLGIWSVH
4 PEA-----I
5 ENPP-----ERDPAVGLTRVISEGESLKT
6 PKRR-----ANWYDRKDLRVWYSERHAPSV
7 LHPKDPEAERRDDHRILPPG-YGLRALKRSPIAPILN-DVVLSGEDQCLVPVHKVLF-IP
8 EAGAKV-----RVATVPSV-VSLQTAARQINGIVLALKHCPTFHGLKNS-----
9 -----KC-FSSKTFAMPDGAKFYSVDLSAASDRLSQ-PLSLSVLWPFIRES-EK---AG-
10 -LE-N--AFEVARLS-CG---PQRI--I-L-----
11 -EDPE--VMELLADPKYSAHLPFFE--TEAGLLMGTPLAWPALSLLHLFAALEAGI---
12 DIH-----
13 -----RVF--IVGDDGGLIAT---APEY--QKYYHIMTNILGLRVNREKTH
14 ES--PMM-GVVAQKV-----LKIRPA--PCP-H
15 AGTPRGKLGIGPL-----
16 TVWDFVRH-----
17 -ELNLNPNEPVYWGK-----VELDQSHCYVPA-----LT--ARSRRRR----DPH
18 APWNPPE--WSDVARWK-----RWRRER-----LPKALR---L---VYPSLVS--
19 Q-SRYRGLRME--APAIL--GGCG-----PL-----TEGAVRK-----
20 -----GLTPV-----QRAGAYAIAR--EIMHGTRKDL--SIDVGPDLTTRT-IP-
21 -----RAVILA-----ANEVFNELPRSETR-----GTGLPKEAA-----
22 -----LRSL-TSFFVRVAWEG-GFRD
23 ANSYL-----GRSLGHRVVVRH-----P
24 GRLLEAKGRIFLRSRGKRA--VPKPLTVVQAFKR-----LPKPVDVSPGVLKKG
25 LNRL-----IHDLGLGRQRKRRLDYVLSAGR-----DDA-----EE
26 GTSSSSVQ-----VP-----EAMSSLLNRLR
27 HGWG-----KVGQ--IQSLGSGGKGKNLTSTHHEA-----
28 -----
29 >GFLP01304447.1
30 -----
31 -----FLALGLSSG
32 LYLTS-----LDSKHFSTLAFELDNRKRSR
33 -----
34 TFHDTIGDFRPR-----
35 -----KLFAKLIT-YQM-----QNPEILNEE-----KLRAVMPEVSI---
36 -----STNRSRSSKFPVT-LGS-----
37 -----LP--HPS--ETVVEETTRKYLEV-VTGPKSSI-DSPPD-----
38 -----LLTSIEEFGRK--YVVD--RLHRVKDKVS-----WVPNLNTASLHVT
39 -----AREG-GRMT-----ELLGADTVLKLLLEYTDPSEPN-----
40 -----
41 -----DLEEAIV--QTSIE
42 IAKL-----
43 KALDGNFEAEVMA--VGEAG-YKARILCKYSAHILLV-GDIIR-----RQLWP-LV
44 SEEPWFDDKDYSFSEKLERAF-HDLPF-----WTGYHEP-----
45 -----EV-VST-----DLSSATDFIPF--EVARALWSGILSG-LGVB---
46 -----FDSW--IFQHVITW-AG--PTKV-----
47 -----KVGD-----SVHTTNRGIPMGTPLSFMTLCLLHHYA---V---
48 -----KST-----EKETGME-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----GPY--VIRGDD--FIGV--LPNKEV---YFRKMEE-LGFVVNHSKTF
4 IS--KVP-GIPVRPS----L-----R--SDTSNE--DTV--
5 -----RRS-----
6 SACGVFAE-----
7 -----TVVQFNFREHPV----TSGK----RKP
8 FTWQPIR-----LSDVTV-----
9 -----IPDV-----PFAGLVR-----
10 ----EEKSSI-----KRSQLVRYG-----EWYRQIF--RRPDGTLLDRGR-
11 ----WVKKAS-----RVVFTAIRRLSKDL-----IMRAVESGVPI-----
12 -----
13 -----HTPLEIGGAGI-----
14 -----PD--ARGRLTPLPDFVT-----RS
15 KITY-----LLNHTGNSRVRSARAT-----LR
16 SAAAEELGG-----
17 -----KIWSS--FDKV-----
18 -----
19 >MH213236.1
20 -----
21 -----MTASLRVYRWTLSTDITAWP-----VTPGVGGVG
22 S-----WISVIPGPATPCVEHMTDRR
23 -----
24 -----AIWT-----AVAITLTAVNPKGYRN-----
25 ---IIPPKLLAK-WVLR-----APVQTIVKGLK----SLAHEYRSKAV---
26 -----LGGPLHFVPEVPLVFQK-----WLKGHVCHATQ-----
27 ---YAQLGRLGRL--MP--KPW--QKITSKAHVHEWIRV-NTRRPDGPIHEPSVDNER--
28 -----MSNDIYRFAHE--FATQ--RKDRLTTEVNFT-----PSESA--STFMTS
29 MSEG-GRLK---EMMN----DARDLIGEFAFDETIVNLAAREGGYKPPPRAQ-----
30 -----C
31 GMQP-----VLDELILENQIQMA
32 DSKE-----
33 -----FSPRVIG--VPEYG-WKTRVLTLPNVALTP-GDLIR-----QQLWP-LV
34 SEEEWL-----DADRVPSW-NLFNN-----MLSRSLDRVGTW-----
35 -----VSC-----DLSNATDYVPH--LYCKALWAGIYDA-LGVS----
36 -EGSW--VRYYTEKM-FS---PMVI-----
37 -AGTK-----GRITFRGTQMGTPLSFMTLCALLHRFAVVRSGF---
38 EYC-----
39 -----PH--LIRGDD--LIGCFHKP----RIYMGMLEE-LGFKINKAKTF
40 LS--TIG-GTFAERT-----FRAERKLIKNRVH
41 GYFFSH-----
42 RKVTAGQW-----
43 -----AYTRAT-----LLQDIPLRGLIKVA-TTGPNRYR-----SKL
44 LMLSEIV---ENVAMGFR-----PKRRAL----IRR LAT-----QGRGDILIA
45 A-LKS-GIPLS--LPREV--GGAG--L-----
46 -----PTVSEILRD----RIYIDLE-----TRRRLG----
47 -----WMISHV-----PKGQLVAK-----LRRRASGGLIG-----
48 -----FT-ERYLEGIKNISLGLRV
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 P-----VTEPERTAHYL-----A
4 AKRRTIHLKPWEVVHRPPC---SFREWGRTRFSFG-----KV
5 GPRW----ALGGDRLSLVSDRLKVATACIVYVPIY-----QGFR
6 GPIRRISD-----
7 -----GIWTEDPGQNEGYZRHQPSDCSAMS-----
8
9
10 >GFJU01140648.1
11 -----
12 -----LAKPMRWFLP-----KNPLSWPAI
13 PMVE-----VGTASDVVKALTREDVKTVL
14
15 -----SFIE-----PCRIVINSIH-----RR
16 PWLDLRLPYRSFVN-WVCCMVHE-----KQEQAIAATAIK-----DLSFAVRMWAIE--
17 -----GRFPETWKSRSAH-LPH-----TFVRWWKSFLCHTG--T--
18 ---RVQFARLGRC--LP--VGS--RAVMDTARANFLAV-VHTE-PKD--VSVD----
19 -----IADELYEFGRS-----FIVSEGLKPVSWE-----PNEST--ATLNFH
20 ASQG-GRMA---ELI-----YDARLPLHLLEEFTP-----
21 -----I
22 TEPN-----ELLDEVIVETAIEK
23 SRIE-----
24 -----DWEPELLV--VPERG-YKARVLVKFPASALLV-GDIVR-----RQLWP-QI
25 EQAEWL-----DLDREVNS-EKLTY-----FLRRCIDPNTGLG-----
26 -----QC-VSA-----DLSNATDYIPH--LYAQALWAGILEK-VD-----
27 -APDF--VKGYVEKM-FS---PMKI-----
28 --EGR-----MTKRGIHMGTPLSFMTLCLMHRYA---L---
29 -----RKAGLHR-----
30 -----HPH--IIRGDD--LIAI--IPQP---DIYFSAMTS-LGFKINKSKTI
31 VS--KNG-GTFVERV---F-----K--FEREDV--IIPH-
32
33 PILRLPPQ-----
34 -----SVITRA-----LEVVDIPLAGMIWPS---ERGSKL-----RTI
35 GRWYNQC-Q-----SRLTPR-----MKRTVS--MLIYMRHKDILA-K
36 A-TKL-RLPIS--TPVEL--GGCG---L-----PNRH-----
37 -----GRNDFDAPHWY---RAIVGWAAC---HESMAKKFSGRVTRLDGKDPNPP-
38 -----YYIQNI-----RENQSRSTYE-----YDPTTTRYYRG-----
39
40 -----MKRRFQIFRWA-----GAL
41 DEDRFKARVEREQSRFISL---PLNQWYEPFKNLK-----RV
42 RPRW-----VPRRRADTDRLIRRIKMFNAE-----YVL
43 PYTSRVQF-----
44 -----LGSDSYKTGGDT-----
45
46 >KX883548.1
47
48 -----MTKSMRRFYLP-----HEPWRWPVT
49 PTPE-----VGRGSVPTLADRDMR-----
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----DF
4 AR----RVWH-----SVVTPIRMVTSGRLSR-----
5 ---IAEFRSFAT-WMVRKVA-----GERSVVCRELK-----NLSFSFRVWAV---
6 -----EDRWTFHSDQ-LPR-----SFVRALKGLAPCSE-----
7 --ARAQFARIGRS--LP--TGD--GKVVEAAIRSYQRI-VVGR-PNP---SWD----
8 ----VADELYAFGKR---WAEE--RAHLVRGRVNWAP-----NENT--ASLNFK
9 ASEG-GRLA---ELI-----ADSAAVIEPLLEFTP-----
10 -----E
11 GEKS-----ELLDEVLIETSIEK
12 ARTE-----
13 ----RPIPKVIA--VSEKG-YKARVLVEFPASTLLP-GDIIR-----RQLWP-ML
14 EEEEWI-----DMDVVPNE-ERLKA-----FLRHSVDHSA-----
15 ----TC-ISS-----DLSNATDYIPH-LYAQALWRGVLD-A-IP-----
16 -APEW--VGDYTDKM-FG---PWDL--L-F-----
17 -PDGK-----TVTSQRGIHMGTPLSFMTLCLFHKA---V---
18 ERS-----GFS-H-----
19 -----HPH--LIRGDD--LIGL--FPDP--ARYCRTMED-LGFKINKSKTI
20 IS--SKG-GVFVERT---F-----R--FTTGRE--PRRF-
21 -----ERR-----
22 SLASFLPD-----
23 -----RMV-----IRETTELKDVPVAGLLNPS---VEGSLL-----RSV
24 GRWYGQL-----PDLSAR-----RRKRLH--TVIWSVFGDTLR-K
25 A-RCS-GIPIS--TPVQL--GGCG--I-----PNRH-----
26 ----GRLDIDASFKM----RRSMALLNK----RSVARQF--AYAVSAIDNLPHD-
27 -----FDRLFL-----KHHEESQRSRSVYT-----YDPQSTGYYQS-----
28 -----ASRR-FKMYRLLGGGA-ESVK
29 -----RRGLREGKLHK-----
30 -----PA--GFRRWV--RLFNS-----LKMS
31 CPRY-----AVGRRANSSRLETYLKISFGTYIPV-----AS
32 PGYSIFGP-----GNGEGGVAHQ
33 -----PQLVV--ALP-----
34 ----->KX883542.1
35 -----
36 -----MTTSSRYFDLP-----SNILSWPVT
37 PGVV-----GAG-----HWHQLPVSALTDRDGVRYSN
38 -----
39 -----SVWK-----AVVTNIVLAHNKKFRN-----
40 ---VVPAKIFAG-WLLRCSRNA-----GGRYYVAKQLK-----QLSFYRQWSVT--
41 -----GRRPTPKSDVPRT-FQK-----WLAGHVCHAEA-----
42 --RAQIARIGRC--LP--EGD--TLVQSKAIKAHRSI-VTRP-PCD---VNME----
43 -----LADDIYYWARK--FSTD--NAHLFKEEVSWA-----PNENS--ASQGAT
44 ILEG-GRME---ELI-----RDAKAESDFESVKSLAEGDF-----
45 -----D
46 PWVS-----KALDDAVYMTACTV
47 PPKK-----
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----EWIPDVMA--IAEKG-YKARVLTKFPAALVV-GDVIR-----RQMWA-MV
4 SSDKPWA-----DMDRQPDD-VKFLS-----TIRSALDRHG-----
5 -----IC-VSS-----DLSNATDYIPH--IYAQALWAGLLEP-HH-----
6 -VMPY--INSYCAIM-FS---SMKL--R-Y-----
7 -PDGT-----VVTTSARGIHMGTPLSFLTLTLLHRFC-----V---
8 -----EKAGYGS-----
9 -----YPH--IIRGDD--LLGI--FPRPEV---YFNVMQQ-VGFSINRAKTI
10 IS--RTG-GTFAERT---V-----R--FSHSLA--TQE-L
11 SNP-----LKR-----
12 TLGQFIPV-----
13 -----NISSV-----KVLQDLPVGGVVRAT--PGKGSLV-----KAL
14 GRWFQSQT-----SNIV-----PRQRK-----AYRAIG-----LNHGDLV--
15 I-RLSTTVPPH--LPLDL--GGAG--L--PD-----RKGRV-----
16 -----GLNGVP-----FAIRAAIGHA--ASHHETAVKL--TGLIARTDGVS-RG-
17 -----FVDVFS-----KKDF-----REQWSKSIW-----
18 -----TTE-PQETEFYQRTR-RYWR
19 -----YFGHRDRPSRP-----
20 -----I--SFRQWRSGLMTP-----RV
21 KARW-----VPRSNSDPSRL-----ANR-----IR
22 SMTGTYIP-----YTQPTTG
23 ST-----YQWAA--KCG-----
24 ----->GGQW01011558.1
25 -----AEHYERFHLP-----VNPLSWPVI
26 PGVV-----RAGGWYNLPSEALTGEGLKF
27 -----
28 VK-----SVWE-----AFVVVFVASHKKTIKN-----
29 ---LVPPKIFLR-WVMSCSKNE-----TSRDFCCEQLK-----KLAFIYRRWAVE--
30 -----DKYPTDHCSIPSS-FLK-----WLKGGMSSCTQ-----
31 ---RAQIARIARC--LP--KGT--TVRMERAIEGHKKI-VSIPARVD----ID----
32 -----LADEIYHFAAK--YTAE--RKEKLIKEVSFH-----PRENT--ASLEFT
33 EKAG-GRLA---DVM-----ANASDWLRRRPIALSEHLYTNKSYLDE-----
34 -----AIVEAAV--KQDLA
35 YPY-----
36 -----RAEVVA--VAEPG-NKVRVVCKFRSVPLLI-SDIIR-----RQLFP-IF
37 EEDENM-----DFDQDIRS-EKFGK-----MIRRCLHNEG-----
38 -----VI-VSS-----DLSNATDYIPH--HYAQALWHGILEQ-FD-----
39 -APEW--VLKHVLRM-LS---PVEF--N-Y-----
40 -PDGS-----VIMSQRGIQMGTPLSFMTLSLLHKFA-----V---
41 ERS-----RNGE-----
42 -----SSY--LIRGDD--LIGV---FKYP--RSYYSEMEN-LGLKINANKTL
43 QS--HRG-GVFAEQT---V-----K--VSWKKS--TRTSY
44 HR-----
45 TLYDYFPA-----
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----AETDEV-----SRIWVLDDIPMKGLLHVD--PGVYGSK----LRMI
4 GRWYSQW--SSYYPICSR-----KR-----AYRAIG-----IAHKELIK--
5 T-ARFLGLPSS--TPLEL--GGCG--L--P-----NRKGALA-----
6 -----LDASFLHRGQ----IGYAASHSDSRFSHFVGKLDGIV-
7 -----KDDGLL----EQSIERSRRSLGSS-----VFETDPF-----
8 -----
9 -----
10 -----NTEFYLKARRA-----
11 -----RAYLSPDYLKERERCT---PLSQWKKSFRKLP-----TV
12 KARW-----VPKRDKDSDLGIDRLIDLKI-----FN
13 GRYVPFTP-----
14 -----
15 -----
16 -----
17 >GBDZ01000351.1
18 -----
19 -----SSPDNTGAMTSQCETTYGLP-----ISVLSWPVN
20 PGVV-----GAASW-----IPNPRTALTGVERGRALAYA
21 -----
22 -----Q-----GVWY-----ALTSNIEAVHPRPFKS-----
23 -----LIPAKILAG-WVVRCSATA-----AGRTQCSGLLK-----RFAFLYRLWAT---
24 -----TDTKPGLVDGIPK-CPQ-----LWIRSLCSHASV-----
25 -----RAQIARLGRM--LP--QGD--AVTVRKALKKHRSV-VTQSTRAN---VF---
26 -----VADEIYAFAKE--YMEN--REEHLREKVSFHP-----TEGT--ASRDYT
27 -----IERG-GRVQ---ELL----DNAWAMMRKPGYIDYQKYPDKS-----
28 -----
29 -----
30 -----
31 -----RIDEFIV--QAAID
32 ADKR-----
33 -----LHHYRVSVMAAAEPG-NKARVLCKFPAVALVP-GDIIR-----RQLWP-IV
34 ASDPDL-----DFDQDPRS-EKFLQ-----MIRRTVDREG-----
35 -----TI-VSS-----DLSNATDYIPH--MYAQALWRGLLES-FD-----
36 -----APSW--VQEHNRM-FS---PMEM--T-Y-----
37 -----PDGK-----KVWSQRGIQMGTPLSFMTLSLLHKFA---V---
38 QAS-----GHGLSAH-----
39 -----IIRGDD--LLGV---FTSP--ADYFRSMER-VGFKINREKTI
40 -----VS--RVG-GVFAEQT---V-----R--VRYTPR--RQP--
41 -----KKI-----
42 TLADFMGP-----
43 -----PREDVESV-----TVLNDIPLKGAVHVD--TKGSVL-----RGL
44 GRWYAQW--SEQYP-----AKCGRT-----AHRAVM-----RVHQALIR-D
45 A-KRW-RVPLH--TPLEL--GGAG--I-----PDRRGRL-----
46 -----GVADLS-FDMR----RKLGYACSHFDRSYARAVRRLD-----GGST-DP-
47 -----FEEHFE-----QMLKKA-----PLGSCVYEYDP-----
48 -----YA-TEYYAACRRKY-YLTK
49 -----VTRPARKDLNP-----
50 -----P--PLRDWL--KQFDN-----CPAV
51 QPRW-----AVSKHADVTLLIHRLKMFGA-----YV
52 PFTTPETS-----
53
54
55
56
57
58
59
60

1
2
3 -----ILARR--KVR-----
4 -----
5 >GBFE01004865.1
6 -----
7 -----GGAMTSQCETTYGLP-----ISVLSWPVN
8 PGVV-----GAASWIPNPRTALTGVERGR
9 -----ALAY
10 -----AQ-----GVWY-----ALTSNIEAVHPRPKS-----
11 -----LIPAKILAG-WVVRCSATA-----AGRATQCSGLLK-----RFAFLYRLWAT---
12 -----TDTKPGLVDGIPK-CPQ-----LWIRSLCSHASV-----
13 -----RAQIARLGRM---LP--QGD--AVTVRKALKKHRSV-VTQSTRAN---VF---
14 -----VADEIYAFAKE---YMEN--REEHLREKVSFH-----PTEGT--ASRDYT
15 -----IERG-GRVQ---ELL-----DNAWAMMRKPGYIDYQKYPDKS-----
16 -----
17 -----RIDEFIV--QAAID
18 ADKR-----
19 -----LHHYRVGVMAAAEPG-NKARVLCKFPAVALVP-GDIIR-----RQLWP-IV
20 ASDPDL-----DFDQDPRS-EKFLQ-----MIRRTVDREG-----
21 -----TI-VSS-----DLSNATDYIPH--MYAQALWRGLLES-FD-----
22 -----APSW--VQE HIDRM-FS---PMEM--T-Y-----
23 -----PDGK-----KVWSQRGIQMGTPLSFMTLSLLHKFA---V---
24 QAS-----GHGLSAH-----
25 -----IIRGDD--LLGV---FTSP--ADYFRSMER-VGFKINREKTI
26 VS--RVG-GVFAEQT---V-----R--VRYTPR--RQP--
27 -----KKI-----
28 TLADFMGP-----
29 -----PREDVESV-----TVLNDIPLKGAVHVD--TKGSVL-----RGL
30 GRWYAQW--SEQYP-----AKCGRT-----AHRAVM--R--VHQALIR-D
31 A-KRW-RVPLH--TPLEL--GGAG---I---PD-----RRGRL-----
32 -----GVADLS-FDMR---RKLGYACSHFDRSYARAVRRLD-----GGST-DP-
33 -----FEEHFE-----QMLKKA-----PLGSCVYEYDP-----
34 -----YA-TEYYAACRRKY-YLTK-----
35 -----VTRPARKDLNP-----
36 -----P--PLRDWL--KQFDN-----CPAV
37 QPRW-----AVSK---HADVTLLIHRLKMFSGA-----YV
38 PFTTPETS-----
39 -----ILARR--KVR-----
40 -----
41 -----
42 >KX883517.1
43 -----
44 -----
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----MAVVIEAIH-----RR
54 PFTNLVPAKMFAG-WLIRASRGP-----GGQKYVCQQIK-----SLAFDYRVWSI---
55 -----TAKKPNYSGD-IPR-----AFRNWLSSLCCCHAE-----
56
57
58
59
60

1
2
3 --TRAQFARVARM---LP--VGN--KLVIDQALRKHRSV-VTR--RPK---LDIG----
4 -----LAHQIYGYARD--YIHE--RKHSLSERVSFHP-----TEGT--ASRDVP
5 ASAG-GRLR---DLV-----DNAHRVMSQRPIVLSEYLISHKK-----
6 -----
7 -----YFDDAIV--EAAIY
8 TDEK-----
9 -----VSTYNVDVLAYPEPG-FKARVLCKFPATALLA-GDIIR-----RQLWP-IF
10 ESDPNM-----DFDQDIRS-EKFQN-----VIRRSLNRDG-----
11 -----TI-VSS-----DLSNATDYIPH--EYAKALWAGILDA-FD-----
12 -FPEW--VENYLERM-FA---PIRM--S-Y-----
13 -PDGV-----TVTSCRGIQMGTPLSFLTLSLLHKFC----V---
14 HKS-----GHE--R-----
15 -----SPY--IIRGDD--LLGV---FSSP--RQYLTVMEE-IGFKINRDKTV
16 IS--KDG-GTFAEQT---V-----K--VTWKAK--ERD-P
17 L-----ARP-----
18 TLYDFIYT-----
19 -----DQKV-----VSSITNLDDIPFKGLIHLD--NKGGRL----RQV
20 GRWYAQW--SPYYP-----PRKGKV-----AYRAIR-----RTIGNVL-R
21 I-ARSLRIPIT--CPMEL--GGCG--I--PN-----K-RGTMQ-----
22 -----LDANFQ-HRSR---VGYAASHESH--NFQLAVRKLD----IGNASDLL-ED-
23 -----YRQHVA-----DLPKGK-----SVYMFDPYS-----
24 -----SED-FSGYRRKRMILL-NNTG
25 NYSAY-----HTKPVPLHRWL-----
26 -----AVFSRCKE--NRARWVPSRRCSG-----SD
27 LRRL-----IDR-----IK
28 MLSPIYVP-----YTTP-----
29 -----NAQIL--SIRI-----
30 -----
31 >GFLP01591397.1
32 -----
33 -----AYSTTMRELLP-----SMLTYVRVG
34 ESHP-----SPSERTGPGSRPRDRKSRAKA
35 -----GSRGKR
36 QRASREDLVWQ-----AVWLGFIAGFISYQ-----
37 ---LGCWRTRE--WVKRSIRR-----DGVHNTCSKLK----DLAVSIRGYVL---
38 -----SDVQPPMVPVPN-----RIQRWLRAEVDKDT--LS-
39 -----GLAFTRFARA--LP--KPD--FRETANAMASHLDL-ITDRRVVP----EG----
40 -----IQDQIEQHVVR--MFGG--QFVGILRHPST-----PISNH--AVGERS
41 KARG-GYNAHIYDCA-----GDPEFLGVDPPLDRIPAYPPTRALWRLADSPPGE-----
42 -----
43 -----GIPEYMASASALSA
44 ARIV-----EQFARDTS-
45 -----RVVHSATG--IGEQG-DKCRIITVPPASLFAA-GDVCR-----SRIWP-RV
46 QQSDAR-----IFGNDEDA-KSSL-----QMGLPR-----GS-----
47 -----VY-VSA-----DLTKATDGFAH--DAVRAVLRGLARC-G-----
48 -----VDNH--TVKLMSET-LGVGDRQHYV--K-Y-KLAELPRRKREWAVQRFETIEE-----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -DGVS--TH-----VLVPLVRGILMGTPCSFIILSILNGWCC---M---
4 PLG-----
5 -----PRT--VICGDD--VASA--CTPEAV--DNYDRRVTV-IGSGLHKRKTF
6 IG--HKG-LLFCELY-----
7 -----
8 VLPGYGEH-----
9 -----RCFEPV-----PLKSLAKDG-----DGT
10 LDTSDFD-----TFIWKR-----MDRACR-----VLWRDVRA-K
11 A-RRLNRWPQ--LPVEL--GGLG--H-----PSSG-----
12 -----KMGAVP-GSVR----NRLATLIKVGPKIARWTVAAGPLDWRSFRSDKE-VA-
13 -----WSLFES-----TDAAVSQLEAD-----SRYVDTSFVSY-----
14 -----HE-ANQYVSVLANQ-----
15 -----LYSAHGGKFVN-----
16 -----EFNRKAMKPGKIQYP--GVGSLQYSAKAPM-----SMV
17 AREY-----LAKLHAGGEYLPTDSVRKIRR-----
18 --RIGKST-----
19 -----
20 -----
21 -----
22 -----
23 -----
24 >GGCO01034162.1
25 -----
26 -----SYTKNMKVRLSSHTRLEFKD-----STAPSPRKS
27 GGCK-----PLSKIRTKHWHSDKERTPEQ
28 -----
29 -----IVWE-----ALWAALVVSDSVN-----DL
30 DVSSMRGSFYLRD-WVARSSR-----RTVGYVVSVLK-----EELNRLRAISL---
31 -----GSDPPPIAHG-FPK-----KLDTFLQKVFTPLG-----
32 -----LFAFSGVSRG--FP-TCVD--KVKVSETLVDHLVN-VANPAPAV---PIN---
33 -----FLRSIEQFVPR--YIQKSVCYDRKLADKMHDPSPVVC---KSSPS--ACYENP
34 RGAG-GFYAYIKKLG-----DKISEGQNDLLADIRAGVKPRPEDICHPILKGIAKASYD
35 PDR-----
36 -----TQDLLSPKEMSDIL
37 DHFY-----FYSGCNEYL
38 DNPDKEIVHRVAV--IPERG-YKNRVVTAPPASILSM-GEVVR-----SSIFP-YV
39 KAHPSC-----DIVKDGDPMVVCF-----PAVDGA-----
40 -----KL-VSA-----DLTKATDGFSH--DViRSVGVGLRNA--G-----
41 -FPQH--VVDFVFVDT-LGAGERPHYA--E-Y--DVHAILRRHSKNTVKYNSLCSFLQSYG
42 WDGIS--SK-----VKVPMRGSPMGTPCSFTLLCIVNSWA-----L---
43 DHC--S-----
44 -----NGP--KICGDD--MLAY--ITAKEF--SKYKVRVAA-IGSGVHPLKTF
45 IS--PYA-GTFCENI----Y-----I--KDPCDS--KFP-S
46 VVAFGRAIKCPIKAMLY-----PQK-----
47 GSNGRIDW-----
48 -----PDEIVHQV-----FTGETRMSKEVKCLSSFND--EPEYHVQ-----
49 RRFEAGP--GDFGYNDCSV-----FRLGRK-----VHRVLR-----TLRRDQIRSA
50 L-KK--GRYPW--LPACL--GGLN---Y-----P-RKNPK-----
51 -----GVRGLD-KHCR----ARLFQFTHSD--LSVVELGKFM-----SSLD-NA-
52
53
54
55
56
57
58
59
60

1
2
3 -----KRPLPE-----HRELEFISDMVADA-----EINDAIVEDTL-----
4 -----ERQGECS-FVDSYSLARHQ-GSVR-----
5 --SNF-----FRSMGGKLDLS-----ASSYSIFNAA-----
6 RIRWPAVKGELGYDKRTP--FSLIIQDLDNLLR-----
7 -----APRYRVLSSFAEKLPGVHPKSNR-----WH-----
8 HLHGA-----
9 -----GGWND--GIL-----
10 -----
11 -----
12 >GAIS01005902.1
13 -----
14 -----E-----
15 WRLQ-----ALSKIRTKHWHSDKERTPEQ-----
16 -----
17 -----
18 -----IVWE-----ALWAALVVSDSYN-----DL-----
19 DVSSMRGSFYLRD-WVARSSR-----RTVGYVVSVLK-----EELNRLRAISL-----
20 -----GSDPPPIAHG-FPK-----KLDTFLQKVFTPLG-----
21 -----LFAFSGVSRG---FP-TCVD--KVKVSETLVDHMVN-VATPAPAV---PIN-----
22 -----FLRSIEQFVPR--YIQK--SVRYDRRLADKMHDPSPV--VCKSSPS--ACYENP-----
23 RGAG-GFYAYIKKLG-----DKISEGQNDLLADIRAGVKPRPENICHPILRGIAKASYD-----
24 -----P-----
25 DRTQ-----DLLSPKE--MSSIL-----
26 DHFY-----FYSGCNEYL-----
27 NNPDKDIIRVAV--IPERG-YKNRVTAPPASILSM-GEVVR-----SSVFP-YV-----
28 KAHPSC-----DIVKDGDPMVVCF-----PAVDGA-----
29 -----KL-VSA-----DLTKATDGFSH--DVIRSVGGLRNA--G-----
30 -FPQH--VVDFVFDT-LGAGDKPHYA--E-Y--DVHTILRRHSKNTVKYDRLCGFLQSYG-----
31 WDGHS--SK-----VKVPLRGSPMGTPCSFTLLCIINSWA-----L-----
32 DHC-----SNGP-----
33 -----KICGDD--MLAY--ITAKEF--TKYKVRVAA-IGSGVHPLKTF-----
34 IS--PYA-GTFCENI-----YIKDPCDSEFP-----
35 -----
36 -----
37 -----
38 -----
39 SVVAFGRA-----
40 -----IKCPIKAMLYPQ-----KGS-----NGR-----
41 IDWPDEI-----VHQVFT-----GETRMSKEKVR-----
42 L-SSFTDEPEYH-VQRKFEAGPGD-----FGYND-----
43 ---CPSVFRRLG-----RKVRRVL-----RTLRRDQ-----IRSALKKGRYP-----
44 -----WLPACL-----
45 -----GG-LNYPRRNPKGVMG-----
46 -----LSGLCRARLFQ-----FTHSDL-----
47 VVELGKFMSTLDNAKRPLP--EHRELEFISDMVA-----
48 -----DAEINDAIVEDTLEKQGECSFVDS-----YS-----
49 LARHQGSV-----
50 -----RSNFFRSMGGKLDLSAE-----
51 -----
52 -----
53 -----
54 >GAIR01012062.1
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
 2
 3 YTKNMKVRLSSHLTRLEFKD-----STAPSPRKS
 4 GGCK-----PLSKIRTKHWHSKDERTPEQ
 5
 6 -----IVWE-----ALWAALVVSDSVDL
 7 DVSSMRGSFYLRD-WVARSSR-----RTVGYVVSVLK-----EELNRLRAISL---
 8 -----GSDPPPIAHG-FPK-----KLDTFLQKVFTPLG----
 9 ---LFAFSGVSRG--FP-TCVD--KVKVSETLVDHMVN-VATPAPAV---PIN----
 10 ----FLRSIEQFVPR--YIQK--SVRYDRRLADKMHDPSPVVC--KSSPS--ACYENP
 11 RGAG-GFYAYIKKLG----DKISEGQNDLLADIRAGVKPRPENICHPILRGIAKASYD
 12 PDR-----
 13
 14 -----TQDLLSPKEMSSIL
 15 DHFY-----FYSGCNEYL
 16 NNPDKDIIHRVAV--IPERG-YKNRVVTAPPASILSM-GEVVR-----SSVFP-YV
 17 KAHPSC-----DIVKDGDM-PVVCF-----PAVDGA-----
 18 -----KL-VSA-----DLTKATDGFSH--DVIRSVGGLRNA--G----
 19 -FPQH--VVDFVFDT-LGAGDKPHYA--E-Y--DVHTILRRHSKNTVKYDRLCGFLQSYG
 20 WDGHS-----SKVKVPMLRGSPMGTCSFTLLCIINSWA----L---
 21 DHC--S-----
 22
 23 -----NGP--KICGDD--MLAY--ITAKEF--TKYKVRVAA-IGSGVHPLKTF
 24 IS--PYA-GTFCENI----Y-----IKDPCD-SEFPSV
 25 VAFGRA-----JKC-----
 26 PIKAMLYP-----
 27 -----QKGNSNGR-----IDWPDEIVHQVFTGETRMSKEKVRLSSF-----TDE
 28 PEYHVQR-KFEAGPGDFGYNDPSVFRRLGRK-----VRRVLR-----TLRRDQIRSA
 29 L-KK--GRYPW--LPACL--GGLN-----YPRRNPK-----
 30 -----GVMGLS-GLCR-----ARLFQFT-----HSDLGVVELGKFNSTLDNAK-RP-
 31 -----LPEHRE-----LEFISDMVADA-----EINDAIVEDTL-----
 32 -----EKQGECS-FVDSYSLARHQGSVRS
 33 NF-----FRSMGGKLDLS-----ASSYSIFNAA
 34 RIRWPAVKGELGYVKRTP--FSLIIQDLDNLLR-----
 35 -----APRYRVLSSFAEKLPGVEHPKSNR-----WH
 36 HLHGGWGG-----
 37 -----GMMGY-----
 38
 39
 40
 41
 42 >GACI01002802.1
 43
 44 -----VRLSSHLTR
 45 V-----DFKDLPTPLPVKGGSKPRS
 46 -----TIISRDPRHRRP
 47
 48 IQ-----IVWE-----AVWAALIISDSYD-----GR
 49 DVSSLRGSFIFRR-WVARSSY-----RSVGYVVAILK-----LEMNRLRAISL---
 50 -----KSSPVPGVPEGFP---K-----RLDQFLDNISPFG----
 51 ---LFTLSTMTRG--FP-VTTD--KSDKAKALLEHEN-ISTPCAPV---PTP----
 52 ----YLRSIDEVPRMFNESKFAKNARDFTKNVQFPAPDTI----KLSSS--ACYEGP
 53 RSRG-GYLG---HIKKEGDDDAKDIRGVRPPPNDTHPLLFIKA
 54 RS---D
 55 YDP-----
 56
 57
 58
 59
 60

1
2
3 NLIS-----DLSPSRL--AEIMS
4 YYYY-----WNNCDKYM**K**
5 SGHRDEQVHRVST--IPERG-YKNRIVTAPPSGVLSV-GEVIR-----SIIFP-FV
6 RDHPSC-----TLVRDGDL-RGARF-----YDRTGC-----
7 -----RL-YSA-----DLTKATDGFSH--EAIEAVGRGLLS--YG-----
8 -FPTSVVEHFLGTLG-VG---SKHY--AEY--NVADLLQGHKEGSRKYNQVCSYLMEKG
9 WDGNS--KV-----MRVLMRRGSPMGTPCSFTLLCILNLWS----L---
10 DRA--C-----
11 -----NGP--RICGDD--MLAY--ITPHDF--AHYSRRISI-IGSGVHVTKSY
12 VS--DIA-GTFCENL--Y-----VRSGTQ--D---
13 -----
14 GIVSFERG-----
15 -----IRCPIKSVLFPQ----KGS----NGR
16 IDWPDPVHVQVFTGETRM-----EEPYRQ----VATVVP-----TEQGYVHTR
17 QRELHQGDFLFEDLPLTV--KIG-----
18 -----VRVARVL-----RTLRRDT----IRACLKKGRNP-
19 -----WAPACL-----GGLNYPRNAKGVK-----GFPKWFRANLY-----
20 -----SF-THSDLSVPE-----
21 -----LSKVLAQIDCE-----
22 -----KKPDSLSKELE--FVSTMVADTEINE-----SL
23 VPEI-----HDRVGDCAFVDMYALSRHYGGIKS-----NFFR
24 SMGGRLSV-----
25 -----ASSSYS-----
26 -----
27 >GAIR01012025.1
28 -----
29 -----
30 -----
31 -----
32 -----
33 -----SYMK-----NMKVRLLL**H**
34 LTRV-----DFKDLPAVPPVKGGARKPRS
35 -----TIKSRDPA*N*RRP
36 IQ----IVWE-----AVWAALVVADPSDG-----IR
37 DVSSLRGSFHLKD-WVVRSSH-----RSVKYTVEILK-----AEMTRLRAHSL---
38 -----LSTEIPVPHG-FPR-----KLDRFLRAAFAPMG-----
39 ---LFTLSTMTRG--FP-VSTD--PKQRSAELMGHMENVMSPPPSVP----AH----
40 -----YLRSSIEEFVPR---MFKESSIYKRSDDLRIVNPAPVCN---KTSSS--SCFEGP
41 RSKG-GYFG--YIR-----TKGDKAGEGKDIKEIREGTTLPPEKTCLPLFRGIAEKS
42 YT-----M
43 EFPA-----PLDEF---ADIIS
44 HYYF-----LNDCEDYLN
45 RDPSQLVHRIGV--VPERG-YKNRIVTSPPSGVLSA-GEVVR-----HVLFP-FV
46 KSHPSC-----EIVRDGDM-VNASF-----YNRPYC-----
47 -----SL-YSA-----DLTKATDGFSH--EAILAVAQGLKT--VG-----
48 -FPTS--VIDIFLST-LG---VSEI--H-Y-AEYSIKDLMAGLRVGSRKYKNRASELMS
49 RGWDGFTDK-----LRIAMRRGSPMGTPCSFTLLCILNLWA----L---
50 DAA--K-N-----
51 -----GPK---ICGDD--MLAY--ITPSEF--NSYSRRIAA-IGSGVHPTKTF
52 VS--RSA-GVFCENI--Y-----IRREDS-DGIACF
53
54
55
56
57
58
59
60

1
2
3 NRGVRIPIKSILF-----PQK-----
4 GSNGRIDW-----
5 -----PDAMVHQMFGETRLSESYSVVETLVPEEQGYKR--IRHRELT----DGD
6 RFYEDLP-HT-----MRIGVR-----IRRVL-----TLRRDAIRAC
7 L-KRN-RVPW---APACL--GGLN-----YPRKNPK-----
8 ----GVKDFP-AWFR----RDLYNFTHSDLGVGELSKVIAAIDCQKKPSKGSEELGM-
9 -----LSTMVA-----ENDINEAIRDDIGE-----DVGECSVSDVY-----
10 -----TL-ARAYGSARARF-----
11 -----FTSMGGTLDTK-----ASSYSAFC
12 VPKIDWPKGSGGYDCKTP---MELVIKLDNLLR-----
13 SPNY-----
14 RITRDFSV-----
15 -----
16 -----
17 -----
18 -----
19 >GGCO01105932.1
20 -----
21 -----YMK-----NMKVRLSSH
22 LTRV-----DFKDLPAVPPVKGGARKPRS
23 -----TIKS RD PAN RRP
24 IQ-----IVWE-----AVWAALVVADPSDG-----IR
25 DVSSLRGSFHLKD-WVVRSSH-----RSVKFTVEILK-----AEMTRLRALSL---
26 -----LSTEVPVPHG-FPR-----KLDRFLRAAFAPMG-----
27 ---LFTLSTMTRG---FP-VSTD--PKQRSAELMGHMENVMSPPSVP----AH----
28 ----YLR SIEEFVPR---MFKESSIYKRSDDPKIVNPAPVCN---KTSSS--SCFEGP
29 RSKG-GYFG---YIR-----TKGDKAGEGKDIAKEIREGTTLPPEKTCLPLFRGIAEKS
30 YT-----M
31 EFPA-----PLDEF---ADIIS
32 HYYF-----LNDCEDYLN
33 RDSSQQLVHRIGV--VPERG-YKNRIVTSPPSGVLSA-GEVVR-----HVLFP-FV
34 KSHPSC-----EIVRDGDM-VNASF-----YSRSSC-----
35 -----SL-YSA-----DLTKATDGFSH--EAILAVAQGLKTV-G-----
36 -FPTS--VIELFLST-LG---VSEI--H-Y-AEYSIKDLMAGL RVGSRKYNNRASELMS
37 RGWDG--FT-----DKLRIAMRRGSPMGT PCSFTLLCILNLWA----L---
38 DEA---K-N-----
39 -----GP--KICGDD--MLAY--ITPSEF--DSYSRRIAA-IGSGVHPTKTF
40 VS--RSA-GVFCENI---Y-----IRREDS-RGIACF
41 DRGVRIPIKSILF-----PQK-----
42 GSNGRIDW-----
43 -----PDTIVHQMFGETRLSESYSVVETLVPEEQGYKR--IRHRELT----DGD
44 RFYEDLP-HT-----KKIGVR-----IRRVL-----TLRRDAIRAC
45 L-KRN-RVPW---APACL--GGLN-----YPRKNPK-----
46 ----GVRDFP-AWFR----RDLYNFTHSDLGVGELSKVIAAIDCQKKPSKGSEELDM-
47 -----LSTMVA-----ENDINEAIRDDIGE-----DVGECSVSDIY-----
48 -----TL-ARAYGSARARF-----
49 -----FTSMGGTLDTK-----SSSYSAFC
50 VPKIDWPKGSGGYDCKTP---MELVIKLDNLLR-----
51
52
53
54
55
56
57
58
59
60

1
2
3 SPKYRITRDFSGKIQGAEHLKSSR-----WH
4 LLRGA-----
5 -----GGWND--GDL-----
6 -----
7
8 >GBTA01004534.1
9 -----
10 -----
11 -----LVPEI-----
12 -----
13 -----KEFT-----
14 -----LKIVSKMYQDWKSKEKFNE---QCEETGRKKINLDLSVPEWSPNLHASAC---
15 -----IENPIKNGQQAY-MHK-----
16 -----LFQVAHTGVG-----EARNARAFAMRIRN-----
17 -----VLQQEMIWLKSMGDWQVG-----NELPS-----TTD
18 RKEG-GVEE-----IPKDDFRFNTSMPE-----
19 -----
20 -----
21 -----GIPDYVS-----
22 -----GSY-----
23 -----LPVGRASV--VYEKG-MKARVVTVHAAFVAL-SCALN-----RTLIN-LL
24 KCYTPS-----KASLLGRD-IELKT-----LKRDGLE-----
25 -----FSA-----DLSTATDLIPF--EVARAVVDGICEA-LN-----
26 -----WGPH--IKEIMYYA-TG---PFRI--E-YDRA-----
27 -----PRSK-----DSIVPDYIISRRGILMGLGTSWPILSILNSFC----A---
28 EYN--C-----SKEGV-----
29 -----RNY--RICGDD--LAAL--WDKN DI--GNYIGNLES-VGLKLNDKKTV
30 TS--DQG-LVFTEEL---Y-----S--IKSRTH--INRKG
31 ERSPD-----QKD-----
32 LYLSRFKR-----
33 -----AIVSVI-----VDAKHNSTTKFTNPE-----ERL
34 PRYLTIG--DCSHAIDEFCP----LWQKSR-----AQKALL-----LMNKKCIR-E
35 I-FRA-GLDPF--LPKYL--GGAG-----VK-----
36 -----TVGNEKRHPLM----KHL SALIQLPYEDIQLVCRRRLQ-----RP-
37 -----WLSSSV-----NKETSAIFS-----
38 -----KT-LNLFEKL PRAEHGSAI
39 PYKEA-----VRKFTGAVLAR-----
40 -----YSLLNFGSCQTN--KVCTLKFIRKQVN-----KA
41 RKTL-----FSKAPRRYHYMSLLK-----FS
42 NALEKLEP-----
43 -----
44 -----
45 >GEWH01003494.1
46 -----
47 -----
48 -----
49 -----
50 -----
51 -----TLNS-----SVAEKSETG
52 KRPR-----SRMEKLLRRCTERDMTTRQL
53 -----
54 -----
55 -----LETL-----GLRLDSLQLGQDLVQAFSTLLDVER-----GN
56
57
58
59
60

1
2
3 KYFERRWFIRLAQ-WLLRAIDSFTDKVKAEGISSAMTWIK-----DALTHLRACAL---
4 -----DAGKVTDPFPSK-APR-----VERRRM-----
5 ---LVQASRCRSRS--LR--AAT--GDKLEQAAQSHKEI-LTSPFKTP----EP----
6 ----ELLSFREYVRA-----KTRASTVQVG-----EVVAS--ASLDTK
7 GSGG-GSLE---EVRLGLVD--SLRSRPFTYPDLLELDRTFPHEMLGGHVLLNRPKVQERF
8 RIGQVGHSSGTL-----L
9 GVPG-----DQVLFLYDQSCEISLDEWEFLREQLFA
10 IVAC-----VLDVDLDS-
11 -----RPACRQCV--IAERG-WKTRVVTPISAASNYL-LTVAN-----KGLLS-II
12 GEHREL-----DNSQRGRPCDNLDW-----SFGKRY-----
13 -----ALTFSA-----DLKSATDYMPH--DLCLVAAEELTRG-----
14 -WPAY--LRRRLVRS-VG---SYDL--R-C-----
15 -LDGG-----TVETRRGILMGVPTSWPILSLYSGWL-----
16 HYR-----SSSDGWF-----
17 -----AVCGDD--YLGC--HTYASY--HKYLRARER-TGAKGSPGKDI
18 LS--TESTGIFAEEL-----VTVGRC-----
19 -----
20 -----
21 -----
22 RVLRTVPI-----
23 -----RPLTGI-----PK-----GDT
24 PSWAMGEAISVLLSEA-----GFSEER-----SGRLVS-----RRFGSEMDK
25 L-RRA-GIDPK--GPRWALAGGFP-----GKPG-----
26 -----PRVL---ETARRVLS-----QGEDTIL-----GWLTALG-IA-
27 -----WSDAPS-----RDLIQGM-----
28 -----
29 -----
30 -----VRDKLGGLKGQ-----
31 -----
32 -----
33 -----
34 -----
35 -----TWRAE-----
36 -----
37 >GDUK01007882.1
38 -----
39 -----
40 -----KVEKWLP-----PFSVFGGSS
41 MASR-----SRPLHVGRIRDLRVEMSQSH
42 -----FPGASMTADQI
43 AQ---QFWN-----VVRSHLLAAFPLVSGLPVA-----KR
44 SGRNARFGLVLLQYHLLRRLHL-----GGVASLVRWLK-----ALAFSCRDSVV---
45 -----SDKKLPRE--FSR-----YFHGVNLPSALS-----
46 -----LTTLAKLGRALPECDAKMGG--GSVIRKALRQHRAD-MTSARVKQ---FGFN----
47 -----PMDRLDLAVAK---ETASFLLRGRLGSPRM-----HDLSS--ASADVS
48 SREG-GLAA---GLR-----PSWARMEEIGPVPWALGHALAEWQAALLRGARNTAWAT
49 AEEERG-----I
50 VLSNGSFLGPLVLLASPDGLWVTRRSEYIRSVLSFHEGLGPVRTGRFTTVQVDHDTV
51 QFWH-----VAGEPDLKRFFA
52 SCARDLPKAKVAL--VRERG-WKVRVVTSLTDQVIA-GHALR-----DLIWP-VL
53 QDDPCF-----DLTGDP-----RIEGMPVDQG-----
54
55
56
57
58
59
60

1
2
3 -----EVVLSA-----DLSRATDVVSI--GLALAVFDQWNRIDPR-----
4 -FDRQ--WHDLARAT-VS---GYRI--Q-Y-----
5 -PDGR-----YVRSSAGWLMGHPSTWALLCLVHKAI----A---
6 EMA-----GF-----
7 -----KKY--RIRGDD--LLAV--GSRESV--EEYFTLISK--YFEVNVKKSF
8 VS--PNG-GTFAEDT----F-----V--VRNGQL-----
9
10 -----
11 RLIAHAAP-----
12 -----PR--GWLAGE---KASLA
13 AAYTQFG--RL-----PRKWRM-----AYRHSL-----LSSGKHQIA
14 E-FRKHRIPLF--LPRVLP--GGVG--V--PH-----P-RGVGG-----
15 -----AVAGS-----VLVSKMVTG-----AGSVS-TA-
16 -----WVNRMA-----YRQREALYSEAREV-----TRMLHQDG-----
17 -----LG-SSTVLTTEEVIN-EITL
18 -----KSNIAMGCTWP-----
19 -----EGWGDKTLLGE--LR-----DPSYA----SI
20 GRDW-----ALFRAKCAGIRPPPH-KVGTGWS-----EQR-----LA
21 DAIRKSSL-----GG-----
22 -----MYWPY--RPDP-----PIRLS-----
23
24 ----->GEUE01057748.1
25
26 -----
27 -----QRMKQPAKPGKFYLP-----NLVESEVPK
28 VTLD-----LILGGKGGAVID-----LPTLSRVRPARPMPSAKGKS
29 -----ADEDNCNVLGNDNNNGSKRAKTRR
30 RD-----ESWE-----RSDKDANVFWGAWLWLDIASSAYERGLDLKNPSEN
31 RS-----AREALSVRG-WIRKSVHF-----NGVDFVQKEFK-----KLAHYRHRAV---
32 -----GDVRDPPAPGL-LDR-----HLSGIPRSVLKGPI--EGR
33 LRACAFIGRLGRC---GP--PPS--KKVEERALKEHRRV-LCRRFKTP----KV---
34 -----QLASIRKSTKA--WCKD--LNEVGPSVV-----AFSSA--ASVEFS
35 RLKG-GQTA---ELF-----DHAKGAIQMYQTLLYGEPIEGSNSTGLR-----
36
37 -----DEDIRIT--VQVAL
38 DSAI-----EKAKSRS-
39 -----PLRVKATV--VPETG-GKARVVTAGPADMvla-GNALR-----RAVWP-IV
40 MGDDSIDTEEAVAQEEVSAGDK-ELLAR-----LLGGGCS-----
41 -----W-YSG-----DLTAASDWLPR--DVCLTIWDAIAEA-AG-----
42 -----FAKGSPEYLLGRKL-LG---NVVI--E-Y-----
43 -----ADGT-----EVTTGAGALMGFPLTWLVLCAYNRAL----V---
44 SMV--S-----KKAI-----
45 -----QRT--VFRGDD--MVSQ--MTPEEG--ARYEELVRL-TGGQPNTAKSF
46 RS--VTG-WVFTEAT----Y-----Q--LTWHT--TL---
47 -----TRA-----
48 DLAERGV-----
49 -----LEHGLGRAGTLRAI-----TVRRVPDCSLRHLLRQR-----VAS
50 QPYLAIA--GPAVSGAVA-----PLLASRLADVRAVGNRLV--RRWMRVNARMVE-S
51
52
53
54
55
56
57
58
59
60

1
2
3 I-QRH-NLGLF--VPREL--GGFG---L---PH-----P-KGWAR-----
4 -----GVTSA--RWAH----KRLAVFFTAT----RGAAVKI-----GCLGYDP-
5 -----WVVDPN----TGRLWGEVSKRLASR-----EKSCLVLLPPE----RV
6 AEPTADRTPESWEQYVPN-----YG-GYYLAVVPGAGRGVAQ
7 PVTEAV-----AQEMALESAWR-----RGLGLDPVVVDKPQ
8 EKSRKSGTSDGCKSKAKKP--SRSALTIISKNHR-AMRKTGKSRFPEKFLPKRALRPE
9 KLQM-----ALMRHKMYVQHGSLFGGPSPLHIV-----FN
10 SEKGRGPP-----RRPGGWKYYH
11 DPKQGRSKTSSD--PVPRS--PAPPLVGGVPLDPAWENSWPSRKSGNTS-----
12 -----
13 -----
14 >GEUE01035234.1
15 -----
16 -----
17 -----
18 -----
19 -----
20 -----
21 -----
22 -----
23 -----
24 -----
25 -----
26 -----
27 -----
28 -----
29 -----
30 -----
31 -----KATV--VPETG-GKARVVTAGPADMVL-AGNALR-----RAVWP-IV
32 MGDDSIDTEEAVAQEEVSAGDK-ELLAR-----LLGGGCS-----
33 -----W-YSG-----DLTAASDWLPR--DVCLTIWDAIAEA-AG-----
34 -FAKGSPEYLLGRKL-LG---NVVI--E-Y-----
35 -ADGT-----EVTGAGALMGFPLTWLVLCAYNRAL---V---
36 SMV--S-K-----
37 -----KAIQRT--VFRGDD--MVSQ--MTPEEG--ARYEELVRL-TGGQPNTAKSF
38 RS--VTG-WVFTEAT----Y-----Q---LTTWHT--TL---
39 -----TRA-----
40 DLAERGV-----
41 -----LEHGLGRAGTLRAI-----TVRRVPDCSLRHLLRQR-----VAS
42 QPYLAIA--GPAVSGAVA-----PLLASRLADVRAVGPNRLV--RRWMRVNARMVE-S
43 I-QRH-NLGLF--VPREL--GGFG---L---PH-----P-KGWAR-----
44 -----GVTSA--RWAH----KRLAVFFTAT----RGAAVKI-----GCLGYDP-
45 -----WVVDPN----TGRLWGEVSKRLASR-----EKSCLVLLPPEVAEPTA
46 DRTPESWEQY-----VPNYG-GYYLAVVPGAGRGVAQ
47 PVTEAV-----AQEMALESAWR-----RGLGLDPVVVDKPQ
48 EKSRKSGTSDGCKSKAKKP--SRSALTIISKNHR-AMRKTGKSRFPEKFLPKRALRPE
49 KLQM-----ALMRHKMYVQHGSLFGGPSPLHIV-----FN
50 SEKGRGPP-----RRPGGWKYYHDPKQGRSKTSSDPVPRSPAPPL-----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
 2
 3 >GBYB01012090.1
 4 -----ASARLGRS
 5 GKWKLPNLPEPVATAVTPPAAAARLLSNLRKDAFPRGVGTGNYLS-----RLEAAGIPA
 6 CSVSRHPRGSEPHPLRKYTPGRGDPACSLTS-----LPHTPLPTSPSRPGRARGGP
 7 -----AGERPRKRKGHSRDPGGCPLGCGAAACRD
 8 LR-----RIWD-----SFRLCILSLIPRLQRSGLDSN-----RR
 9 DYRMAVSILKVRS-WLSINWRS-----SGRNETIRQIK-----AYAEFCRARALD--
 10 -----AEQSRRGAPRGFP-VRA-----FEHGLGSLLRGASS--GRT
 11 A--LAQLARLGRS---MP--PAD--QAILTAALEQHRRV-LTTADPGV---PLT----
 12 ----LLDGLFHWAASS--WSGS--RPASLRNSVVA-----SFSSS--ASEGMS
 13 RAKG-GQRA---ELE-----QLARAEIDAIRADIAGHPDTPGFFDVDEGDYAGLYFVE
 14 -----W
 15 DLQT-----VTDSAMARVREEAA
 16 DTRY-----
 17 -----AIPVRATA--LPELG-NKARVVTTAPPAHWGII-GDAMR-----KVLWP-LL
 18 ETDPRIDLSSGRRLDGAAASFH-DQVVK-----SLRGAA-----GQ-----
 19 -----WM-YSA-----DLTAATDLMPE--NVILALWHGVHLG-LG---IP-
 20 -EESF--FARAGDKI-LG---CVDV--S-Y-PDLAA-----
 21 --PGE--KP-----IVVRSMRGCMGLNSWFLNLNLNLA-----V---
 22 DIA---C-L-----GGIAALTDPQDGGLDEAEVRRIVGL-----
 23 -----APA--IVRGDD--LAAA--LTERQA--TAYEELIAA-TGGEANRAKSY
 24 RS--HSA-FVLAEKS---F-----L---VDREVR--PLS-G
 25 GQKRGFISFPAGLAVDAPMLA-----RDP-----
 26 RALEALGF-----
 27 -----GGPLTGKDLGGATWTDAAVEVVSCLTALQDIPVRHLIP--APSKEGL-----
 28 PVYVSLP---AAAADVLYAA---EEEGSP-----LFPGMC---RAVLSVNSEYVQ-R
 29 Y-RDY-RIPLL--LPREV--GGAG---F---PH-----P-GGFSK-----
 30 -----ALASGG-RGHW----LRATLRVTTY--GVEQRARRHL-----DE-DV-
 31 -----WRVDGS-----NSDRRAAARLIRAR-----EAAAMNSGQLG-----
 32 -----NMVPVPL-EDEV
 33 -----TREVAFESLWR-----DLFCPADRKERCTR
 34 RDRGVRLATISKRLSKADM--AARRLAFNDRFLV--RS-----LV
 35 RAEY-----QADVAMARGAESVYVPLSERDP--R-----LMPLSDDHPSLV
 36 GLHRRRYR-----AGRASSHGRNGCGGVSQVG
 37 GVP-----DNWES--RLDDDFVYGGLDNEAGISGHCLPKRSTLERAVRSRGQ
 38 GAVAVRTLRSGMVLADSIHYTESRSKRRGKPPTS-----
 39 >KF298275.1
 40 -----
 41 -----GVMIRSRRKRNHQRWRLP-----STLVLAP EG
 42 RITP-----RAHSQPNEAGGAQG-----AAQSNVAQVTAERAANKPTQ
 43 -----ANSQRTRHARAPKDPKEQM
 44 GQ----ATWD-----AFRLAIWALIPSLQKKGLALKG-----PR
 45 NLATVRGIQSVHK-WISVSMAR-----SGPEWTAKQVK-----EYSNYCRARSL---
 46 -----RDARTPCPRGFP-VKA-----FEVGLIKAISDPN--GGR
 47 PA-LAQLARLGRA---MP--VAT--DRVGRQALEKHRET-LSRATDVP----PA----
 48 -----LIARFGAWAEQ---WARQ--QADKTGLTKEGSRLFQAVET--KFSNS--AAEAFS

1
2
3 AASG-GQQA---ELR-----ENYQRTMVDVQAEYLLATEGFVWDPEVPTLGAEDSDLMN
4 IT-----L
5 CARE-----EAEDYYL--RRHRW
6 RGVD-----
7 ----GMPVRATI--VREQG-MKARVVTACPAWAVVC-GDACR-----KTLWP-LL
8 AADPRVDLSGARPVAEQLDRFN-DQIAL-----SLVGSI-----TP-----
9 ----EF-FSA-----DLTAATDLMPF--EVSNALWAGLCRG-LD--LL-
10 -PTSQ--ISRIGHAL-LG---PVEV--S-Y-PDLAA-----
11 --KGE--PI-----PRILSKQGCMMGLPLSWTILNLYNLAV----A---
12 DFA-----VVAPSVVRVGV-----
13 -----APV--IARGDD--LVAA--YLPEEA--DRYTQLLRE-SGGEVNVLKSF
14 RS--SDS-FVLAERT----F-----KVTLRD--AVPLP
15 VHRRGFELRARN-----THA-----
16 PLLTQVPQ-----PSGTSRSRP
17 NVLGGGPAAESILGKTRETRGKVVVSIKMFDDCPLRTLVGKG--PGYAAGA-----VI
18 PAYISVP--SAASASLA-----EFEGR-----FYPALA--QGLLSVHKSLVG-E
19 F-RRS-AIPVF--YPREL--GGGG--F--PH-----P-RGFGH-----
20 ----AVCSAG-LLGR-----KRAAFAMTC-----FTHKARK-----KHGLGDDP-
21 -----WASRHR-----LAELERARQQLSAE-----EARAQTRAASEVPGAPRV
22 IPDDWEPKV-----WD-QIKDLEPPRGA-RIPV
23 PMEDAV-----VQRAAHGETWA-----
24 ---SMILQEVGQQRDRYP--SLGTIA--KRLQA--IRKVTAERKFSNRFLPKRGKLS
25 RERL-----LERVAMARGGETRYVPYVDFGGGG-----MV
26 LVCRGFDT-----YAAARYEPES
27 Q-----PERYP--ERRQKLRPGETEGEEMVEVPVAVRTTRAGLGLADVAFVR
28 HGRVKRVKRRLPSQFPR-----
29 >KP642119.1
30 -----
31 -----MVLSGTCAQPNSGVASA-----VKAVASPV
32 GASA-----KLRKGPVVGTDR-----PKSSLSSPRGHVGRKPRETV
33 -----KPKPDPLQGQ
34 AQ----RAWD-----VWRALFALMPSLQKKGLDVRKAKPAKKPV---TA
35 QHATQRAIQVLYK-WLASSFVR-----SGPEWTCKQIK-----EYANYCRSRSL---
36 -----GDERTDAPRGFP-VKG-----FEAGFIKACIMDPN--AGG
37 TA-LAQLARVGRA---MP--MGT--TRVAVAALRKHRDT-LTSTTVVP----EE---
38 -----TQVNRLRLWAER--WTRD--RLAAGARAREGTAT-----SFSRA--ASASVS
39 AADG-GQLA---ELR-----QLPRFVDHRQMLRDLLLEEQGFDLDSDD-----
40 -----I
41 FLSA-----TCDEGDD--LQLLA
42 DCAI-----ETARDRCMEGS-
43 -----PIPMVATT--VLELG-MKARVVTKPPAWAVVA-GDACR-----QSVWP-LL
44 ESDRRIDLGSARPNTESLDRFH-DNLAH-----SLVGAATP-----
45 -----QF-FSA-----DLTAATDDAPL-----
46 -----RLLGRYL-LG---PVTY--G-Y-PTLAAIEHDR-----L
47 YQAGE-----PTEIVSVRGCMGMLPLSWTLLNLYNLAV----A---
48 DMA---C-----SPVGVVQVGV-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----APA--IARGDD--LVAA--IPPGEA--DRYEELIAL-TGGEANRLKSF
4 RS--ATA-FVLAERT-----FRVETQEIPDRVR
5 GLPRGWRW-----RRVSRGDPALP
6 PLLTREHQ-----
7 --PGLRGQPSRNLEGA----REVVGVRMSVDLPLRALIAGS---ASFAGG-----AAV
8 PTYISLP---PAATSCLS----EFEGRTR----FYPAVA---RGLLSVHRDLVS-E
9 M-RRS-AVPLF--YPREL--GGGG---F---PH-----P-GGFAGA-----
10 ----GLASAG-LLGL----QRAGLALTS----HGWAARK-----KAALLE-DP-
11 ----WVPRRQ-----NAQLNRAREQLLAS-----ERRAWAQSVLKRGTVAS
12 LEAAGLAERAKELKL-----ALATP-VPEYERVPRGA-RIAV
13 PLEDET-----IRQAARAGLWE-----
14 DAFLSPTVDDGGRSKTKYP---SLGDIARRLKRIR----EVTSKCKFDPRFVPNKDRVQ
15 RDSF-----FEKLRMLRGSETVLVPVDSRPVVRV-----TVWD
16 PARHRADP-----ELDNDDCGY--
17 -----PGRRR--LLRALDREEGGDPELVPDPVPPLLQRGAGATLGDLVVFR
18 PAKSRHGPKSRAVTGARREKRSRRTPQHP-----
21 >KP642120.1
22 -----
23 -----MLATGVMAKPQRPKVPNWFPLPKPSVGASAKLRKGPVV
24 GTDR-----IPKSSLSSPRGHVGRKPRKT
25 -----VKPKPDPLQGQ
26 AQ----RAWD-----VWRALFALMPSLQKKGLDARKAKLAKKPV---TA
27 QRATQRAIQVLYK-WLASSFVR-----SGPEWTCKQIK-----EYANYCRSRSL---
28 -----GDERTDAPRGFP-VKG-----FEAGFIKACIMDPNAGGTA
29 ---LAQLARVGRA---MP--MGT--TRVAVAALRKHRDT-LTSTTVVP---EG---
30 ----TQVNRLWAER---WTRD--RLAAGARAREGTTT-----SFSRA--ASASVS
31 AADG-GQLA---ELR-----QLPRFVDHRQMLRDLLLEEQGFLDLGDD-----
32 -----I
33 FLSA-----TCDEGDD--LQLLA
34 DCAI-----ETARDR
35 CMEGSPIPMVATT--VLELG-MKARVVTKPPAWAVVA-GDACR-----QSVWP-LL
36 ESDRRIDLSGARPNTESLDRFH-DNLAH-----SLVGAA-----TP-----
37 -----QF-FSA-----DLTAATDLMPF--SVWSLWDGLCDG-LG---AA-
38 -ADAP--LRLGRYL-LG---PVTV--G-Y-PTLAAMEEHDR-----L
39 YQAGE--PT-----EIVSVRGCMGMLPLSWTLLNLYNLA-----A---
40 DMA---C-----SPLGVVQVGV-----
41 -----APA--IARGDD--LVAA--IPPGEA--DRYEELIAL-TGGEANRLKSF
42 RS--ATA-FVLAERT---F-----R---VETQEI--PDRVR
43 GLPRGWRW-----RRVSRGDPALP
44 PLLTREHQ-----
45 --PGLKGQPSRNLEGA----REVVGIRMSVDLPLRALIAGS---ASFAGG-----AAV
46 PTYISLP---PAATSCLS----EFEGRTR----FYPAVA---RGLLSVHRNLVS-E
47 M-RRS-AVPLF--YPREL--GGGG---F---PH-----P-GGFAGA-----
48 ----GLASAG-LLGL----QRAGLALTS----HGWAARK-----KAALLE-DP-
49 ----WMPRRQ-----NAQLNRAREQLLAS-----ERRAWAQSVLKRGTVAS
50 LEAAGLAERAKELKL-----ALATP-VPEYERVPRGA-RVAV
51
52
53
54
55
56
57
58
59
60

1
2
3 PLEDET-----IRQAARAGLWE-----
4 DAFLSPAVDDGGRSKTKYP---SLGDIARRLKRIR----EVTSKCKFDPRFVPNKDRAQ
5 RDSF-----FEKLRLRGSETVLVPVDSRPVRV-----TVWD
6 PARHRADP-----ELDNDDCGY--
7 -----PGRRR--LLRALDREEGGDPELVPDPVPLLQRGAGATLGDLVVFR
8 PAKSRHGPKSRAFTGARREKRSRRTPQHP-----
9 >KF298284.1
10 -----
11 -----GVMAKPQRLFRSQQRWYLP-----AVLSLEERS
12 GVTMPLGTFPVKTGVAKGDPCHP---VPQGS-----AKRANRHPNSPPTRRPKPER
13 -----RARTQRTPTQPGVELQRQ
14 AQ----HIWD-----TFRLTIWVLLPSAQKKGIA-----VS
15 AKATIRSIKGVD-WIRNSMIR-----SGVEWTSKAIK-----EYANYCRARAL---
16 -----GDDRVGRPRGFV-VRT-----FERGYIAAVKCPG--QGR
17 LH-LAQIARIGRA---MP--KGT--RNVGIISSLRKHRGV-LSQ--PME---TDKD----
18 -----LLENLRSWATQ---WAQE--RVSERGTVVNPSV-----SFSRS--ASATVS
19 AMSG-GQLA---ELS-----QLESVKAHHRLIRELLLEEQGFEWNPAGTSDDS-----
20 -----I
21 GLLT-----VADEVAD--LQILA
22 DCAI-----SDALDAAREQI-
23 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GNACR-----KTVWP-LL
24 EADRRIDLSGLRPTAEVLDEFH-DNLAH-----SLVGAR-----TT-----
25 -----QF-YSA-----DLTAATDLMPF--DVSRALWNGLCDG-LG---AS-
26 -PSAP--LRSLGSYL-LG---PVRV--S-Y-PDLAALKERDR-----L
27 FKAGE--ET-----EILSSRGCMGLPVSWTVLNLYNLAM-----A---
28 DMA---C-----TPAGVPTLLGV-----
29 -----APA--VARGDD--LVAA--IPEFAA--DRYENLIAA-TGGEANRLKSF
30 RS--TFA-FVLAERT---F-----EVG-VKSVKDLTALTRR
31 GFKVRQVSRAP-----RTA-----
32 PLLCRFDE-----
33 ---LGPQQIPAFIQEM--DRPVEVVALRMSCDLPPIRSLLGGM---AGFAGG-----EAV
34 PNYVAIP---SAAAACLA-----EFEGR-----LYGAIC---HGLLSVHRDLVK-E
35 M-RRS-AIPLF--YPRVL--GGGG---F---PH-----P-RGFAA-----
36 -----AVASAG-PLGL---QRAGLYLTA-----SGYKARR-----KVGIGH-DP-
37 -----WVPLAD-----SRRMDVARDRLVAT-----DRRLWAKSVLDRLPDLLD
38 LEDPYLDKRLSERAE-----ALSVP-CVGFETVPRGA-RVAV
39 PLDDAV-----ISEAATIAAWE-----
40 ---DAVLGPCSAERRYP--TLGGIAKRLKAIR----QKTVEARFPMKFVPNMKRMG
41 REGF-----FNRVKLLLASETVLVPYSNRNTAR-----VTVFD
42 PERHCTVQ-----EPEDEGRVEGDLG
43 L-----PIRRL--ELRERHGGGVRVLRDGATLGDVAVIRHVPQKRRRHRR
44 PGPSNRNVL-----
45 >KF298276.1
46 -----
47 -----GVMAKPQRLFRSQQRWYLP-----AVLSLEGRS
48 GVTMPLGTLPKAAGRVAKGVPCHPDPQGS-----AKRANQHPNSPSTRRPNPER
49
50
51
52
53
54
55
56
57
58
59
60

1
 2
 3 RARAQRTPTQPGVELQRQ
 4 AQ---HIWD-----TFRLLTIWVLLPSAQKKGIA-----VS
 5 AKATIRSIKGVYE-WIRNSMIR-----SGVEWTSKAIK-----EYANYCRARAL---
 6 -----GDDRVGRPRGF-TGL-----FERGYIAAALKCPG--QGR
 7 LH-LAQLARVGRA---MP--KGT--RNVGISSLRKHRGV-LSQPMETD----KD----
 8 ----LLENLRSWATQ---WARE--RVSERGTVVNPSV-----SFSRS--ASATVS
 9 AMSG-GQLA---ELS-----QLESVKAHHRLIRELLLEEQGFEWNPAGTSDDS---
 10 -----I
 11 GLLT-----VADEVAD--LQILA
 12 DCAI-----SDALDA
 13 AREQIPIPMTATV--ISELG-MKARVVTKPPAWAVVA-GNACR-----KTVWP-LL
 14 EADRRIDLGLRPTAEVLDEFH-DNLAH-----SLVGAR-----TT-----
 15 ----QF-YSA-----DLTAATDLMPF--DVSRALWNGLCDG-LG--AS-
 16 -PSAP--LRSLGSYL-LG---PVRV--S-Y-PDLAALKERDR-----L
 17 FKAGE--ET-----EILSSRGCMGLPVSWTVLNLYNLAM---A---
 18 DMA---C-----TPMGVPTLLGV-----
 19 -----APA--VARGDD--LVAA--IPEFAA--DRYENLIAA-TGGEANRLKSF
 20 RS--TFA-FVLAERT---F-----EVGVKSVKDL-TALTRR
 21 GFKVRQVSRAP-----RTA-----
 22 PLLCRFDE-----
 23 LGPQQIPAFIQEMDRP-----VEVVALRMSCDLPIRSSLGGM---AGFAGG-----EAV
 24 PNYVAIP--SAAAACLA-----EFGTR-----LYGAIC--HGLLSVHRDLVK-E
 25 M-RRS-AIPLF--YPRVL--GGGG---F---PH-----P-RGFAA-----
 26 ----AVASAG-PLGL---QRAGLYLTA---SGYKARR-----KVGIGH-DP-
 27 -----WVPLAD-----SRRMDVARDRLVAT-----DRRLWAKSVLDRLPDLLD
 28 LEDPYLDKRLSERAE-----ALSVP-CVGFETVPRGA-RVAV
 29 PLDDAV-----ISEAATIAAWE-----
 30 ---DAVLGPCPSAERRYP--TLGGIAERLKAIR-----QKTVEARFPMKFVPNMKRMG
 31 REGF-----FNRVKLLLASETVLVPYSNRNTAR-----VTVFD
 32 PERHCTVQ-----EPEDEGRVEGDLG
 33 L-----PIRRR--ELRERHGGGVRVLRDGATLGDVAVIRHVPQKRRRHRRT
 34 PGPSNRNVL-----
 35 >MF176344.1
 36 -----MSLPERVGGTIPRTPAGSREP-----ASAQVGRPQ
 37 PSPK-----GLARK--TNHTQ-----IPPTIVDPKSSPRRARPPKK
 38 -----ACAPGKATTKPPASAQQKQ
 39 AQ---HIWD-----AFRLLSLWVLLPRLQKHGLS-----TK
 40 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
 41 -----KAQQCEPPRGFP-IRA-----FERGYIRAMLACPG--QGR
 42 LH-LAQLARVGRA---MP--IGT--AKVAIASLRKHREV-LSQ--PMD--TEPA----
 43 ----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
 44 ALKG-GQLT---ELR-----QLPAVAEHYALIQELLGDQLDPV-TGEPFVWDPTYTFDAD
 45 GGA-----I
 46 GLLT-----SASEADD--LQILA
 47 DCAL-----RTATEHAADNT-

1
2
3 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
4 EGDRRIDLSGVRPTAEVLDTFH-DNLAH-----SLVGAR-----TT-----
5 ----QF-YSA-----DLTAATDLMPF--DVSRAMWDGLCDG-LG---AT-
6 -ATAP--LRTLGRYL-LG---PVQV--S-Y-PDLRALPASSK-----L
7 YVAGE--RV-----ECLSKRGCMGLPVSFTVLNLYNLAM----A---
8 DLA---C-----TPEGSPVLVNV-----
9 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
10 RS--ADA-FVLAERT---F-----EVGVLRPNV--ELRQR
11 GYVVRRTY-----RTA-----
12 PLLAQFDS-----
13 AELHGGTGDARRLGG-----PEVVAIRMACDVPIRSLLGGG--PRTAGA-----NPV
14 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHRLVA-D
15 L-RRS-AVPLF--YPREL--GGAG--F--PH-----P-KGFAA-----
16 ----AVASAG-ELGL----KRAGLALT-----AGHKAQQ-----KVGLLS-CP-
17 -----WSPKLT-----DRRAKEARRRIAS-----DQRAWANSVLARHPDLPR
18 GGSAWTDRAAGELRK-----ALSIP-ARGFETVPRGA-RIAV
19 ALEDAV-----IREVAAAAAWE----DAFLPGPARGWSGVPPPMQ
20 EPSRKRSARRKEESSGMYP--SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
21 RARF-----FDRLKALRELETVLVPAQSAGARL-----VVFD
22 PVTHRVE-----SAGEIDGGEGRMG
23 F-----PLWRH--TLQERKGKRGAQGQEAAPVPVATTRSGLSLGDVMAVR
24 RLPRRPRGVRRLPNPGGAHSRNVR-----
25 >MF176385.1
26 -----MSLPERVGGTIPPRTPAGSHEP-----ASAQVGRPQ
27 PSPK-----GLARK---TNHTQ-----IPPTIVDPKSSPRRVHPPKK
28 -----VCAPGKATTKPPASAPQKQ
29 AQ-----HIWD-----AFRLSLWVLLPRLQKHGLS-----TK
30 AVTTIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
31 -----KAQQCEPPRGFP-IRA-----FERGYIHAMLACPG--QGR
32 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPA----
33 -----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
34 ALKG-GQLT---ELR-----QLPAVAEHYALIQELLGDQLDPV-TGEPFVWDPYTFDAD
35 GGA-----I
36 GLLT-----SASEADD--LQILA
37 DCAL-----RTATEHAADDT-
38 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
39 EGDRRIDLSGVRPTAEVLDTFH-DNLAH-----SLVGAR-----TT-----
40 ----QF-YSA-----DLTAATDLMPF--DVSRAIWGDGLCDG-LG---AT-
41 -ATAP--LRTLGRYL-LG---PVQV--S-Y-PDLRALPASSK-----L
42 YAAGE--RV-----ECLSKRGCMGLPVSFTVLNLYNLAM----A---
43 DLA---C-----TPEGSPVLVNV-----
44 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
45 RS--ADA-FVLAERT---F-----EVGVLRPNV--ELKQR
46 GYVVRRTY-----RTA-----
47 PLLAQFDS-----
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 AELHGGTGDARRLGG-----PEVVAIRMACDVPIRSLLGGG---PRTAGA-----NPV
4 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA---EGLMSVHRGLVA-D
5 L-RRS-AVPLF--YPREL--GGAG---F---PH-----P-KGFAA-----
6 -----AVASAG-ELGL----KRALGLTT----AGHKAQQ-----KVGLLS-CP-
7 -----WSPKLT-----DRRAKEARRRLIAS-----DQRAWANSLARHPDLPR
8 GGSAWTDRAAGELRK-----ALSIP-ARGFETVPRGA-RVAV
9 ALEDAD-----IREVAAAAAAWE----DAFLGPARGWSGVPPPMQ
10 EPSRKRSARRKEESSGMYP---SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
11 RARF-----FDRKLALRELETVLVPAQSAGARL-----VVFD
12 PVTBHRVSE-----SAGEIDGGEGRMG
13 F-----PLWRH--TLQERKGKRGAQGQEAAPVPVATTRSGLSLGDVMTVR
14 RLPRRPRGVRRLPNPGGAHSRN-----
15 >MF176278.1
16 -----
17 -----MSLPERVGRTIPPRTPAGSHEP-----ASAQVGRPQ
18 PSPK-----GLARK---TNHTQ-----IPPTIVDPKSSPRRARPPKK
19 -----VCAPGKATTKPPASAQQKQ
20 AQ-----HIWD-----AFRLSLWVLLPRLQKHGLS-----TK
21 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
22 -----KAQQCEPPRGFP-IRA-----FERGYIHAMLACPG--QGR
23 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPA---
24 ---LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
25 ALKG-GQLT---ELR----QLPAVAEHYALIQUELLGDQLDPV-TGEPFVWDPYTFDAD
26 GGA-----I
27 GLLT-----SASEADD--LQILA
28 DCAL-----RTATEHAADNT-
29 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
30 EGDRRIDLSGVRPTAEVLDTFH-DNLAH-----SLVGAR-----TT-----
31 -----QF-YSA-----DLTAATDLMF--DVSRAMWDGLCDG-LG---AT-
32 -ATAP--LRTLGRYL-LG---PVQV--S-Y-PDLRALRASSK-----L
33 YVAGE--RV-----ECLSKRGCMGLPISWTVLNLYNLAM---A---
34 DLA---C-----TPEGSPVLVNV-----
35 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
36 RS--ADA-FVLAERT---F-----EVGVLRRPNV--ELRQR
37 GYVVRRTY-----RTA-----
38 PLLAQFDS-----
39 AELHGGTGDARRLGG-----PEVVAIRMACDVPIRSLLGGG---PRTAGA-----NPV
40 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA---EGLMSVHRGLVA-D
41 L-RRS-AVPLF--YPREL--GGAG---F---PH-----P-KGFAA-----
42 -----AVASAG-ELGL----KRALGLTT----AGHKAQQ-----KVGLLS-CP-
43 -----WSPKLT-----DRRAKEARRRLIAS-----DQRAWANSLARHPDLPR
44 GGSAWTDRAAGELRK-----ALSIP-ARGFETVPRGA-RVAV
45 ALEDAD-----IREVAAAAAAWE----DAFLGPARGWSGVPPPMQ
46 EPSRKRSARRKEESSGMYP---SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
47 RARF-----FDRKLALRELETVLVPAQSAGARL-----VVFD
48 PVTBHRVSE-----SAGEIDGGEGRMG
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 F-----PLWRH--TLQERKGKRGAQGQEAAVPVPVATTRSGLSLGDVMTVR
4 RLPRRPRGVRRLPNPGGAHSRNVR-----
5 >MF176306.1
6 -----
7 MSLPERVGRTIPPRTPAGSHEP-----ASAQVGRPQ
8 PSPK-----GLARK--TNHTQ-----IPPTIVDPKSSPRRAHPPKK
9 -----VCAPGKATTKPPASAQQKQ
10 AQ----HIWD-----AFRLSLWVLLPRLQKHGLS-----TK
11 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
12 -----KAQQCEPPRGFP-IRA-----FERGYIHAMLACPG--QGR
13 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPT----
14 ----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
15 ALKG-GQLT---ELR-----QLPAVAEHYALIQUELLGDQLDPV-TGEPFVWDPYTFDAD
16 GGA-----I
17 GLLT-----SASEADD--LQILA
18 DCAL-----RTATEHAADNT-
19 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
20 EGDRRIDLSGVRPTAEVLDTFH-DNLAH-----SLVGAR-----TT-----
21 -----QF-YSA-----DLTAATDLMPF--DVSRAMWDGLCDG-LG---AT-
22 -ATAP--LRTLGRYL-LG---PVQV--S-Y-PDLRALPASSK-----L
23 YVAGE--RV-----ECLSKRGCMGLPISWTVLNLYNLAM----A---
24 DLA--C-----TPEGSPVLVNV-----
25 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
26 RS--ADA-FVLAERT---F-----EVGVLRRPNV--ELRQR
27 GYVVRRTY-----RTA-----
28 PLLAQFDS-----
29 AELHGGTGDARRLGG-----PEVVAIRMACDVPIRSLLGGG---PRTAGA-----NPV
30 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHRGLVA-D
31 L-RRS-AVPLF--YPREL--GGAG---F---PH-----P-KGFAA-----
32 -----AVASAG-ELGL-----KRALGLALT-----AGHKAAQ-----KVGLLS-CP-
33 -----WSPKLT-----DRRAKEARRRLIAS-----DQRAWANSVLARHPDLPR
34 GGSAWTDRAAGELRK-----ALSIP-ARGFETVPRGA-RVAV
35 ALEDAV-----IREVAAAAAWE----DAFLGPARGWSGVPPPMQ
36 EPSRKRSARRKEESSGMYP---SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
37 RARF-----FDRLKALRELETVLVPAQSAGARL-----VVFD
38 PVTHRVSE-----SAGEIDGGEGRMG
39 F-----PLWRH--TLQERKGKRGAQGQEAAVPVPVATTRSGLSLGDVMTVR
40 RLPRRPRGVRRLPNPSGAHSRNVR-----
41 >MF176257.1
42 -----
43 MSLPERVGRTIPPRTPAGSHEP-----ASAQVGRPQ
44 PSPK-----GLARK--TNHTQ-----IPPTIVDPKSSPRRVHPPKK
45 -----VCAPGKATTKPPASARQKQ
46 AQ----HIWD-----AFRLSLWVLLPRLQKHGLS-----TK
47 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
48 -----KAQQCEPPRGFP-IRA-----FERGYIHAMLACPG--QGR
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPA----
4 -----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
5 ALKG-GQLT---ELR----QLPAVAEHYALIQELLGDQLDPV-TGEPFVWDPYTFDAD
6 GGA-----I
7 GLLT-----SASEADD--LQILA
8 DCAL-----RTATEHAADNT-
9 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
10 EGDRRIDLSGVRPTAEVLDTFH-DNLAH-----SLVGAR-----TT-----
11 -----QF-YSA-----DLTAATDLMPF--DVSRAMWDGLCDG-LG---AT-
12 -ATAP--LRTLGRYL-LG---PVQV--S-Y-PDLRALPASSK-----L
13 YVAGE--RV-----ECLSKRGCMGLPWSVTVLNLYNLAM----A---
14 DLA--C-----TPEGSPVLVNV-----
15 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
16 RS--ADA-FVLAERT----F-----EVGVLRPNV--ELRQR
17 GYVVRRTY-----RTA-----
18 PLLAQFDS-----
19 AELHGGTGDARRLGG-----PEVVAIRMACDVPIRSLLGGG--PRTAGA-----NPV
20 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHRGLVA-D
21 L-RRS-AVPLF--YPREL--GGAG---F--PH-----P-KGFAA-----
22 -----AVASAG-ELGL----KRALGLTT----AGHKAQQ-----KVGLLS-CP-
23 -----WSPKLT-----DRRAKEARRRLIAS-----DQRAWANSVLARHPDLPR
24 GGSAWTDRAAGELRK-----ALSIP-ARGFETVPRGA-RVAV
25 ALEDAV-----IREVAAAAWE----DAFLGPARGWSGVPPPMQ
26 EPSRKRSARRKEESSGMYP---SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
27 RARF-----FDRLKALRELETVLVPAQSAGARL-----VVFD
28 PVTHRVSE-----SAGEIDGGEGRMG
29 F-----PLWRH--TLQERKGKRGQAQQEEAVPVPVATTRSGLSLGDVMTVR
30 RLPRRPRGVRRLPNGGAHSRNVR-----
31 >MF176365.1
32 -----
33 MSLPERVSGTIPPRTPAGSHEP-----ASAQVGRPQ
34 PSPK-----GLARK---TNHTQ-----IPPTIVDPKSSPRRVHPPKK
35 -----VCAPGKATTKPPASAQQKQ
36 AQ----HIWD-----AFRLSLWVLLPRLQKHGLS-----TK
37 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK----EYANYCRARAL---
38 -----KAQQCEPPRGFP-IRA-----FERGYIHAMLACPG--QGR
39 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPA----
40 -----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
41 ALKG-GQLT---ELR----QLPAVAEHYALIQELLGDQLDPV-TGEPFVWDPYTFDAD
42 GGA-----I
43 GLLT-----SASEADD--LQILA
44 DCAL-----RTATEHAADNT-
45 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
46 EGDRRIDLSGVRPTAEVLDTFH-DNLAH-----SLVGAR-----TT-----
47 -----QF-YSA-----DLTAATDLMPF--DVSRAMWDGLCDG-LG---AT-
48 -ATAP--LRTLGRYL-LG---PVQV--S-Y-PDLRALPASSK-----L

1
2
3 YVAGE--RV-----ECLSKRGCMGLPVS梧LNLYNLAM----A---
4 DLA--C-----TPEGSPVLVNV-----
5 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
6 RS--ADA-FVLAERT----F-----EVGVLRPNV--ELRQR
7 GYVVRRTY-----RTA-----
8 PLLAQFDS-----
9 AELHGGTGDARRLGGA----PEVVAIRMACDVPIRSLLGGG--PRTAGA----NPV
10 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHRGLVA-D
11 L-RRS-AVPLF--YPREL--GGAG---F--PH-----P-KGFAA-----
12 ----AVASAG-ELGL----KRALGLTT----AGHKAQQ-----KVGLLS-CP-
13 -----WSPKLT-----DRRAKEARRRLIAS-----DQRAWANSVLRHPDLPR
14 GGSAWTDRAAGELRR-----ALSIP-ARGFETVPRGA-RVAV
15 ALEDAV-----IREVAAAAWE----DAFLGPGARGWSGVPPPMQ
16 EPSRKRSARRREESSGMYP--SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
17 RARF-----FDRLKALRELETVLVPAQSAGARL-----VVFD
18 PVTHRVSE-----SAGEIDGGEGRMG
19 F-----PLWRH--TLQERKGKRGAQGQEAAPVPVATTRSGLSLGDVMTVR
20 RLPRRPRGVRRRLPNPSGAHSRNVR-----
21 >KX883461.1
22 -----
23 MSLPEGVSGAIPPRTPAGSHEP-----ASAQVGRPQ
24 PGPE-----GLARK--TNHTQ-----IPPTIVDPKSSPRRAHPPKK
25 -----ARAPGKATTKPPASVKQKQ
26 AQ----HIWD-----AFRLSLWVLLPRLQKHGLS-----TK
27 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
28 -----RAQQCEPPRGFP-IRA-----FERGYILAMLACPG--QGR
29 LH-LAQLARVGRA--MP--IGT--TKVAIASLRKHREV-LSQ--PMD--TEPA----
30 ---LVEKLRAWAEI--WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
31 ALKG-GQLT--ELR----QLPAVAEHYALIQUESTGDQLDPV-TGEPFVWDPYTFDAD
32 GGA-----I
33 GLLT-----SASEADD--LQILA
34 DCAL-----RTATEHAADNT-
35 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
36 EGDRRIDLGVVRPTAEVLDTFH-DNLAH-----SLVGAR-----ST-----
37 -----QF-YSA-----DLTAATDLMPF--DVSRAMWNGLCDG-LG--AT-
38 -ATAP--LRTLGLYL-LG--PVQV--S-Y-PDLRALPASSK-----L
39 YVAGE--RV-----ECLSKRGCMGLPVS梧LNLYNLAM----A---
40 DLA--C-----TPEGSPVLVNV-----
41 -----APA--IARGDD--LVAA--IPAGEA--TRYEDLIAA-TGGEANRLKSF
42 RS--ADA-FVLAERT----F-----EVGVLRRSNV--ELKQR
43 GYVTRRTF-----RTA-----
44 PLLAQFDS-----
45 AELHGGTGDARRLGDA----PEVVAIRMACDVPIRSLLGGG--PRTVGA----NPV
46 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHRGLVA-D
47 L-RRS-AIPLF--YPREL--GGAG---F--PH-----P-KGFAA-----
48 ----AVASAG-ELGL----KRALGLTT----AGHKAQQ-----KVGLLS-CP-
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----WSPKLT-----DRRAREARRRLIAS-----DQRAWANSVLARHPDLPR
4 GGSAWTDRAAGELRK-----ALSVP-AREFETVPRGA-RVAV
5 ALEDAV-----IREVAAAAWE----DAFLGPGARGWSGVPPPTQ
6 EPSRKRSARQKEESLGMYP--SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
7 RARF-----FDRLKALRELETVLVPAQAAGARL-----VVFD
8 PVTHRVP-----SAGEIDDGEGRMG
9 F-----PLWRH--TLQERKGKRGQAQQQEAAVPVPVATTRSGLSLGDVMTVR
10 RLPRRPRGVRRRLPNRSGAHSRNVR-----
11 >KX883537.1
12 -----
13 -----MSLPEGVSGAIPPRTPAGSHEP-----ASAQVGRPQ
14 PSPK-----GLARK---TNHTQ-----IPPTIVDPKSSPRRAHPPKK
15 -----ARAPGKATTKPPASVQQKQ
16 AQ----HIWD-----AFRLSLWVLLPRLQKRGLS-----TK
17 AVATIEAIKALYS-WIGNSVAR-----SGVEFTAKQVK-----EYANYCRARAL---
18 -----RAQQCEPPRGFP-IRA-----FERGYILAMLACPG--QGR
19 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPA----
20 -----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
21 ALKG-GQLT---ELR-----QLPAVAEHYALIQUELLGDQLDPV-TGEPFVWDPYTFDAD
22 GGA-----I
23 GLLT-----SASEADD--LQILA
24 DCAL-----RTATEHAADNT-
25 -----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
26 EGDRRIDLSGVRPTAEVLDLFH-DNLAH-----SLVGAR-----ST-----
27 -----QF-YSA-----DLTAATDLMPF--DVSRAMWNGLCDG-LG---AT-
28 -ATAP--LRKLGLYL-LG---PVQV--S-Y-PDLSALPASSK-----L
29 YVAGE--RV-----ECLSERGCMMGLPISWTVLNLYNLAM-----A---
30 DLA--C-----TPEGSPVLVNV-----
31 -----APA--IARGDD--LVAA--IPAEEA--TRYEDLIAA-TGGEANRLKSF
32 RS--ADA-FVLAERT---F-----EVGVLRRPNV--ELKQR
33 GYVTRRTY-----RTA-----
34 PLAQFDS-----
35 AELHGGTGDAARRLGDA----PEVVAIRMACDVPIRSLLGGG---PRTVGA-----NPV
36 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHRGLVA-D
37 L-RRS-AIPLF--YPREL--GGAG---F---PH-----P-KGFAA-----
38 -----AVASAG-ELGL----KRALGLALT-----AGHKAQQ-----KVGLLS-CP-
39 -----WSPKLT-----DRRAREARRRLIAS-----DQRAWANSVLARHPDLPR
40 GGSAWTDRAAGELRK-----ALSVP-ARGFETVPRGA-RVAV
41 ALEDAV-----IREVAAAAWE----DAFLGPGARGWSGVPPPMQ
42 EPSRKRSARQKEESLGMYP--SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
43 RARF-----FDRLKALRELETVLVPAQAAGARL-----VVFD
44 PVTHRVS-----SAGEVDGEGRMG
45 F-----PLWRY--TLQERKGKRGQAQQQEAAVPVPVATTRSGLSLGDVMTVR
46 RLPRRPRGVRRRLPNRSGAHSRNVR-----
47 >KX883538.1
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 MSLPEGVSGAIPPRTPAGSHEP-----ASAQVGRPQ
4 PSPK-----GLARK---TNHTQ-----IPPTIVDPKSSPRRAHPPKK
5 -----ARAPGKATTKPPASVQQKQ
6 AQ---HIWD-----AFRLSLWVLLPRLQKRGQLS-----TK
7 AVATIEAIKALYG-WIGNSVAC-----SGVEFTAKQVK-----EYANYCRARAL---
8 -----RAQQCEPPRGFP-IRA-----FERGYIHAMLACPG--QGR
9 LH-LAQLARVGRA---MP--IGT--TKVAIASLRKHREV-LSQ--PMD---TEPA---
10 ----LVEKLRAWAEI---WTRE--RLGAGGRTLTNPAA-----HFSRS--ASATVS
11 ALKG-GQLT---ELR-----QLPAVAEHYALIQELLGDQLDPV-TGEPEFWDPYTFDAD
12 GGA-----I
13 GLLT-----SASEADD--LQILA
14 DCAL-----RTATEHAADNT-
15 ----PIPMTATV--ISELG-MKARVVTKPPAWAVVA-GDACR-----KTVWP-LL
16 EGDRRIDLSGVRPTAEVLDAPH-DNLAH-----SLVGAR-----ST-----
17 ----QF-YSA-----DLTAATDLMPF--DVSRAMWNGLCDG-LG---AT-
18 -ATAP--LRKLGLYL-LG---PVQV--S-Y-PDLSALPASSK-----L
19 YVAGE--RV-----ECLSERGCMMGLPVSWTVLNLYNLAM----A---
20 DLA--C-----TPEGSPVLVNV-----
21 -----APA--IARGDD--LVAA--IPAEEA--TRYEDLIAA-TGGEANRLKSF
22 RS--ADA-FVLAERT----F-----EVGVLRRPNV--ELKQR
23 GYVTRRTY-----RTA-----
24 PLLAQFDS-----
25 AELHGGTGDTRRLGDA----PEVVAIRMACDVPIRSLLGGG---PRTVGA-----NPV
26 PDYVSIP---PAAAACLA-----EFEGR-----LYRSVA--EGLMSVHTRGLVA-D
27 L-RRS-AIPLF--YPREL--GGAG---F---PH-----P-KGFAA-----
28 -----AVASAG-ELGL-----KRALGLATT-----AGHKAQQ-----KVGLLS-CP-
29 -----WSPKLT-----DRRAKEARRRLIAS-----DQRAWANSVLARHPDLPR
30 GGSAWTNRAAGELRK-----ALSVP-SRGFETVPRGA-RVAV
31 ALEDAV-----IREVAAAAAWE----DAFLGPGARGWSGVPPPMQ
32 EPSRKRSARQKEESLGMYP---SLGEVARRLKTIR----SQTQAARFNRTFLPKRAVQE
33 RARF-----FDRLKALRELETVLVPAQAAAGARL-----VVFD
34 PVTTHRSE-----SAGEIDGGEGRMG
35 F-----PLWRH--TLQERKGKRGQAQQQEAAVPVPVATTRSGLSLGDVMTVR
36 RLPRRPRGVRRLPNRSGAHSRNVR-----
37 >KX883539.1
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 -----
47 -----
48 -----
49 -----MGREETVKRVK-----QYAEYCRARSID--
50 -----ARQAEAGPPRGFP-FHA-----LEGGSFKALLRCPV--QGP
51 AA-LAQVARLGRS---MP--LPT--HKQQGRALIEHCNE-LTAGPHEE--IPVE---
52 -----LLDAFYYWTRT---WGAG--KGAALRDEVT-----NFSSA--AAESVS
53 RARG-GQRE---ELR-----LLAAPHFEAIREALAIDDIDFRVEEGDYAGI-----
54 -----Y
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 SEED-----DLEIITQAAIERVR
4 ADIR-----AGR-
5 ----RPDVRAAV--VPELG-AKSRIVTAGEAHWVVI-GDAIR-----KCLWP-AL
6 EAEPRIDLSSGRDVAGTA AHLH-DEVAK-----ALTGAE-----AV-----
7 ----EF-YSA-----DLTAATDLMPH--DLIQAIWRGLMEG-LG---VS-
8 -DDDL--MFTAGLAL-LG---PVNL-H-H-PDLD----D-----
9 LDRPG--KY-----AYGVTHKGCMGFPLSWYILNLYNLAC----A---
10 DLA--L-----SAERGRFLAEVAPRLVGA-----
11 -----APA--VVRGDD--LCSA--HTPAEA--DRYERIIRA-TGGRANLSKSY
12 RS--RKG-FILAERT----F-----L---VTTART--KCSAA
13 RSRLGTVVVRPGVH-----PRA-----
14 PLLTRAPD-----AFES
15 GALGYRSIDPRSRRRG---AATVVGLTMLSDVPVRHLMPSP-----DS----EGL
16 PAYVTLP---PAASAVVD----EIKDIK-----RREAIC--YGILSVNAPLVR-Q
17 F-NRF-KIPIF--YPREL--GGAG---F---PH-----P-KGFGS-----
18 ----AIRSAG-GMGH---LRATLAMTTFTRRARQAYQLDR-----DP-
19 ----WICDRQ----VEFERSRAQTGLIAE-----ERRAAESGQLG-----
20 -----LATSVP-LEDAVAARA AEGQLWA
21 EVMLPQSRE-----LASARRGRRGPRLSSVRKG LSDVFNLAKVHAKFD
22 PKFLANPRIDKETWMDRVA---
23 MAKGSSTVYRVVVRPSGRVVMHGIRDTHREIRAREML
24 KGGW-----TFGGQRDLRQNTAAGLEGGEIA-----VP
25 VPTQRAGA-----
26 -----TLWDV--LSVSCKRARKRKKG N PRL-----
27 -----
28 >GENC01006608.1
29 -----
30 -----RVS RQGSDKAWFLP-----DNVRWTQIV
31 GNLR-----PPGESAGTAEMVESLPASQK
32 -----SKVGRKKEKFPGRATCAA
33 GQ----RIWN-----AFRLAFWAIAPS AQRKGLNPRS-----QR
34 QARSIQAFRTLEK-WIAISLLK-----CGTEWTCKAVK-----DYSEYCRARAL---
35 -----RAKQCVRPREFP-VTG-----FEIGLVRGAINDPLAGAEA
36 -----LAQLARLGRA---MP--QGT--DKQVIRALEKHRDT-LLGETRVD----EE---
37 -----TLDRYHAWAER---WSSD--RRRFAPERVAV-----KFSDS--AAESVS
38 RADG-GQKA---DLL-----ENARPHFEALRAALAESQGWDAPDDYIGLYSEED-----
39 -----
40 -----DVETIAQMALETVR
41 AAVR-----
42 --SGEPVPARASA--IPEYG-SKVRVVTACPGPFVVA-GDACR-----KIVWG-LL
43 EADPRVDLSGRRPLGDSISQFQ-DEVAR-----SLVGAV-----TT-----
44 -----EY-YSA-----DLTAATDLMPF--QLSAAIWTGLCDG-LG---ES-
45 -QDSD--LRSIGLHL-LG---PVQV--R-Y-PDLK----EG-----S
46 VEEGF--RV-----PTIISQQGCMMGLPLSWTVLNVFNLAT----A---
47 DLA--I-----GAPEDRV TDLIGR-----
48 -----APA--VARGDD--LAAA--FLPEEA--DRYEALISA-TGGRVNLSKSF
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 RS--SIG-FVLAERT----F-----MVKRSV--DRYQR
4 RAPRQAVLIQPHRPEGIP-----QNA-----
5 PILAQDPS-----
6 --APTMMSGVGRDVRGH----RTESLRMLPDCSLKGLLGGS----AST----DGA
7 PRYVTLP---VSAASSLY----EFEGR----FYGALC--RGLLSVHKDLLG-T
8 F-REF-GVPKF--YPREL--GGGG--F--PH-----P-RGYLR-----
9 ----AMESAG-GIGH----LRATLNITRFTESSRRAAALHM-----DP-
10 -----WRTERD----SRTLERARSQIAE-----ESRAREAGLLG-----
11 -----SF-VSKFQEEAE-----
12 -----VTRAAGQSLWE-----TTLSQKKRDEK
13 FPRLDVVRHRMLEITEKTKAIKFNRFLRRQLHR-----DEYLQIASHARGA
14 PTRY-----VRMERAPGLVALRDIVPGQQA-----PN
15 SFQTGTVP-----QHGNARGNRG
16 ERAEMSLVPTLRSRASLGQVARILPARASRKRGARRGPNPRPSDSSRGE-----
17 -----
18 >GEND01011317.1
19 -----
20 -----DLSRQGSDKAWFPL-----DNVRWTQIV
21 GNLR-----PPGESAGTAEMVESLPASQK
22 -----SKVGRKKEKFPGRATCAA
23 GQ----RIWN-----AFRLAFWAIAPSQRKGLNPRS-----QR
24 QARSIQAFRTLEK-WIAISLLK-----CGTEWTCKAVK-----DYSEYCRARAL---
25 -----RAKQCVRPREFP-VTG-----FEIGLVXGAINDPLAGAEA
26 ---LAQLARLGRA---MP--QGT--DKQVIRALEKHRDT-LLGETRVD---EE---
27 ----TLDRYHAWAER---WSSD--RRRFAPERVAV-----KFSDS--AAESVS
28 RADG-GQKA---DLL----ENARPHFEALRAALAESQGWDAPDDYIGLYSEED-----
29 -----
30 -----DVETIAQMALETVR
31 AAVR-----
32 --SGEPVPARASA--IPEYG-SKVRVVTACPGPFVVA-GDACR-----KVVWG-LL
33 EADPRVDLSGRRPLGDSISQFQ-DEVAR-----SLVGAV-----TT-----
34 -----EY-YSA-----DLTAATDLMPF--QLSAAIWTGLCDG-LG---ES-
35 -QDSD--LRSIGLHL-LG---PVQV--R-Y-PDLK-----EG-----S
36 VEEGF--RV-----PTIISQQGCMMGLPLSWTVLNVFNLAT-----A---
37 DLA---I-G-----APEDRVTDLIGR-----
38 -----APA--VARGDD--LAAA--FLPEEA--DRYEALISA-TGGRVNLSKSF
39 RS--SIG-FVLAERT----F-----MVKRSV--DRYQR
40 RAXRQAVLIQPHRPEGIP-----QNA-----
41 PILAQDPS-----
42 --APTMMSGVGRDVRGH----RTESLRMLPDCSLKGLLGGS----AST----DGA
43 PRYVTLP---VSAASSLY----EFEGR----FYGALC--RGLLSVHKDLLG-T
44 F-REF-GVPKF--YPREL--GGGG--F--PH-----P-RGYLR-----
45 ----AMESAG-GIGH----LRATLNITRFTESSRRAAALHM-----DP-
46 -----WRTERD----SRTLERARSQIAE-----
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----ESRAREAG-----
4
5
6
7
8
9 ----->KX883556.1-----
10
11 -----MQEVNIS-----TLTRQGVSN
12 DRPLSATEKVDADCMKILHESSGTDPNTAARG-----APHPSTEVEGEESRQRGRSK
13 -----AAPAPAREGSGADHRSK
14 NPAADNVSVWG-----PGGRALLSDSLLSSVLSLGSSSTRLITGLRGV-----TL
15 AIFTAEQDEVLARSLAYRTCRHLARRALHEGVEAVLSVMG-----RWFSSARESFL---
16 -----NGAKNPSLNPV-MRR-----FSPGTRKDRVTC-----
17 ---VAQFSYKGKRA---LP--VAS--DRVAQESLRVHERR-LTTPLSTP----DP----
18 ----VLADFRRFCRA---WAEG--HLDPNPLPLGVV-----GLSQA--ASYTHS
19 VRKG-GHMA---RCR-----EILEEMNASNYCEPPTRPP-----
20
21 -----G
22 LLPQ-----EWWDGWVK--ERLLT
23 HCRE-----RYSTPGRE-
24 -----PPKGTAVI--VYERG-LKARIVTKIETEACIL-GHQAR-----MRLVS-AM
25 RKMPEL-----SALSGNH-KEFLT-----AFNGQEG-----
26 -----WI-LSS-----DLTAASDLLPL--DLVSAGVEGICDS-GR-----
27 -LLPD--EELGLRVC-TG---PFDL--E-W-----
28 -VRGS-----PRRSSTGILMGAPPTWCLLSLIHLFW---L---
29 EQA-----RRVYGNNSGGQ-----
30 -----LPA--RIFGDD--LVAA--FDKRQK--LAYERSVTV-CHGLLSKGKHS
31 FH--DSH-GVFLERL-----FVASRHIVEVDHT
32 NRGPL-----
33 QVMGRLTT-----
34 -----RHNKIKSI-----SLLRTLPLRPLTVLQVSVAAGRRHVS-----GTL
35 PTPLAVG--GVSDGLMQH-----GFPPEL-----IRRCQI-----ALWPGLPA-F
36 Y-RKL-GIPAF--LPTLL--GGGG--LI-----LPEGWEA-----
37 -----PITRFS-KLVR----ASTTALVTG-----NGRLSYVPLARLPPPVLQMA----
38 -----LEEADQ-----LLPAHPHRVSHV-----TPKPWLGHIAW-----
39 -----HD-MGVWNDFRITA-----
40 -----AMVAQEGLYLLQ-----
41 -MPMECLGRPLNITVGRWK---DQVWKTRLALCK-----
42 -KAW-----WASVTRARYTIARAVRKCTEWPHV-----WL
43 PGSPDPDH-----PGVYGI--
44 -----PMYIDSKPVMEA KAAIALRQGRPTSR-----
45
46 ----->KX883562.1-----
47
48 -----MKILHESSGTDP-----NTAARGAPH
49 PSTE-----VEGEESRQRGRSKAAPAPAR
50 -----EGSGADHRSKNPAADNVSV
51
52
53
54
55
56
57
58
59
60

1
2
3 WGPGRALLSD-----SLLSSVLSLGSSSTRLITGLRGV-----TL
4 AIFTAEQDEVLARSLAYRTCRHLARRALHEGVEAVLSVMG-----RWFSSARESFL---
5 -----NGAKNPSLNPVMRR-----FSPGTRKDRVTC-----
6 ---VAQFSYKGRA---LP--VAS--DRVAQESLRVHERR-LTTPLSTP---DP---
7 ----VLADFRRFCRA---WAEG--HLDPNPLPLGVV-----GLSQA--ASYTHS
8 VRKG-GHMA-----RCREILEEMNASNYCEPPTRPPGLPQE-----
9
10 -----
11 -----WWDGVVK--ERLLT
12 HCRE-----RYSTPGRE-
13 -----PPKGTAVI--VYERG-LKARIKVTEACIL-GHQAR-----MRLVS-AM
14 RKMPEL-----SALSGNH-KEFLT-----AFNGQEG-----
15 -----WI-LSS-----DLTAASDLLPL--DLVSAGVEGICDS-GR-----
16 -LLPD--EELGLRVC-TG---PFDL--E-W-----
17 -VRGS-----PRRSSTGILMGAPPTWCLLSLIHLFW---L---
18 EQA-----RRVYGNNSGGQ-----
19 -----LPA--RIFGDD--LVAA--FDKRQK--LAYERSVTV-CHGLLSKGKHS
20 FH--DSH-GVFLERL-----FVASRHIVEVDHT
21 NRGPL-----
22 QVMGRLTT-----
23 -----RHNKIksi-----SLLRTLPLRPLTVLQVSVAAGRRHVSK-----GTL
24 PTPLAVG--GVSDGLMQH-----GFPPEL-----IRRCQI-----ALWPGLPA-F
25 Y-RKL-GIPAF--LPTLL--GGGG--LI-----LPEGWEA-----
26 -----PITRFS-KLVR----ASTTALVTG-----NGRLSYVPLARLPPPVLQMA---
27 -----LEEADQ-----LLPAHPHRVSHV-----TPKPWLGHIAW-----
28 -----HD-MGVWNDFRITA-----
29 -----AMVAQEGLYLLQ-----
30 -MPMECLGRPLNITVGRWK---DQVWKTRLALCK-----
31 -KAW-----WASVTRARYTIARAVRKCTEWPHV-----WL
32 PGSPDPDH-----PGVYGI--
33 -----PMYIDSKPVMEAKAAIALRQGR-----
34
35
36
37
38
39 >GENC01041260.1
40
41
42
43
44
45 -----LRSKVDQWFDQPSEFGDTL-----FR
46 DFPSFRDDTRFTN-----QGDENAAQY-----EYA-----
47
48 -----MSLVAADENS--LP---E--AKLPSRGVTEYVVD-VTADVHLE---RER---
49 -----ISKIVRDISLR--W-----LA
50 KLTG-----
51
52
53
54
55 -----PPRASVVA--LRERG-FKARIKVSPVELVEC-GHLLR-----SLVWP-ML
56
57
58
59
60

1
2
3 RNDPHV-----HGSLSGARLESFIN-----DLKERPLVNPAWVGRPSI-----
4 -----ISA-----DLTAATDGLYR--WAISAVWEGVCDG-AG-----
5 -FGED--VRKMGRVL-LG---PMKM--S-Y-----
6 -PASA--GV-----DEFTSKRGCLMGLPLSWFVLNVINIWA----A---
7 KSA---V-E----K-----SCRGRVGKPAD-----
8 -----FRF--AICGDD--LVGI--MPARAQ--RLYQRNIEV-VGSGLSEGKHL
9 ES--DRL-AFFTEQA---A-----T---FTRVTR--SA---
10 -----PRG-----
11 TLAEKAGL-----
12 -----REEYLEPV-----EMVGAIPVKSIHPSHFAGLRQGP----AAT
13 PSWAAAG---PALSSLA-----PWVHDS-----TLKVVS---RLGRYLPEW---R
14 L-LASSGIPPE--LPRQL--GGGG--F---PA-----RIPRRGFLA-----
15 -----APKQWRS---ALTCALLTGD---HKWLSRL-----R-NS-
16 -----WLTTGV-----YGDVMAEALSMAEE-----EFQDYPLAELE-----
17 -----IP-LSEVRTTRFTGGGLAI
18 TEA-----DLRTRVATVWA-----
19 ---PALTIASPFIEGRXY---CERFHFSRKRLS-----MVK
20 RKRL-----PGHLQPAQFNDTEAVSRLISLSEG-----RS
21 FLVPRESL-----PPGLNIDFIG
22 TRRPCG-----DGEDR--AIEVLT-----
23 -----
24 ----->LA534138.1
25 -----
26 -----
27 -----
28 -----
29 -----
30 -----
31 -----
32 -----
33 -----
34 -----RGVDRARE-----
35 -----
36 ---LLQFSSFARA---LP--PAD--DLVCKEALRDHREI-LNT--PVS---TPVA-----
37 -----VVSELRSYATR---WGQR--MARFANTSVAKL-----SASSS--ATLDYS
38 RRLG-GMRS---DLKGM---VDGFEEPASDVTEFHPYPTFSDPTRFTATGEMNRAMHP
39 APT-----R
40 GSVE-----YVVNVLEDPDLDERSRVARIIR
41 DVSL-----RRF
42 RAREGPLPCKATT--VKERG-YKCRVVTKSPADVVEV-GHLVR-----SVVWP-ML
43 EHDPRV-----QASLDGGR-LEEVF-----KGFSDRQIECPASLGS-----
44 -----LLLVSA-----DLTKATDGFSR--ESIYAVWDGVCEG-AD-----
45 -LPSD--VHALGRRI-LG---PMSV--E-Y-----
46 -EDDE--LE-----DLESSSRGCLMGLPLSWFVLNIINLWA----C---
47 ESS---I-R-----ECCHRLSLPSEVCKDL-----
48 -----LRF--AICGDD--LASV--FPAEAH--DGYERRISD-VGSGLSDGKHL
49 VS--LHV-LLFTEQM---C-----W---FEFEER--PAPPY
50 -----
51 -----
52 -----
53 TLLAWLKP-----
54 -----GKAPEGFRATYTCFTAVRMIDYVPVRSLVHPGHFAIKRVMG---PISFEL
55 -----
56
57
58
59
60

1
2
3 PSWATSG---PAITSAIP-----EWIGVE-----KRRRIH---QVVKVLRPEARA--
4 --LASVGIPPF--VPREI--GGGG---F-----PPARPGR-----
5 ----VLRDCP-RSYR----LFLFSLLVGDQKKAETVARII---NLWRTCGVAG-DL-
6 -----LSDALA-----EAEFEVE-----QKPLWKSPSDG-----
7 -----SL-AERLSTVIDDHLTME
8 D-----DALLRLASVWA-----PALGIAGFQQGRP
9 YCKKFGAFAKQYRALVGAG---SARITGRGNPLRN-----ED
10 EAAV-----ALRLKEICAGPVVFVPRDRIPPGL-----GVV-----VQ
11 GMCSRKRA-----
12 -----VGLHR--KIRSNRMESA-----
13 -----
14 -----
15 -----
16 >LA710658.1
17 -----
18 -----PPQKKTLHDNP-----RVDVWWNAC
19 EAVL-----GSRGAIKPGRGATRQVLPR
20 -----STRQRAAFLSVSLPSAAHTKL
21 LS----LVWE-----ALLGVAAYRPLPVWGS-----RF
22 GNKRCLELQLRAL-WLVRTVDG-----QGIGATIAWLK-----SAASGARNTCV---
23 -----TGAALLPSTRFL-VRK-----ILIGDTRLES-----
24 ---LDQLSFLSRS---LP--EGD--SVVESRTLIAHRDN-MTTSFRTS---PN---
25 -----LISNLRGFAMA---FATR--HLKKEDLREEV-----VPTPS--ASLDSS
26 RKEG-GCRE---EVR-----RMFMRWLIECPTHVVPHDAWNQAIFS-----
27 -----S
28 EWLD-----PSEVYGLRRAQATS
29 HVAR-----AAAYK
30 TAAAETLEHRVLT--VPERG-WKRRVVSAPPATVA-GTVLN-----RAMLR-GV
31 RRHGPC-----AMFLRGDR-KGAVE-----QAVSHASVGC-----
32 -----RF-VST-----DLTAATDRLPL--DLVRAVVVDGLCDGWSG-----
33 -LPPV--WAEALYAL-TG---TQTL--R-Y-----
34 -PWGQ-----TIQSESGVLMGLGPSWPIMSIHAWW---V---
35 ELA-----ASRVGIPRSVAR-----
36 -----RTT--AIGGDD--LLGA--WPPRLE--ESYRRLVLE-TNGKPSKGKDF
37 SS--DTS-GNFTEMT---F-----WVVGEA-----
38 -----
39 DGTPQIRW-----
40 -----SAAIPTKGLVGTS-----IDEL-----GAA
41 YESLGSE-----PGRCLR-----GRRVLK-----ALRPHAWR--
42 V-CREAGVSVN--APRLL--GGAG---L--P-----PLRGSLA-----
43 -----RVDFK-KWHA----LALGKFLYG-----SGQDQIP--FSPPSWVEAADPAV-
44 -----WEARQV-----AEQRLRAE-----AEIGIVSF-----
45 -----D-TRPLCGTPGSK-HV--
46 -----VQSISDQMAWW-----AGARVFS
47 DTPFPPVATEMVLKKYHR---LIKRCSSRTKEG-----IPSSLAVKSGR
48 NSRF-----ALTAKARRNRDRWVIRNALFDNGI-----PV
49 DPRHRPII-----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3
4
5 >U90136.1
6 -----
7 MHHKVNKTQREVHFPMD-----LLQACGASA
8 PRPV-----ARVS RATDLD RRYRCVLS---LPEERAR SVGCKWSSTRAAL
9 -----RRGLEELGSREFRRRLRL
10 AD-----DCWR-----AICAAVCTGRKFPSFSVTD-----RP
11 ARAR LAKVYRMGR-RLLVGVV-----CRGESVVSSDLK-----QECADLRRVIF---
12 -----EGSTRIPSSSLW-----GLVGVLGWTSPER-----
13 ---AMQLTFIGRA---LP--YGS--PDVERRALASHAAT-LSIPAECH---PN---
14 ---YLVAAEQFAKS---WADD--NLPRKFRIYPI-----AVQES--SCMEYS
15 RAQG-GLLQ---SFR-----KG FVGYDPAAPSADPDDL-----
16 -----
17 -----ELAKERG--FSRIR
18 ASWY-----STFRYRGEL
19 KSTNQSLEARVAV--VPERG-FKARI VTHSASRVTF-GHQFR-----RYLLQ-GI
20 RRHPAL-----VDVIGGDH-RRAVE-----TMDGDFGLLRPDG-----
21 -----RL-LSA-----DLTSASDRIPH--DLVKAILRGIFSDPDR-----
22 -RPPGTS LADVFDLV-LG---PYHL-H-Y-----
23 -PDGS-----EVTVRQGILMGLPTTWPLLCLIHLFW---V---
24 ELS---D-W-----A-----PARPNHNSRGFVLG-----
25 -----ESF--RICGDD--LIAW--WRPERI--ALYNQIAVD-CGAQFSAGKHL
26 ES--KTW-GIFTEKV----F-----T--VKPVKM--KVRVR
27 SEPSLKG YVFSRSSAFSCRMGGKGITGI-----RAARLYTIGAMP
28 RWSRRIRD-----
29 --VYPGSLEHRTASQRYGEPVTVYRFGRWSSAIPLRWAVRAP---TRTVGN---PVQSL
30 PDWFTVG--PAASSVAAD-----SNAFGA-----VSRVLR-----RMFPGLPR-K
31 L-ASA-GIPPY--LPRVF--GGGGLVKS-----TGLTT-----
32 -----KIGAVASRRWM---SRIGHDLYRS---RERKSTL-----G-RV-
33 -----WTLSTS-----PAYAASLHEVEKFM-----DRPDIILTRKC-----
34 -----RNP-MLKHARELGLFEEVF-
35 -----ESRVGGGILWA-----SLNGKALVES
36 HS PSLQVSRNLRRSLACP--SGGFLRPSAPIGK-----LVQRHTLPR
37 GTVW-----FLESSATDSARQGGMGLPPPPPPLGG-----GGMAGPPPPPFM
38 GLRP ESSV-----
39 -----PTSVP--FTPSMF SERLA AALESLFGRPPPS-----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 >MF176258.1
47 -----
48 -----
49 -----MPLVAR-----PTSMRVATRSEVSRRVNLA A
50 -----
51 -----ETWK-----AICAAVATRKWPKFNLDF-----QT
52 DRRRFCTFQALAR-HLLSALA-----RGPVETIKRVK-----SWTAQIREDVF---
53 -----LHGRCP-PSA-----FRGLVGALGWTDA--NS-
54 -----HLQLSYIGRA---LP--YGD--ATICEHALKSHADV-LGSLGVTP-----PS-----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----LLGAARGWAKE---WAIK--NRLSLDLIREP-----PIQES--SCLEY
4 RREG-GYAR---AIK----EHTARCLTLPPSSVRHSMK-----
5 -----E
6 GTRA-----SNPGRVL--KDLAF
7 ADLS-----
8 -----NTARVVA--VPERG-WKARIVTAAHAAARVAF-LHSLR-----HGLFA-AL
9 RNDRRT-----KLVAEGLH-REAVE-----ALFATGVDPYC-----
10 -----TI-VSA-----DLSAASDTLHG--DLLQAIVEGLGDA-LGC-----
11 -TSKF--MDEFTNAA-VG---SYTL--S-Y-----
12 -PDGS-----TVTTQRGALMGLPTTWPLLCLTHLFW---C---
13 DQS--L-----VQLHGRGGCREKETPGSSR-----
14 -----EAE--VICGDD--LAAA--WRPDRV--AAYEEIAVL-CGAVFSPGKHL
15 KS--KVY-GIFTEDI----Y-----S---VQVSCV--PTRVT
16 REVSKTWTGPVDKVMNLVNELGQLLRWRAQRYKVGKTGAFQFP---PRN-----
17 VTLSYVSK-----
18 -----ETGLKMNPV-----SVLFREWSTAIPRLWAVRAP--KHAPGMR-----GDL
19 PPWVTVP--LAAHAVASA-----PGRWSK-----VCRVVK-----LCYPGMAK--
20 F-FASHGIPPY--LPRVLP-----GGGG--L-----PTPLG-----
21 -----DRVKIG-RVAS----RRIRKALGG-----ALYRSTDPTTFGSIWTTALS-PA-
22 -----YALALA-----DAQDLDAMHPDA-----FKVRQTRSLRP-----
23 -----P-FKDNGNYADFV-----
24 -----EMRAGKGSLWA-----LRQGAQLVGRSFTPSP
25 ARVSNALRTKVRKALAKRG---YLRSSAPVSKLIE-----RASVKPS
26 DHMW-----IRERPVLSPGRFGGR-----MS
27 GKKEEPVS-----
28 -----FQWKK--RESVKTGS-----
29 -----
30 >GDUK01014407.1
31 -----
32 -----SCMRWLRLMLKTGTSP-----AADAAGSVG
33 HS-----GPTESWTRSGSAVARAPLGS
34 -----DVLASGMSLEKRKALASRIK
35 RANLRLRSLLG-----TWDAGLTAFFGRQIRR-----EV
36 ADSVVVWYDHVRS-WLVRVAAT-----QGVEVAAAELK-----KFSSDARRAWVEGS
37 -----PPRHFFWRSCPEW-LRR-----KQGL-----
38 -----WAQLSYLGRA--LP--LGV--DFHEKKALDQHYAD-LTSSFSCS-----EE-----
39 -----VLASARAFAER---WARS--HLSRAPGPEDWLEF-----PSGSS--ATFGRK
40 RREG-GFAT-----DCQILLESAAEVDTVPPDHVPV-----
41 -----G
42 PWSG-----DVTAFRLVTAALVE
43 AHGR-----
44 --RDTLPKGGRVAV--VPERG-HKVRVVSAMERHALIL-GELAR-----KRLFK-GL
45 RKWGPL-----AAVLRGDS-RSGVL-----RLLGVPG-----
46 -----EF-VSA-----DLRAASDLIPL--DLASAIVDGLEAS-GRL-----
47 -MPAE---IWGLRLGTG---PQEL--T-W-----
48 -----PSGK-----TAVTSRGILMGLPTTWALLNLIHGWA---W---

1
2
3 SQA-----AKAVPVQSLPRASAARGGR-----
4 -----PALSCAICGDD--LVGV--APPVVI--DAYEAAMRS-IGAEMSAGKHF
5 RS--TRR-AVFLEEL----W-----E--FRGSAV--SVVDG
6 NPIYRTIR-----EKG-----
7 KRVRRQVV-----
8 -----NSTRLLRAD-----RLVRATAIPLKGMVSAPGVPG-----IDP
9 PDWWVAG--LAESALADR-----GFSRAK-----ISSVSR-----SLRPDLPV-R
10 F-RAS-GIPPF--LPRAL--GGAG--L-IAPH-----
11 -----TRCEAP-RLHR----KAIATLI-----WGYDPDSTLAFDRVWSDSVP-GD-
12 -----WRELAI-----DMVDAE-----FTPEVTRVLPV-----
13 -----GSRP-TGSWVAIGHPDQ---
14 -----VREKAVADNAS-----ALSSSLGPSPDVVG
15 YPTIPVLARRLRKVREDLA-----SRWASAKPTKR-----PVDELLLWRQRAAD
16 LRVW-----VPLTDQTSRWNDWL YVSATSPYKKS-----WR
17 RLVISSVS-----REDLPAVT
18 -----EWLPGGRIPSSTMDSPPWEWR-----
19 -----
20 ----->GFKT011160020.1
21 -----
22 -----
23 -----PMKEPVDCRLS-----TPAGFSGTV
24 PPP-----GRTKAARQGTIPVRRSRGSA
25 -----SALPGKIYGWSRRQRDR
26 FA----MLLS-----SFDAALAAYSGVVVSRGT-----RS
27 LPPSLRLFRTMTRKWLSVTARG-----NGVEFAIASAK-----EFSAACRAGWISGT
28 VP-----DHFFMKWLPEPVR-----RKSGL-----
29 ---WAQLSFIGRS---LP--EGG--DRHEIEALANHCAA-LSSSFEVP----AD----
30 -----VLTSLRNYSED--WARR--HLAADPDPSLLCEP-----CTGNS--ATFERT
31 RREG-GFAQSITDLV-----SSPTDNLLPLESMPFGPTQQQ-----
32 -----
33 -----ALPVHVLEVSLRY
34 RNGS-----
35 -----DPKGRVSV--VRERG-HKVRVVSAMETHELVL-GHAAR-----RRLFK-GL
36 RRERRL-----KDTLKGD-FATT-----AFVGCAG-----
37 -----TV-ISS-----DMKSASDLIPL--SVASAIVDGLEAS-GR-----
38 -LLPV--EVAGLRAC-TG---PQHL--V-Y-----
39 -PDGS-----EITTRRGILMGLPTTWAILNLMLWC----W---
40 DSA-----DRQYRLEGHPFRAT-----
41 -----VRSDCRVCVGDD--LIGV--GPDSLL--RSYDRNLGL-VGMILSPGKHF
42 RS--NRR-GVFLERL----L-----E--FQTRKT--VYEHA
43 VIYRKVG-----HRR-----
44 VPVDRSHI-----
45 -----PVVTRV-----TVLNTIPLKGLVRAS-----VLGR-----DDP
46 PVWWAAA--VAESSLLS-----DYPRK-----IFAAAR-----TLRPGLSR-Q
47 F-RRL-GIPPF--LPREL--GGAG--L-----VGSSD-----
48 -----RVDAP-AFHR----KAISSLV-----WGSDATAAYSFIRMWQGFEG-HP-
49 -----WKTAAS---QETDTWFADYKVTRPGK-----MYPDRYGFQDG-----
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----ES-LRTKSTMLNSA-----
4 -----VYETFLGPDPE-----ATHYPSL
5 RIVASRLAKVRKDLVSRWP--SVRPVG--KDLGT-----ILEAFEESKL
6 CTLW-----VTPYDASGYFDDSLLMDESVYQRR-----FR
7 QLVIAGLM-----REGRMGDLL
8 F-----PSWLP--PSTVVS-----
9
10 -----
11 >M64034.1
12 -----
13 -----MKEPVDCRLS-----TPAGFSGTV
14 PPP-----GRTKAARPGTIPVRRSRGSA
15 -----SALPGKIYGWSRRQRDR
16 FA----MLLS-----SFDAALAAYSGVVVSRGT-----RS
17 LPPSLRLFRAMTRKWLSVTARG-----NGVEFAIASAK-----EFSAACRAGWISGT
18 VP-----DHFFMKWLPEPVR-----RKSGL-----
19 ---WAQLSFIGRS--LP--EGG--DRHEIEALANHCAA-LSSSFEVP----AD----
20 ----VLTSLRNYSED--WARR--HLAADPDPSSLCEP-----CTGNS--ATFERT
21 RREG-GFAQSITDLV-----SSSPTDNLPPLESMPFGPTQQQ-----
22
23 -----
24 -----ALPVHVLEVSLSRY
25 HNGS-----
26 -----DPKGRVSV--VRERG-HKVRVVSAMETHELVL-GHAAR-----RRLFK-GL
27 RRERRL-----RDTLKGDFAATTK-----AFVGCAG-----
28 -----TV-ISS-----DMKSASDLIPL--SVASAIVDGLEAS-GR-----
29 -LLPV--EIAGLRAC-TG--PQHL--V-Y-----
30 -PDGS-----EITRRGILMGLPTTWAILNLWHLWC----W---
31 DSA-----DRQYRLEGHPFRAT-----
32 -----VRSDCRVCVGDD--LIGV--GPDSLL--RSYDRNLGL-VGMILSPGKHF
33 RS--NRR-GVFLERL--L-----E--FQTRKT--VYEHA
34 VIYRKVG-----HRR-----
35 VPVDRSHI-----
36 -----PVVTRV-----TVLNTIPLKGLVRAS----VLGR-----DDP
37 PVWWAAA--VAESSLLS-----DYPRK-----IFAAAR-----TLRPGLSR-Q
38 F-RRL-GIPPF--LPREL--GGAG--L-----VGPSD-----
39 -----RVDAP-AFH-----KAISLV-----WGSDATAAYSFIRMWQGFEG-HP-
40 -----WKTAAS--QETDTWFADYKVTRPGK-----MYPDRYGFQDG-----
41 -----ES-LRTKSTMLNSA-----
42 -----VYETFLGPD-----ATHYPSL
43 RIVASRLAKVRKDLVNRWP--SVKPVG--KDLGT-----ILEAFEESKL
44 CTLW-----VTPYDASGYFDDSLLMDESVYQRR-----FR
45 QLVIAGLM-----REGRMGDLL
46 F-----PNWLP--PSTVVV-----
47
48 -----
49 >AF039063.1
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----MKEPVDCRLS-----TPAGFSGTV
56
57
58
59
60

1
2
3 PPP-----GRTKAARPGTIPVRRSRGSA
4 -----SALPGKIYGWSRRQRDR
5 FA----MLLS-----SFDAALAAYSGVVVSRGTR-----RS
6 LPPSLRLFRAMTRKWLSVTARG-----NGVEFAIASAK-----EFSAACRAGWISGT
7 VP-----DHFFMKWLPEPVR-----RKSGL-----
8 ---WAQLSFIGRS--LP--EGG--DRHEIEALANHCAA-LSSSFEVP----AD----
9 ----VLTSLRNYSED--WARR--HLAADPDPSSLCEP-----CTGNS--ATFERT
10 RREG-GFAQSITDLV-----SSSPTDNLPPLESMPFGPTQGQ-----
11 -----
12 -----ALPVHVLEVSLSRY
13 HNGS-----
14 -----DPKGRVSV--VRERG-HKRVVVSAMETHELVL-GHAAR-----RRLFK-GL
15 RRERRL-----RDTLGDF-EATTK-----AFVGCAG-----
16 -----TV-ISS-----DMKSASDLIPL--SVASAIVDGLEAS-GR-----
17 -LLPV--EIAGLRAC-TG--PQHL--V-Y-----
18 -PDGS-----EITRRGILMGLPTTWAILNLNMHLWC---W---
19 DSA-----DRQYRLEGHPFRAT-----
20 -----VRSDCRVCVGDD--LIGV--GPDSLL--RSYDRNLGL-VGMILSPGKHF
21 RS--NRR-GVFLERL----L-----E--FQTRKT--VYEHA
22 VIYRKVG-----HRR-----
23 VPVDRSHI-----
24 -----PVVTRV-----TVLNTIPLKGLVRAS---VLGR----DDP
25 PVWWAAA--VAESSLLS-----DYPKK-----IFAAAR-----TLRPGLSR-Q
26 F-RRL-GIPPF--LPREL--GGAG--L-----VGPSD-----
27 -----RVDAP-AFH----KAISLV-----WGSDATAAYSFIRMWQGFEG-HP-
28 -----WKTAAS--QETDTWFADYKVTRPGK-----MYPDRYGFLDG-----
29 -----ES-LRTKSTMLNSA-----
30 -----VYETFLGPDPD-----ATHYPSL
31 RIVASRLAKVRKDLVNRWP--SVKPGKDLGTL-----EAFEESKL
32 CTLW-----VTPYDASGYFDDSLLMDESVYQRR-----FR
33 QLVIAGLM-----REGRMGDLL
34 F-----PNWLP--PSTVVSGFP-----
35 -----
36 >MF190030.1
37 -----
38 -----MGNKVGFRSSRSCDR-----YLKRAPLPG
39 DRKG-----RVRGRVIDKLEV-----RTRSRKSKGSRMRGARPPMA
40 -----ESYKIKPGWRKGFMTPGRE
41 IE----GLWV-----ALIATLTSISPRICRFNMRC-----RA
42 DRDRHGAMVRLAV-WFTRRVD-----RSSYDLLKSAK-----KLTAGFREASI---
43 -----LRRPYNWVDAPPG-WAK-----RFRPILKGIRLSEA-----
44 -V-LDQLSYMGRA--LP--KGG--PRVLQEALKDHKDC-MLKPSPLP----EE----
45 -----VFGAARDFACR--WADR--YVASGGRLSHP-----GTTES--SVLGST
46 RRDG-GISS-----SLCGLTYSVKPLPMVDVQDEVRALVSSLKCCTPSEC-----
47 -----
48 -----DVDAALR--TTSIR
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 EWAI-----DESSPDV-
4 -----IPENKVLA--LPERG-YKTRVATIPDLPRQIL-GGFVR-----SYLIP-AL
5 RRTPEC-----ADTLRGDD-ERAAK-----RIGSSQRSFST-----
6 -----DMIVSS-----DLTAASDTLSR--ELLALTAGLVDS-GK-----
7 -FPSW--FGPALRQL-VG---PMLL--R-Y-----
8 -PDGS-----SITSQRGTLMGLGHTWSMLSLSHLFW----L---
9 DYS-----NKTVHKTIARSVE-----
10 -----RTA--AVCGDD--LIVR--APWRWI--NSYHRIAKM-CGAQFSSGKHF
11 IS--RER-YCFLEQL---W-----T---LKSKEIVPAYTRA
12 LDK-----RIP-----
13 KGGDQFPP-----
14 ---LGTDPRKSVRRV--MKFHIPSCKVRSTAIPLAGLTSGAGSDPEKPLV-----VGT
15 PTWAKLG--LAVENTSRT-----PVAARK-----VRAVSR-----ALYPRLPR-W
16 M-RKH-GIEPY--LPRAY--GGGG---L-----LPPGNVS-----
17 -----HLGALA-PQVR----KLVTVTAYG----RCPKTAARLSRIYLPMNRTA-TN-
18 -----FTEAQE-----HTDIRFKRVGAI-----LCRRGKVPF-----
19 -----NGHAVK-LGTLTDAPER-A-LAI
20 -----ARELSMVLFPP-----PRNRK
21 KSTISPFQVGRAFSAMRK---AIKKWPSAHP CSA-----KVTRRVLLQKARAL
22 EEEF-----VWWSPRRRIPEEGFGGWG-----LA
23 PAAQRVLK-----
24 -----DSL RWSSDC-----
25 -----
26 >KX883516.1
27 -----
28 -----
29 -----
30 -----
31 -----
32 -----MKPPRPTEVHKDFRLP-----VLLRACR VG
33 GTWS-----RDVSRYLVPALQDG VGV---RPGE GGNPQTRPRKPAK SVD
34 -----LERTRRYDLTVWRKEFRSRGRL
35 VE---GAWA-----AVHAALVAAHPDLGRADMST-----RA
36 GRDRIWA WRRRLAV-WMVRWID-----RDPRDMV RGWK-----QLAATFREASI---
37 -----MQKPYDWAAAPAG-FPK-----RLRPVLAGVRLTPP-----
38 -V-LDQLSYIGRA--LP--QGD--AWVQRKALDDHREC-LSSA-PRP---LANH----
39 -----VLA AARDFARS--WASR--HVRPSGRISHP-----KVLES--ATIQEP
40 RRSG-GVAA---HLC-----KQGFDLPPFEGLTDEL RDQIRAMGESAPS DCDLEA---
41 -----I
42 GRST-----AFRDWAL--GQLCC
43 PDAS-----
44 -----PPRSRVEA--IPERG-FKARVVTVPDLPESIL-GSALR-----CYLLP-AL
45 RRTPEC-----ADVLSGKE-ELAAM-----RIASSRRPIE-----GD-----
46 -----WI-VSS-----DLTAASDTLSF--ELLQALVDGLEES-GT-----
47 -LPPW--AIRALKVL-VG---PQEL--V-Y-----
48 -PDGS-----VLVTRRGTLMGLGHTWSILSLSHLFW----L---
49 DWA-----HRSVDPRVADVVR-----
50 -----RTA--AVCGDD--LVVR--APLKWV--QSYHQIAEA-SGARFSTGKHF
51 VA--KRR-YVFLERI----W-----S---LKSSLQ--RVRLP
52 RERTR-----RAG-----
53
54
55
56
57
58
59
60

1
2
3 RWVREFPP-----
4 -----LGARPVKVRPT-----VVR SIPSRSVPSTAIPLAGLTAGEVDCHFGPVRNLVL
5 PNWVRLG--MAVETLLHS-----PEAARK-----VRKVVR-----ALYPGLPR-W
6 M-RRH-GVQPY--LPRFL--GGGG--LL-----PPGRNVT-----
7 -----RIGRLP-RPVA-----KAVTTWVYVG---RSPRENLSL--AKIYLPTTRRI-TN-
8 -----YAVGEQ----HTAVRFQQVSAVLCRA-----GKKPFDGHGRR-----
9 -----LGTLSEAPEAA-ALAI
10 -----ARELSMVLLPE-----PSEERAYR
11 RLLRPSQVGQTFRRLVL-R---ANRRWPGSKPICP-----RAPRRALLARCKAV
12 EEGF-----VWWANPRLLTQFGFGGWG-----LE
13 PAAQRVVK-----
14 -----DALGWSTFRPTPGRGTH-----
15 -----
16 >GEND01007370.1
17 -----
18 -----LDHEPPTMGSSKKTARETHVSNSVSLH-----ALLRNPRVG
19 GSWS-----KDVSL---YLAPA-----RRSGVQGESGERGLPSRTGL
20 -----ATHANLEHRGAS-----RRGQGHAKRGPEPRRKAPKGGYTR
21 QR---VCWS-----REFKSSGRKLEAGWNALTAALAACHPAALRFDSLVA
22 DRDRVGVALKRALV-ACVRRLD-----RGVGETMAELK----SLATTFRASI---
23 -----RQVPYSWDTCPAG-LRR-----HLEAALRGFRLTTE-----
24 --TLDQLSFLGRA--LP--PGD--STVEERALRVHRDV-ITSD-PAP---IPPG-----
25 -----ILKAAEAFACt---WALT--YLHRGGRGLAHP-----KTMES--STITSR
26 RREG-GVAA-----DLCMKSMDDPLPETVQAEVRRQ---VLELEGPLPSECD
27 LSA-----A
28 FRTT-----ALRDWAV--TTLTE
29 PGYV-----
30 -----PTH-KVLP--IPEAG-WKCRIATVPELPASVA-GGLCR-----SYLLP-AL
31 RRTPEC-----ADVLRGEE-DKAAR-----RVAASRRPIE-----GD-----
32 -----WI-VSS-----DLTAASDTLAF--GLLQALVVRGLALS-GR-----
33 -LPPW--VVRSLLEV-VG---PQRL--T-Y-----
34 -PDGE-----TVITARGTLMLGHTWSLLSLSHLFW----L---
35 EAA-----HASVPRPVSEQVR-----
36 -----LTA--AVCGDD--LVVR--APLGWI--RRYEEVASL-TGAQFSQTKHF
37 VA--KRR-FVFLEEL---Y-----T---LRGTQQ-AASTVL
38 GPTS-----RGG-----
39 RLVRAWPA-----
40 PGTRPPKVRVTRLRTY-----ASGTVKSTAIPLAGLVAGDVMTHQVMA-----DSV
41 PLWSKIG--LAVESLAST-----RERALR-----VRRVTR-----HLYPALPS-W
42 L-LRH-GIRPT--LPRAL--GGGG--LL-----PNSGNCM-----
43 -----RLGRLP-----RRVAKMVTAVTVGRCPTAARL--SRIYLPLSRRA-TN-
44 -----FAEAKV-----HTDIRFKRVGAV-----LSRSGREPW-----
45 -----NGKGVR-LGTLTEAPEQA-ALAI
46 -----ARELSMVLLPE-----PKKGGS
47 NAALSPFQVGKQFLLALT---ALRRWPGAKPVA-----RAPRKSVSDRTRAV
48 EEAf-----VWWAPSHRIPEEGFGGWG-----LA
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 PAAQRVLK-----
4 -----DSLGWSA--PSP-----
5 -----
6 >GENB01008321.1
7 -----
8 -----LDHEPPTMGSSKKTARETHVSNVSLH-----ALLRNPRVG
9 GSWS-----KDVSL---YLAPA-----RRSGVQGESGERGLPSRTGL
10 -----ATHANLEHRGAS-----RRGQGHAKRGPEPRRKAPKGGYTR
11 QR----VCWS-----REFKSSGRKLEAGWNALTAALAACHPAALRFDSLVA
12 DRDRVGVALKRLAV-ACVRRLD-----RGVGEELMAELK-----SLATTFRASI---
13 -----RQVPYSWDTCPAG-LRR-----HLEAALRGFRLTTE-----
14 --TLDQLSFLGRA--LP--PGD--STVEERALRVHRDV-ITSD-PAP---IPPG-----
15 ----ILKAAEAEFACT---WALT--YLHRGGGRLAHP-----KTMES--STITSR
16 RREG-GVAA-----DLCMKSMDDPLPETVQAEVRRQ---VLELEGPLPSECD
17 LSA-----A
18 FRTT-----ALRDWAV--TTLTE
19 PGYV-----
20 -----PTH-KVLP--IPEAG-WKCRIATVPELPASVA-GGLCR-----SYLLP-AL
21 RRTPEC-----ADVLRGEE-DKAAR-----RVAASRRPIEGD-----
22 -----WI-VSS-----DLTAASDTLAF--GLLQALVRGLALS-GR-----
23 -LPPW--VIRSLEVL-VG---PQRL--X-Y-----
24 -PDGE-----TVITARGTLMLGHTWSLLSLSHLFW---L---
25 EAA-----HASVPRPVSEQVR-----
26 -----LTA--AVCGDD--LVVR--APLGWI--RRYEEVASL-TGAQFSQTKHF
27 VA--KRR-FVFLEEL---Y-----T--LRGTQQ-AASTVL
28 GPTS-----RGG-----
29 RLVRAWPA-----
30 PGTRPPKVRVTRLRTY-----ASGTVKSTAIPLAGLVAGDVMTHQVMA-----DSV
31 PLWSKIG--LAVESLAST-----RERALR-----VRRVTR-----HLYPALPS-W
32 L-LRH-GIRPT--LPRAL--GGGG--LL-----PNSGNCM-----
33 -----RLGRLP-----RRVAKMVTAVTVGRCPTAARL--SRIYLPLSRRA-TN-
34 -----FAEAKV-----HTDIRFKRVGAV-----LSRSGREPW-----
35 -----NGKGVR-LGTLTEAPEQA-ALAI
36 -----ARELSMVLLPE-----PKKGGSSH
37 NAALSPFQVGKQFLALTQ--ALRRWPGAKPVSA-----RAPRKSVSDRTRAV
38 EEA-----VWWAPSHRIPEEGFGGWG-----LA
39 PAAQRVLK-----
40 -----DSLGWSA--PSP-----
41 -----
42 >GBWZ01004682.1
43 -----
44 -----GSRTP-----WRKSLPQMN
45 RGSR-----
46 -----QK
47 VR----ATWS-----VIRAALAAASGFVRCDLST-----LA
48 ARDSANGLCRLAW-WLWRTTAH-----SGIEYTLKLLK-----TWLSETRRWWV---
49
50
51
52
53
54
55
56
57
58
59
60

1
 2
 3 -----EEVPKPKRGGG-MAR-----FLHGNFRSMVDSNA-----
 4 -K-ALQLSYAGRA---LP--VGG--PSVILRSRGLHQDA-FASNPVTP-----PE-----
 5 ----LLTSARIFARE---WAKR--HLSSSPSFARA-----GLNES--SSLEST
 6 RAKG-GVLA---QTR-----ERIYAIPDEALRERKSSLQSREEFADLSSC-----
 7
 8 -----
 9 -----DLDIVAR--ESLLR
 10 DSAL-----EELNRSLENG-
 11 -----NPQAEVRY--CLERG-LKCRVVTKSPWALVYL-AHSVR-----HELLN-AL
 12 RNDKRT-----KHVLKGDH-QEAVE-----DVFSTPRRID-----NR-----
 13 -----QI-VSA-----DLTSASDLLPL--DLVGALVDGLRDT-AR---LW-
 14 -WPEW--MLKVLGIV-TG---PMDL-----WYPAE-----
 15 -LGED--EL-----TIVGTRCGILMGVPTTWSILSLVQLFW----A---
 16 EEA----W-----RSRFPNPAWGHPSTM-----
 17 -----PET--AICGDD--LVGW--WPKSVI--KTYEAVAKR-CGAQFSAGKHF
 18 HS--RTH-AVFTEEI----Y-----R---LRSISE--ETPRC
 19 PDWESLPKLQQRWVRSGTKRR-----PRK-----
 20 RLVNFLRT-----
 21 -----ARLGGP-----VGKGVVKYPSIPLAGLARPSVLPGSYQTV-----
 22 PWWIAIG--PTYEAVAAQ-----TGNPSA-----ARRVLR-----TLHPNAWR-W
 23 M-RER-GIFPE--LPREL--GGGG---L---P-----PLRGSAD-----
 24 -----VLKLP-PFLR----FGVAKLLYG-----TKFRDLR--SPGALWSTTFA-PE-
 25 -----WRVMAA-----DHAEVK-----FKTSAVNCHPG-----
 26 -----RVP-FHGKGELGGVKALED
 27 VTTRL-----AGELTLMMEHP-----KDGR
 28 QAFSVVPRFIARSVRKFYS---ARGKVARRSKMKGIDPSGVPR-----AL
 29 LKRW-----ASLGQTRWFSEPAAGEGFVLGI-----PG
 30 QVKRRIAP-----A-----
 31 -----LGWAR--P-----
 32
 33 >GBPT01024369.1
 34 -----LRM
 35 AVRKPSTQAARERVTTLSLVEAALIGSRLVPLAGIVRPARARAL-----RSDRGATLN
 36 PKRRGVGQPLEPIWAQGSTAVPIPNPSSSPHKACLKPISRQQDGKPPKASAMKKGHKIKD
 37 -----HFRRFRLSPRSSEKVRASANPPKEDHDVDWSSVRRLYPEWGRISKRGHRQ
 38 LK----TVWA-----AFRAALAAVGSTRVSLRT-----NS
 39 GREALRKWCKLAH-WLVRTTGL-----SGITHVLATLK-----AWTATVRRWWV---
 40 -----EGVREQDRQGG-LAK-----FLAGSLRFLVTRDD-----
 41 -A-AVQLSYVGRA---LP--QGT--LSVQRKALVSHKKA-LGEDYRTP-----RH-----
 42 -----ILREAREFAAG--WGDR--YLRSPVLMNC-----GFSQS--ACLEVP
 43 TSKG-GTLEATLERT-----NALQTDVNWSKAVADWRDTLKSS-----
 44 -----I
 45 PTLP-----DCDREILARGTALR
 46 DRSI-----VELREA
 47 VSQSGHPRAAVAV--LPERG-SKARVVTKSPWALVQL-AHPIR-----GWLLS-GL
 48 RRDPRI-----AKVLEGDH-HGAAT-----GLVDVPRVLADGR-----
 49 -----EF-LSA-----DLTSASDLIPL--DLAAALIEGLMEG-TGDT---
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
 2
 3 -LPDW--AREVLRRA-TG---PMLV--D-YSTA WP-----
 4 -PSGG-----RRPLPGPCFISRRGALMGLPTTWSLLCILQLFW----A---
 5 DRA----W-----RLTFPNIAQRTDPGNPGMT-----
 6 -----PAT--IICGDD--LAAW--WPKQVS--DAYEQTVVS-CGGIFSVGKHF
 7 RS--RAV-VVFTELL----FHAEKSTVRALAGAGPDKISAPAI--ASWERE--GRPFL
 8 MKVRRPHAKSTARRRSER-----GRR-----
 9 PMIPKMTR-----
 10 ---GIGPHDLSDHRRV-----VTRLVPFKSIPLKGLLEAEGGGPLADNH----VRQ
 11 PRWVNIG--PVADRISKM----TGNPYA----VRRVLR----TLFPHYWG-W
 12 F-QKH-GFVPT--LPRFL--GGPG--L--P-----PRTGDLD-----
 13 -----RLALP-RDLL--LKVRTLVHG----TKWHLLR-----SPK-AP-
 14 -----WTQMGA-----GFYRQMVAGH-----AEARLQTHARC-----
 15 -----ARVGRVP-YNGKGVLGGV-DVYE
 16 DL-----ATRFENEATVG-----IFG
 17 CRDDSRRSAFKIRPGQVSQ--ALRKFWKGVKVPTGEKHLRGGEERRGRLLNKT SARNL
 18 ERRW-----IELTSSRAWHADPMAAQEGFVLGL-----NR
 19 RVKLDLAI-----KLGQTGRAGGT
 20 ES-----PGPSA--EEKSPAGDAPAASGSPTAS-----
 21 -----
 22 >GBPU01010966.1
 23 -----MPSN
 24 KPNTAGVSLEKLT LVEAYLLGSALVPLSDSLPSRSRGGAVT PRP-----NVGPRNPVG
 25 PRRS-----TLHLPLSGTRSSGANPI-----KIPPLLPKF PARSSGKMRLI
 26 -----DLIKTSSKSVSSKRKT-----PVLSKKNGKAKRKGRSSAFSHKW
 27 VRDNRGRVRWDWTKLRALAPELASLDRKARGLCEQA WTAVRAALA ADGRRRATLMT
 28 SSGR
 29 TA--VRSYCRLAK-WLVHTTA F-----SGITHALKTLK-----DWVSLARQYVV--
 30 -----EGVRPKDRKL G-LAR-----FFRGDFRH YV DSDK-----
 31 -A-LLQLSYIGRA--LP--LGG--PSVVRKALIKHKEV-YSTEHHTP-----AH-----
 32 -----LLEAARRFASE--FAQD--RLGVVPLLSS-----GFSES--ACLERP
 33 AAKG-GSAA--HTH-----AVYNSLVS APEWEEAKREWLA EIAKEDLNIPDCRDIA
 34 VR-----V
 35 AILR-----DFSLRKL--KESLE
 36 ETSG-----
 37 -----IPKAEVTV--VTERG-LKARVVT KSPW ALVQL-GHPVR-----AWLLS-AL
 38 RKDPRT-----SGVLEGDH-LNAAK-----GVVDAPRLAD-----TR-----
 39 -----RF-LSA-----DLTAASDRLPR--DLV SALVEGLIEG-SG--ET-
 40 -LPEW--ARWVMRVS-VG----PMLL--D-Y-SKA WNP KLG-P-----P
 41 LPGDD--KI-----I--SRRGILMGLPLTWVLLSLTQMFW-----A---
 42 SQG----W-----KAILAHMTRSAGPSAT-----
 43 -----PST--VICGDD--LLGW--WPQVVI--DQYQSV ALA-CDAQFSEGKHF
 44 VS--TRR-AVFTEL M-----F-----E--ARIRPA--SNVSP
 45 RIALWER-----RRF-----
 46 PRLTFAPT-----KKPKWRKVL
 47 LGPQPQC PGQR DTRRV-----VTGLR RTPLL T RGM VEPE--AGLDSSY-----VRQ
 48 PRWINLG--LVYQEVAQ-----SGRPAA-----ARRVLR-----TLYPGYWS-W
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 Y-RQK-GFQPT--LPRFL--GGPG---L---P-----PLTGDVN-----
4 -----RLHLP-KPLR----LGVAALY-----TPWRDIR----CPSSLWSTGTAK-
5 -----WVREMA-----ADSAEWKFHKFAKS-----SRVGGKVPY-----
6 -----GGLGVL-LGGPQAVEETTRLAN
7 ELSLIL-----DCDPKESSRWK-----KKP
8 REVAKALNGFYARVLKRKP---GVRNWVRKGVSKGVPVG-----AL
9 EARW-----RRVQNAYAWYSHPTAAAEGFC-----LAL-----PQ
10 GVKSNLSV-----ALRQSGVPARV
11 ERKKS-----KMARE--PSLDGPDDLGLQSLAEFPPLPSAGASRRR-----
12 -----
13 -----
14 >GBWH01015246.1
15 -----
16 -----
17 -----
18 -----
19 -----
20 -----
21 -----
22 -----
23 -----
24 -----GRA---LP--VGT--QVVIIRRSLEAHKVN-LGEEYHVP----EP----
25 -----ILERARAWAKC--FAKD--KLRSVPLITC-----GFTQS--ACLENP
26 SSKG-GSFAITLSRC-----ATLETETEEQRGGRRTQHEWKKSIDKEVLDAP-----
27 -----
28 -----DCDAHIATRTAALR
29 ELAL-----EEIRKQ
30 VEQHGYPMAEVA--LPERG-MKARIVTKSPWSLVTL-SHHIR-----VWLLH-GL
31 RSDVRT-----SGVLDGKE-KEAVE-----GLVDVPRALAD-----SR-----
32 -----KF-LSA-----DLTTASDPLL--DLVAAIVDGLVEG-SE--GT-
33 -LPAW--AQEVFRLS-TG---PMVL--D-Y-SKAWDGSV-----P
34 LKGPT--QM-----I--SRRGILMGLPTTWTLLSLVQLFW-----
35 SDT---A-W-----RNGFPEIMARRGFSAPS-----
36 -----PAT--VICGDD--LAAW--WPEKVI--EEYEALART-CGALFSAGKHY
37 RS--PKN-LVFTEVI----FRSVVSTAQDHPPKRSFWERERLPFALIGKTVR--SSNVG
40 AKAGQVEGRNLTKKE-----RRAKLRQSERTI
41 TLVRRDG-----
42 -----RRVSSLSSKDDRRVVRKLAVMPAVPLRGLVEPE--LGLNESF-----IRQ
43 PRWVN LG---PAVDCAIQI-----SGLPHA-----VRRILK-----ILYPGYWA-W
44 F-RQQ-GFEPT--LPRFL--GGPG---L---P-----PLTGSPA-----
45 -----RLRLP-RRLR-----VAIGILLYA-----TPWKEIR-----SPAAV-
46 -----WTCLGA-----GSVRESAREHTEWK-----FSKSAVACRKG-----
47 -----KVP-FNGNGAVLGGHSAVED
48 -----VTTRM-----YNELALILDPE-----DRR
49 SRFKVVKPREVSSAVRKFYD---RLLPRDPKQRVKT-----HPYITRGIAAPSVPVSAL
50 RRRW-----EGILGQRTWYSTPGASSEGFCGL-----AA
51 GVKHELSL-----
52 -----ALGHSGRRTRRSPS-----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 >KX883605.1
4 -----
5 MVARLVEDRRRDNRSGAAAKQ-----PASRLGGFM
6 GHRR-----GDRQRVAALFPWRRRLKGSE
7 -----
8 -----WT-----RLSEAWRVIQAAIAAGFLTVGLED-----PA
9 SREFLRGLCRLAW-WMVKTGHY-----TGVERVLSDLK-----GFAGKARLWVM---
10 -----EETPPTLRRGS---FAR-----WFRGIGRSLVDRDE----
11 -R-AMQLSFIGRA---LP--AGG--PRVVAQNLKFREV-VTA--PYE--TAPH----
12 -----LLSEMEAFAKA---WASR--YLQPAEIPPP-----WMGEG--ATFERT
13 RTEG-GFASQSVEA-----GLFDQAYRVDADWDELIRIRAAEV-----
14 -----A
15 GMTD-----CDVDIACR--EAQTR
16 EVAM-----GTLSRLRR-
17 -----PIAVEALA--IPERG-LKARVVTKCPWALVYL-GHFLR-----SWLLQ-GL
18 RRDPRI-----SAVLEGEHQRAVLD-----LLKSETAEVTD-----
19 -----SILLSA-----DLTAATDRFPH--DLVQACTRGILEG-AR----
20 -CPLNSDWTQVFADL-TG---PLLV-----V
21 PPRGE-----GSDYLTARGIMMGLPTTWCFNLNFNLFW----S---
22 ERA---W----SLCPRVNPVIHRASLFRPGHHART-----
23 -----PAT--VICGDD--LAAV--WPHDVC--DQYERIAVA-CGASF SKGKHF
24 RS--ARY-LLFTEEP----Y-----EVLVSETERR--HGRAV
25 ERPQEWS-----
26 TLADYLFP-----
27 -----SRTRSALEPSGRGSGWHYTGFRRMKSVPLRGLVRPT---HKPFDH-----HTI
28 PPWAALP--FAVQASVRA-----SGDGAR-----VRRILK-----VLHPGIWA-W
29 A-RNN-GFSPT--SPTEL--GGYG---L---P-----PVHGRRG-----
30 -----EGYLLP-PWHR----RGISVLLRH----TKWKDLI--SPSTA WTLIESGSH-
31 -----WRGMGE-----EHTRHKLA VS-----AVACKRGKVPF-----
32 -----EGKGSE-LGTVDCVEQLTTRMAR
33 -----ELRMLLGPERQ-----PRDEF R VAPRQIA VRV
34 RKMFRGLEANKGVPGGRLP---SVSLVPASGSVQR-----L
35 LSEY-----ESRLTERCYFSAPGPAREGFVLWL-----PG
36 SARRDLAV-----A-----
37 -----LGWTR--SSAHADD-----
38 -----
39 >KX883602.1
40 -----
41 -----MPRNKARHRWVCLYELTLSVATLLP-----ARP GTGPYL
42 GGTR-----PNG PDTGRFLPRF-----PRHPNIRCPK ERGNAATGKT
43 -----TNGNVNGILSKV LRRPL
44 GN---QRWT-----RLEEA WRTILAAIAACGFVRLDSSR-----RA
45 HREALRGLSRLST-WIVKTAAL-----SGTEEVLRELK-----TWSARLRAHV V---
46 -----NSLPPHKRRRG---FAR-----YFQGVLRSLIDKDS----
47 -R-ALQMSYIGRA---LP--VGS--FRETEGAIRKFREV-TQV--PFE--THPD----
48 -----LLDSIRDWGAH---WAAR--FLDTPAMPPL-----WRSSS--GCVDFS
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
 2
 3 RLQG-GLA-----EAAVRAADFAAAEVESREDWDDL RNTAAE-----
 4 -----V
 5 PEMT-----DCDV DIAVREAQLR
 6 DVAL-----TVAAHLPD-
 7 ----LPPVEVCT--VPERG-MKCRIVTKAPWCLHI-GHFLR-----SWLFG-GL
 8 RKDRRV-----QGVLAGDH---AG-----ALRGQTERVYSQKPTEEE-----
 9 ----NICLSA-----DLTAATDLFPQ--DLVAALVDGLLSG-AR-----
 10 -KPPTAEIQEVFRKL-TG---PLNA--R-L-----
 11 -PDGD-----GFAIRRGIMMGFPTTWTLNLNVNLFW---S---
 12 ETA----W-----SASGSLPFQDVALSAQPAAS-----
 13 -----PRT--IVCGDD--LASV--WPPKVA--DRYEEVAEA-CGAKFSVGKHY
 14 RS--RDY-ILFTEEI---F-----K--VSWEKQEYSLMDS
 15 AKP-----RRT-----
 16 TLADFFPV-----
 17 -----PRAPRKSTM RKR YRVQAKLAPVRCVPLRGLVRPL---HAPKDR-----RLR
 18 PSWAALP--AAVEAAMD---LGAPVA-----VRRVLR-----VLHPGIWC-W
 19 A-RKR-GFSPT--LPLVV--GGYG---L---P-----PLRGSRV-----
 20 -----RVSLP-KWLR---YGLAAWLLN-----SKWKDLE--GPSRCWEGVCRPVS-
 21 -----WQTMAR-----EHADVRIASS-----TFAVRKGRTPP-----
 22 -----PGFHL-LGGL EAIQNLSVKFEG
 23 ELCV-----LLGTEEGSLKR-----QFGLAP
 24 RRVANSLRKFFARMVSDHP--SRRPFPSL GRTA-----L
 25 VQRW-----RSRAEERALYSTPRSAVEGLVSL-----PV
 26 RAKRLLAE-----
 27 -----ALGWV-----
 28 -----
 29 >IABX01132835.1
 30 -----
 31 -----PLPKQQHRWVHLS-----RVSLVGAAS
 32 VGAR-----RGAPLSIFFERASTDRQKGRLNPESAADYCLAEIARVSR
 33 -----LAWTPDDTGIPKGPLYKVF
 34 VSRPMMPRALRS-----SLVETWRVVLSSLAAAAGFIRLDVSK-----WA
 35 GREALHGLCDLAW-WIIRTSHL-----TGIDLMFSQLK-----GYCSALRLAVV---
 36 -----SEIPPAERRRG---FAR-----FFRGVLRSCVDRDA-----
 37 -R-ALQISYIGRA--LP--KAS--PAQCTAALENFRAV-LKRPHHTS---PE---
 38 -----ILDRAYQFAVE--WAGI--HLTSPEIPNP-----FLGAG--ATLETT
 39 RAQG-GFSE---ETR-----RKAGQFDWQLGAPDQEMWDERRMLLNE-----
 40 -----I
 41 GSLT-----DCDV DIAVREAMTR
 42 DVVL-----DRIHRM GD-
 43 -----FPEVSAIA--LAERG-RKTRVVT KAPWEIVYL-GHFLR-----SWL LD-GL
 44 RRDSRV-----STVLMGDH-AEAVK-----RLSREALYCR-----GD-----
 45 -----EI-LSA-----DLTAASDLLPL--DLVRAITDGLWEG-TRKG---
 46 -CRPEERFRLVLDRL-TG---PLRV--E-M-----
 47 -PDLY-----DAFISERGIMMGLPTTWCLLN LVQLFW---C---
 48 EFA----W-----MSGQQYRHMHDPLRGSPRSS-----
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

1
2
3 -----PNT--AICGDD--LVAY--WPSDII--GRYEQVVKE-SGGMFSVGKHY
4 RS--GRY-AVFTEQF----Y-----D---CKVTRFSAPGPAL
5 RKGGER-----
6 TLADFLFP-----
7 -----C MPSKGGEITPWNVQVGEVRNVPTLPLRGLVRPSHGPDRHLM-----
8 PSWAALP-LAVEHTMMYR-----PENVKA-----LRRRALR-----ILHPGIWC-W
9 A-RSR-GFAVG--LPTSL--GGYG--L-----PTLSGCP-----
10 ---SRCRLP-KWVQ----VGVAALLFK-----VPWRDII--SPARYWQMVRQNLP-
11 -----HSTMAA-----DHARYKFETS-----AVACRPGKSPF-----
12 -----NGNGHG-LGGLGAPDEVAVRLAN
13 EL-----FMILEDLPPRR-----PELRVRP
14 RTIASGVRKFYQRLLTKRYP--VGIRGVKRVGKTWM-----DL
15 QRQW-----DSRQNERLWWTNPRSASEGFVLGL-----PR
16 SAKRRLAS-----
17 -----ALGWV-----
18 -----
19 ----->KX883500.1
20 -----
21 -----
22 -----
23 -----
24 -----
25 -----
26 -----
27 -----
28 -----
29 -----
30 -----MEAVPVTSSRRRCG-FTK-----YFRGSLRSIIDRDC-----
31 --RALQFSYVGRA---LP--EAS--PGQCDAALTKYREE-LSQPATTP-----DS-----
32 -----ILDSAEEWARR---WASR--FLSGPVVPPT-----WLSGG--ACKEVP
33 RAKG-GLGT---RSV-----DVARRATDAICEPSEPEWESLRERLNAVEGLT-----
34 -----
35 -----DCDV DIVCRERLVS
36 NQAL-----VELSELPR-
37 -----PIAVDAMA--LPERG-GKVRIVTKCPVALVYL-GHFLR-----VWLLE-GL
38 RRDERT-----KSVLDGDH-AGSVG-----GLVGGAGVRLLE-----
39 -----HLAVSA-----DLTAASDLLPH--DLCQAIVRGVLRG-SR-----
40 -TRPAPHILDEVWADL-IG---PLTI--Q-M-----
41 -PDGS-----EFVNQRGIMMGLPTTWTLISLVHLYW---A---
42 EWA--W-A-----STVPRHISLGLPSAA-----
43 -----PRT--AICGDD--LAAV--WPTSVV--QRYESIVKS-CGGQFSAGKHY
44 KS--RTY-LMFTEEA---F-----R---LRVEMR--DYDTR
45 RVDGRRP-----GEA-----
46 TLADFLPS-----
47 -----WMVEEPSRDSRW-----CTGV DHLPLVPLRGLVRPL---HLPKDR-----QPL
48 PVWVAVP--HCVEAACTQ-----SGDAGA-----VRRILR-----VLHPGIHR-W
49 A-RSR-GFAVN--APTS--GGYG--L--P-----PVRGDPK-----
50 -----RARLP-KWFR----WGISSLRA-----TPWKDLI--NPARAWAPVMRDIP-
51 -----HRDMGA-----DHASEKLKVG-----AVAARKGRVPF-----
52 -----M-GRGADLGYDAEDRLTV
53
54
55
56
57
58
59
60

1
2
3 RL-----SNELLFILD RD-----TLPSSFFLQGP
4 RKVSGAVRKMFKQHLRRFP---IPEGSVVVA SETPM-----ASL
5 RRRW-----EEVRTARSWWSSPRGAAEGFVLGIS-----SS
6 AKRAVASA-----
7 -----LGWAS--MEPAHSGGPVSD-----
8 -----
9 -----
10 >GFNE01046070.1
11 SDLSGVRRPQTRKVLLSRITYVEATRLQAARCRS VLPSVSGRRTPAVCRSRGGWVTRR
12 R
13 KCRPVAPASNAFLSSRTGGGLSLHVADSTLSSLSQPPSISLLS-----SKAFSTGLA
14 HSPTLIPLLMQDPTWPSDRPSAAVPAKAVTTAKHADPRPPGRAPAPQPPPNNPHPPGPK
15 GRVPARKVGPRPPAGAGGGPGAGDAARKHQPPRAPFGTQREHAVVPGLVGTPRTGQAF
16 RT
17 LC----GAWD-----ALRACIAAAGFVEFDLRE-----RS
18 SRDAVGGLAQLSG-WLTTATRG-----SGLGYTIKTLK-----TWSAQFRSWVF---
19 -----NRTPTHKRRGG--LAR-----FLRGKLRLSLVDRDE---
20 --RALQMSYLGRA---LP--VGS--DGVCEKSLKNHRDA-LTRPHPTD---PD---
21 -----LLARAYRFARL--WADR--HLPEVPNPSHV-----RLGHG--ACKEAT
22 RATG-GFAR-----AIVHLAGRDLPENTEARQTIMDLRQEEDLGPE SITEC-
23 -----
24 -----DLDSFVA--ESELI
25 RAVV-----KPL
26 LDAPRPPRHS AIP--VAERG-FKCRVVTKSPAALT AG-LHHIR-----TWLMA-GL
27 RSDKTS-----STTISGDHRAAVEA-----ILAGDPVPREG-----
28 -----YCFLSA-----DLTAASDLLPR--DLLIALVDGLSDADHS-----
29 -APQW--YFDLLRLA-CG---PQWV--S-Y-----
30 -PWGD-----TILTERGTLMGLPHTWCMLSLVHLFW----I---
31 DQA--V-W-----TTPATRIPPALL-----
32 -----KRS--AICGDD--LAAH--WPDRV I--GRYEDVATR-CGAEFSVGKHY
33 KS--RDY-VCFTESL----AKARCNPGCNPPTCSGSGGRVPRY--RRTRPR--PAPQT
34 PIKR PYGSASPIPSPLASAVAADPHPALLAKSLGLTPPASAPP SRSIPTPTSKPKGRSR
35 PLLPNLPPKGRRRRGGRPVALRAPDPSPRDPFPYPPTVPGRPDPGPRWPRGPKRWPRLG
36 AGSGAKLPRRRRRPPRSNRRIMRRVVGVWRWSNAVVRGLIRPNVVPFGREAL--
37 GDTMGPA
38 PAWAVIG--PAFWAIVG-----ENPPPP-----RLNAAR---RALKALHPHL DA-W
39 G-RHNFGCLAT--VPREL--GGLG--L--P-----KSKGDAL-----
40 -----KVGKLP-KWLR----QALAVDLYG-----TRAGESP--PPIGTWSGMKPGHH-
41 -----RSMAEE-----HSRARLQRM-----TIRCRA GKVPF-----
42 -----KGAP-TPATLLGPEVD-----
43 -----VEELSTIAMS R-----
44 -----SLTYILGAELLGDRL--SVNPRRLADRI R-----SMVSR
45 STRW-----AGSRGICQSAPWAKLIARHRNLVQN W-----LW
46 LKPGADEG-----FAGALIDKAAKRS
47 LAGT-----LGWAI--PRASAGGSPARPTS QELSGP-----
48 -----
49 ----->KX883526.1
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 -----MP-----PRKSNGRKG
5 GR-----APRTCVALEAVLSAGAAAGL
6
7 -----TS
8 AEVRTLPIGVLA-----SEDNPDLGALK-----ALSAAYRAHWI---
9 -----ATGDP-LAG-----
10 -----QLSFVGRA---LP--PAP--ASARLAALVQHRED-LTSTFRTP----PS---
11 -----VLALAREFATR---WARR--HLARAAPVLEEAS-----WPTPS--SCYERS
12 ASKG-GLLR---HLL-----ECTSFEPAPPGFGLPDGPVTQLWVQSA-----
13
14 -----RFREYGLRGLERCR
15 SAAG-----
16 -----PPTHRVAV--LSERG-LKTRVVTV GPAWLQVL-GHCVR-----QRLLR-AL
17 RSNRG-----YAPLAGARDDEIWA-----SLEGGHSE-----
18 -----AV-VST-----DLTRASDLLPL--DLLSAIIDGLQES-GK-----
19 -MSTL--EVEILRIS-AG---PQRL--E-Y-----
20 -PHGW-----GTYSSSRGALMGLPTTWCLLSLVHLFW----M---
21 DQV---R-----LTSRAGPMRAG-----
22 -----HRF--SVCGDD--ALLA--TTEAGA--QRYSALVRL-CGGQPSEGKHY
23 VS GGVR CRGV FLEKL---L-----E--FDTEGG--RL---
24 -----TGM-----
25 RRLAAMPV-----
26 -----KGVT SV-----SLPRDFTGALPIQCRS-----KGLR LLV-----TLD
27 SVLED LG-----PAGL VP-----VRNW VR-----TRASWLGRF
28 A-REQ ALLTPG--APLR Q--GGY A-----FGDP N-----
29 -----DR TPPE EERR----FRAISLLPRPSIGA-AALRSVD-----PL-
30 -----WRLASR-----ETLAGE-----DLPLSSGKVVE-----
31 -----IGRH PPLS-----
32 -----GEGPLLPTQWR-----
33 -----QSEL DRLAEG YP--AVAIQSQEERHTS-----AC
34 VALF-----AALKSLMGPPLEVYTFR-----VSQ-----LR
35 AAVLKATA-----TG-----
36 -----EALRE--SLGPTA-----
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 >GEZJ01006036.1
4 -----AGG
5 QDNVPQTGGPTRPARNYLLSSVKSLASVGRKV----LGKPRTLNLGVLAQMDSRHK
6 KPLSKPQKKHL-----
7 -----
8 -----RELRRNW-----SATRG
9 FLELTGVLKGKIF-----APNNRKAIREQVS----FLLEVLE----
10 ---SGL---QVTAKRWSSLLRSKSI-----YQGTS
11 DSPC-TGIGTRGKTPMALF-----LASTYSRACVVDP----PTKDEV-QTMV
12 LAAIERISTREPL-----PEPD---VLRHLKGF--IADLFRSQRMKPRRAPVLP---
13 -----VPAPTSCYERSGSEGGSI-----IALT-RFG
14 RQDLSEAVHEHIARLFAESQRRAQPEERHDYNWTDLDEPLYEGIPRDPQEMIRCLDRQ-G
15 A-----MDLMMC-PPPPP----RLLT
16 LPETFKKAV---KDSSRPE-----ER-----
17 -----RA---KALPIVQQDGKIRVATLHMSGVV-WSARAMS--AFLLPRLRNLT
18 FS-----KDMLRNRTVQLESHRRSD-----DLLVYSADLSKSTDPI--SIEL
19 SRFVLQEITRHTG---KPTWWNDALEAVITNHQI-EYTA-----
20 -----DSVKKTVM--SECGALMGLPGWTVLSVLNAFAAAEAAAGA-----
21 -----R-Q--GDYAVCGDD-LIGLWPTEVCDRYEGNLASRLKANKEKSFRS-
22 --THR-GVFCERLVLRSDNRASAR-----
23 -----SCVRLG
24 EASGA-----
25 -----RAIDGGKGRVV
26 V-----DALREVAS-----SSEAHRIIRQTAYRTIMRMALSQ-
27 TAKGSMKE-GGGGIGQADAVTLISYLLHGPTQLYKVETDSRIAAMKTLLR-----
28 KAPNLSTGVPVEDVLTAGRAQI-ELNRRLETAERGPRPKYLPQSQR-----
29 -----RRIHSRRKDAKKAIADYGGVINAVER-----GLLASGCYVR
30 ENNRIR-----RDLTHAVRG----RAYTRALRIARISSWRTKV-----
31 CPHAAKEAFEVS-----FPEHQPRVTSDELPM-----
32 -----QAWDSRLAD-----
33 -----
34 >GGIC01436417.1
35 -----
36 ----KNQRISNKKIKKNNISKSKIYKKIVS----PEVKNYIDKLGLKYKIEDMEKI
37 KDKEV-----
38 -----
39 -----EVLKDYV-----DEITH
40 QHIKELKDERKR-----SRNECVKDIRLFV----YLVRSL-----
41 -A-SGPNQLIKGLKEKLTSVRRYCMGDIIEEQEIVDKHFSQVLARDQEKRHLSKKERQQRR
42 LDVRSDRQKILGLYPQWLELPGIRKFTRH--QHIQMTFISRSLP----EGTRNV-KDEA
43 IEGHKKMVSSGGK-----MDEE---IRRDLVRW---VREYCRGKVSDIQLLLAG---
44 -----KWSATPSSCLEATRKEGGHLMGSYQLYNMISPCRHTLF
45 NQSRSEDIERMNWERHGGDRFCVVSLVATSLMIHESEPLKEFVTKYLHEAKQIPSLP--
46 -----NVIQERHNLKIAPGFSFRSKTD
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 CKTFVKQQKRRRLGIKDKVVWDEKLGRMVDNFDDRRILSPFLPSGDQSEVHSTLGGILSTE
4 E
5 FSDYAKSCFDEGGLRSRVS AIAERGWKS RVTNHSELEI-QSGNIIR--NLLNSILSKDK
6 RFL-----GSRRTHFDDQEVL DAMGRNSVGG----HPLVLSED MTAATDSI--YMDV
7 FATICWNTLCSTIG---LK FSSIHIAGLIMCAPRLY-EYKDL-----
8 -----TKVR---TNRGIPMGTPISWPLLNL ANMFCWERANI-----
9 -----DRAFEI-I--PPV TILGDD-LLGVADLT YINKYSDNLKRIGGISNPRKRVLS-
10 --VFY-GLIGERIYQFDVEDNEKSLVIKH KETQLEAEQTNYQVHAEIAGKSSIKTLLSE
11 DGKRKDLVELGSSDYRISSSQARYTESLKMML EDQDCHTGRRSKIVYELEIVKYIESDLI
12 KEDRK DIEEKLRFVVKKD TEDRQKEIENQIINICKIIYKCYGKGRSFSGIIGPFEN YKG
13 IVEWLFGSIVIMIRGV TYSISTRRKDCVT SKDG IRRPLWVRQTN G KLEKKAYEEGPV
14 NLDPGF DGVELHLNQYQKKFKTLLRQWHSPRRN RGLVISKRLKCTS LK-
15 GIVSGTNKLGA
16 TDINK---GIKPRWTC-MGSIEHFVLES--LGDSNEVRQILNRIRSIFHPDIIREL RGH-
17 RLEPFIPE-SMGGGGLVYKNKLDSTRWPCS-SVFKKYINKILFGQNVKDYLE-----
18 KINTSLV--I SEKQDWWE LASE-MG-RADFEDVKIVER RLLHPSIEG-----
19 -----ITIGNHQDLVTENILQYQGD FIRQ---MGPDPFCKKVSKLVRG
20 KTKQAM-----KKEYESWASMRGF---KGSIEEFIRFRDIYLSRF-----
21 --MAITIRQEEL-----STYL RRTKGGFPLIVP-----KEHTHLNYITTK
22 TNTSQT DQPRGQAETV LNADVT VVITDGLLETIVPSEPQRQLIAAA--GRSMHLTHIL
23 PGLYFYTNPAVWF KMLESTP-----
24 >GDKJ01026489.1
25 -----STWGGMVAKTGLGSTSRKTAASGSISVC----PPRSTQQKPSGKVAKKVSHNLSQ
26 LTTLGSLLL V WQDLV KSN-----
27 -----KALTRVQ-----RALHL
28 YKA-AADGKGNNLKVFREKKRRSRTMKLRFKAIETAFLPILHA----WLLV-----
29 ---NGGASCCSRPNNIGKWMSF-----TSYMM
30 AMLT-DSKYLKSITHQWREESIRHSNTRVAKRLFQASTLGRAVP----FPHTMLQEDRA
31 YQGAIDAWTQGP A----F--VSKQ---HLDQIAEF---ISKLPIRRSKEKPG--LHD---
32 -----L-----MLQN KSCIENT QAHGGIA-----HYLYTQEV
33 EENKKRRDARLFKAGDLIQGQATTIRT KILMESFSRGEELDSRLLEGQRATMSQFKRL-Q
34 Q-----EELEDI-RRQYR-----YQDL
35 L--DKARARFENTCHATQPM-----
36 -----NPVIVREL G-KLRCVTVHPADEV-ITARTIT--TKLLRVLR GCV
37 ST-----RDTLRG-QCPTVRTETG-----RPVLYSADLSSATDHV--DHRL
38 AQFIGEKLLDKMG----LESSKPLLKSIFSPKEL-RFTKKEY-----
39 -----KKHNGTLTSRGVHMGLGPSWVILCLINMYAAHNAGY-----
40 -----N-K--DEVQICGDD-LIALG SERAAQRYQQT MKE LGMV VNETKS FVG-
41 --KR--GVFCERFLEIKVDRIGCYASAL-----
42 -----DV GHIA
43 EASAV-----
44 -----RTTSDPDKSRLAAVERLYKE---RHLPTLARKVADQLAPRG
45 VGK----GPIPMGGNGRGA PTT PQLVH-----FMKCGPIRLVQTQ-
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 ---FSPKLELNASRVPIKDGIPWSGLRTEMMRHDAQMSYFKNESKPLTIIRAE----
4 HFSRLAR--GYVRSTATVHLKS-QSGTYLSSLPRAARQKVKLLLKRF-----
5 -----TDSSNTINSNLSPKATRAALA-----
6 -----KALTGPTAE-----RFVNLSTASNLLGLHH-----
7 -----CPDADIDLALM-----RSRAKMTNDGIRLGPK-----
8 -----EAEELRRITNLDL--L-----
9
10 -----
11 >KX883566.1
12 -----
13 -----MPLQSMPSQKVNFGLSAVA---ASLRRRLAREGRTVGSSQYATFAR
14 LAEKLRLGLSPRQRRR-----
15
16 -----
17 -----CLSTASW-----VAGGL
18 SVSTTDRSKGS-----SLGHRLAVVKFAI----KLWNNII----
19 ---SDVEQFVVDYKRVCNSAAYAAV-----RG
20 YTSIVAAGKCRDYLHRAFGPCGIIPHKGRR-SLFALASTTRCLP---PASLTPSRVKIE
21 LEALKSRLTSVPP---P--VSGA--LLSELVSF--SHDLLRDAGPPVQVEVTGN---
22 -----ALGPPGKTSLDRAQLYGSYA-----
23
24 -----FKGVIMSKTL-----L
25 M-----MSKSAI-PPDV-----FKED
26 LEENVQ-----
27 -----RVSLLAEAGDKVRAITVPSSLAVQLAGRSIN--RALIRCLRRSD
28 VF-----GPGFGKRPGTLMMSGANLG-----DKHIVSCDLTAASDYL--NQDV
29 AIAVLADYLRSHA---SGIWYPWASALLGPQRI-RS-----
30
31 -----RD--DTYT---TSRGVLMGTPLAWPILSICHAFACYKALG-----
32 -----NTY-Q--FLCРИVGDD-GLLICNASQYRAYVSVMNRLGFVINEHKTFQG-
33 --KQC-GILAGRLYVLK-----
34
35 -----TTRAKF
36 QDLNS-----PT-----
37
38 -----KTMAWRMVDCQA---PYVP---ELLAARGA---
39 -----QRFGRSDA-PTAPIRVPGLS---QRENKAALRAFK-JLNPRFTDAMLRA
40 GRPTPLTT-VQ-----RRGLLVLKNAAP---SSGARSLWKSSLAKRGHYASLMTSYLR
41 QLMDINL--PACSKAPVNNGSYL-AR--MSGTVSQLVTARLNLPKDSR-----
42
43 -----MILRPSYVLSRLGKRLDKK-----
44 -----MHTPKQSSV-----LRLLRGLEAE--MVPPSH-----
45 -----VSVMEKL-----LRSMVGPDDGARPISG-----
46 -----VIGDFSQLGI-----TSPTELYPQFL--HWIEEGAP--
47
48 >KX883471.1
49
50 -----KSAPNGK
51 DTRLKPAKPAVIKIDLSHPAGSRVDPKVADKN---KPRLEFASLDSLRQVSLLARSSL
52 GVQRISTIPLSPDPSKSAA-----
53
54 -----QLFRRFAARGKQNIAKFQRLW-----STVLN
55 AV--SLAFKIDQ-FDLRTPQ-----GRIAYTDLQCAF-----HLLARFT-----
56
57
58
59
60

1
2
3 -T-VGYQDSVRVVAFWALQARQMAL-----NSAVA
4 TSG--HRPNLVKFLRGRLDIRRASQR----TKFLSGVGRALP---FAKIDAIR-IDAE
5 IKAHRARVTV-PG---T--TDPL---ILGKYAEWLQKKLNLLDSRADLPSMKNIA---
6 -----PL-----TAARSSAVGAPSSKGGRY-----GFYK---
7 --SLFEQGPQFQ-----ERILEH-A
8 I-----KEVEDC-PTPLP-----V---
9 -----
10 -----
11 -----KLTVIPDKGDKARTVSTTLPGIQ-VLGEQLN--QSMMVLLRRLA
12 PS-----AIPLSENPDVSRRLNNG-----HGTMLSTDLSAASDYI--DHTV
13 LKVTRTRLVARKLG----WQFNSKLVRLLTGPFMF-TIDQ-----
14 -----EGVKDSYL--GLRGAPMGLPLVWPTLSLMNMFCAYRIS-----
15 -----PNS-K--RSYAVLGDD-HVARWPKADIDQYFRNMSDIGLKINKAKTFLS-
16 --KRD-AVFAEDLYSLVNGKVKR-----
17 -----
18 -----
19 -----
20 -----
21 -----
22 -----TVPKLSAVMSTRQQ---
23 -----GTDPQIFA-IIASLQSALAN--AKNFKERKQIKKIFF-SQWKKLTRQIRRA-
24 QVPLQLPT-QLGGGLGIAP----FKLSKLT-RKQRSRLAYLLSL---PEYASKSARSDQR
25 SYFRSLD--QFKNANRFAHAIR-NKDKLVKEAARQVI-EQISDEEERTSR-----
26 ---SFNATGCAVEDCAPQLPEGGVPAEAVLKELLALKLTRL-----TQEGL
27 VQATDR-----SVSSGTVLR--RL---TRVWERLPLCLKPVGARG---YTKLLS
28 LCVQSSQILSTQ----SVAKLITTNLPLYAV-----
29 ---GVFYKDGGMPIQCS-----PNSNDAYDSAT---AG-----
30 -----
31 -----
32 >LC150604.1
33 -----
34 -----MASQFCDSIA---RRCGEVIRITLHRSNDHTTSPPH
35 GGPRLLEGNLETFPPIAKRD-----
36 -----
37 -----KHAKRVM-----QAFYA
38 AYIAMGYHCHLR-----DRSLTRFFEKLHYI-----RAWH-----
39 ---EGLDRVCPDIKKQVTLLRS-AL-----IAETG
40 VSPDVVLGKLLPRSWANLAHPR-----KLFQLTCLGRALP---RPTQKV-IESS
41 KREALDGFTS-PA---SVTTDIN--VLHELVPM--ALNLRRYSAKADTTIFRGS--
42 -----SSTEQSTLGPVFGTSATLFNSRKEGGRA-----GIIH-----
43 --DACRSRPRQV-----GAIVALKRQAAIRKTAAP
44 LR-----FDIRSL-KSEPL-----IGED
45 SRRVLNDVRHRFPNNLGES-----
46 -----EIAVVPELGFKARVVTKSHPITV-VAGHQLR--RRLYPTLYQSR
47 VF-----SRALGDKPKALRFARRS-----GAVFFSADLSRATDGL--SHHT
48 VGVFCRNAEIDPD-----IIFRNMSV-----
49 -----DG--HALK--RGIFMGLPMSWTILSYIHRACDSDVDP-----
50 -----I--QNYYLKGDD-LVAHWTARQISLYRERVASVGMPNESKTFVG-
51 --PRK-AVFCEGYYESSKASFNKK-----
52 -----RMRTEI
53
54
55
56
57
58
59
60

1
2
3 WLVRQ-----
4 -----
5 PSISLR-----
6 -----RFFPDSEA-AALRLSEGLDE--VDYNLRSQYNRLVY-CFYPNYIRISRKI-
7 RLNPFLPP-QLGGSGLLIPDLNSQVHS---VFDRMRLNSIFGGVKRFNASI-----
8 -----PVGGSSSNMKQVN-IL---LSRVRMVP-----
9 -----RGSGLTCSHFDKWLEFATSAAAFR-----
10 -----DAKIGIFPK-----DPSDHTFFAEAKRISRDR-----
11 -----RRGVDFPID-----YRTVRSVCRSLEPLLS-----
12 -----TVPLRHTCE-----
13 -----
14 -----
15 -----
16 >GBSG01013692.1
17 -----
18 -----RMVRHTNKRLTSFKSPP---CGAGSDSRKNDRTLHRGRRK---
19 -----
20 -----
21 -----SVEQQVW-----TACYS
22 SLVHS-----GWDNTNMAWFHD----WVSKTLY-----
23 -S-RGVRFTVGELKKLSHGVRFASLRA-----TGTVP
24 DAPRIPKKVYSAFLYLSVREKR-----NAFAFSRLSRGLP----CPPKED-EIKA
25 LRDAKEMSVSYHP-----TSGK---AIE-----AIMWHILQERGGKT--RLP---
26 -----D-----RLPNTLSSCFERSTTQGGVD-----GHLR---
27 --HLGFGELINI-----QCQYMGCSSADMIRAKYSSFATDTLGQF-C
28 L-----KSICDV-ADNFD----STKD
29 DEALRCLGVLKLRSDRPD-----PR-----
30 -----C-----KSHCLRTPGYKCRVIGVPDALTF-VEGTWVRWSSQLLPRKHDFP
31 SK-----NTCPAVKGRTAS-----DKRYCSVDSLKATDGL--SHDA
32 VRAVIHALAVSGR--IRKEDLSMSLRSLGVGDTEM-IW-----SWRIP-----
33 -----KEPVEELR---CVRGSPMGTPLSFIVLSWINSCATESF-----
34 -----THSSIHGDDAVGTMYNTAELEYKEFVRDIGATVNVSKYIS-
35 --PTS-FTMCERMYAPE-----
36 -----KTKGMF
37 TAFCP-----
38 -----
39 -----PSCPPIPGGVVVPVPPQ---
40 -----AHIPNLYL-----RRAERVQR-TIHPWVSKLPILR-
41 -----LPT-SVGGFGY---TGYGLKVSKAT-RCRLLATACSRNIADVTKEILKQSTFRGEG
42 LFPRSAV--PTPKNSSKYYKYL-NL---IIKDKRFSTVGS-----
43 -----EEFEDVVLASDFVTFREKVILQSY-----
44 -----LAIGGKLGR--GS---SERKPERTRARALFKGKA---PCNI
45 APLTVKRGIGAL-----SKLSDRVKNQKVVR-----
46 -----TDVACVIRGRT---THSK-----
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----
53 >GFLP01591397.1
54 -----
55 -----AYSTMRELLPSMLTYVRVGEHPSPS---ERTGPGSRPRDRKSRKAGSRGKR
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 QRA-----
4 -----SREDLVW-----QAVWL
5 GFIAF-----GISYQLGCWRTR-----WVKRSIR-----
6 -R-DGVHNTCSKLKDLAVSIRG-----YVLSD
7 VQPP-MVPGVPNRIQRWLRAEVDKDTL---SGLAFTRFARALP---KPDFRE-TANA
8 MASHLDLITDRRV-----VPEG--IQDQIEQHV--VRMFGGQFVGILRH--PST---
9 -----PISNHAVGERSKARGGYN-----AHIY---
10 -----DCAGDPEFL-----GVDPPLDRIPAYPPTRALWRLADSP
11 PG-----EGIPEY-MASAS-----ALSA
12 ARIVEQFAR---DTSRVVH-----
13 -----SATGIGEQGDKCRIITVPPASLF-AAGDVCR--SRIWPRVQQSD
14 AR-----IFGNDEDAKSSLSQMGLPR-----GSVYVSADLTAKATDGF--AHDA
15 BRAVRLRGLARCG-----
16 VDNHTVKLMSETLGVGDRQHYVKYKLAELPRRKREWAVQRFET
17 IEE-----DGVSTHVL-VPLVRGILMGTPCSFIILSILNGWCCMPLG-----
18 -----PRTVICGDD-VASACTPEAVDNYDRRTVTIGSGLHKRKTFIG-
19 --HKG-LLFCELYVLPGYGEHRCF-----
20 -----
21 -----
22 -----
23 -----
24 -----
25 -----
26 -----
27 -----EPVPLK-SLAKDGDGT--
28 -----LDTSDFDTFIW-----KRMDRACR-VLWRDVRAKARRL-
29 NRWPQLPV-ELGGLGHP---SSGKMGAVP-GSVRNRLATLIKVGPVKIAR-----
30 -----WT--VAAGPLDWRSFRS-DKEVAWSLFESTDAAVSQLEADSR-----
31 -----YVDTSFVSYHEANQYVSVLANQLY-----
32 -----SAHGGKFVN-----EFNRKAMKPGKIQYPGVG-----S
33 LQYSAKAPMSMV-----AREYLAKLHAGGEYLP-----
34 -----TDSVRKIRRRI--GKST-----
35 -----
36 -----
37 >GGCO01034162.1
38 -----
39 -----
40 -----SYTKNMKVRLSSHLTRLEFKDSTAPS---PRKSGGCKPLSKIRTKHWHSDKE
41 R-----
42 -----
43 -----TPEQIVW-----EALWA
44 ALVVSDSVNDL-----DVSSMRGSFYLRD---WVARSSR-----
45 -----RTVGYVVSVLKEELNRLRAISL-----G
46 SDPPPPIAHGFPKKLDTFLQKVFTPL---GLFAFSGVSRGFP----TCVDKVKVSET
47 LVDHLNVNVPAP---A--VPIN---FLRSIEQF---VPRYIQKSVCYDRKLADKM---
48 -----HDPSPVVCKSSPSACYENPRGAGGFY-----AYIK---
49 --KLGDKISEGQ-----NDLLADIRAGVKPRPEDICHPILKGIAKASYDPDRTQDLLSP
50 -----KEMSDI-LDHFY-----FYSG
51 CNEYLDNPDKEIVH-----
52 -----RVAVIPERGYKNRVTAPPASIL-SMGEVVR--SSIFPYVKAHP
53 SC-----DIVKDGDMPPVVCFPAVD-----GAKLVSADLTAKATDGF--SHDV
54
55
56
57
58
59
60

1
2
3 IRSVGVGLRNAG----FPQHVVDFVDTLGAGERPHYAEYDVHAILRRHSKNTVKYNSL
4 CSFLQSYGWDGISSKVK-VPMLRGSPMGTCSFTLLCIVNSWALDHCS-----
5 -----NGPKICGDD-MLAYITAKEFSKYKVRVAAIGSGVHPLKTFIS-
6 --PYA-GTFCENIYIKDPCDSKFP-----
7 -----SVVAFG
8 RAIKC-----PIKAMLYPQKGSNG
9 RIDWP-----
10 -----DEIVHQVFTGETRMSKEKVCLSSF--NDEP---EYHVQRRFEAG
11 PGDFG-Y-NDCPSVFR-LG-----RKVHRVLR-TLRRDQIRSAKK-
12 GRYPWLPA-CLGGLNY--PRKNPKGVRLD-KHCRARLFQFTHSDSLSELGKF---MS
13 SLDNAKR--PLPEHRELEFISD-MV----ADAEIND-AIVEDTLER-----
14 -----QGECSFVDSYSLARHQGSVRSNFF-----
15 -----RSMGGKLDL--S--ASSYSIFNAARIRWPAVK----GEL
16 GTYDKRTPFSLI----IQLDLNLLRAPRYRVL-----
17 ---SSFAEKLPGVEHPK-----SNRWHHLHGAG--GWNDGIL---
18 -----
19 >GAIS01005902.1
20 -----
21 -----EWRLQALSK----IRTKHWHSDKER-----
22 -----
23 -----
24 -----TPEQIVW-----EALWA
25 ALVVSDSYNDL-----DVSSMRGSFYLRD----WVARSSR-----
26 -----RTVGYVVSVLKEELNRLRAISL-----G
27 SDPPPIAHGFPKKLDTFLQKVFTPL----GLFAFSGVSRGFP----TCVDKVKVSET
28 LVDHMVNATPAP---A--VPIN---FLRSIEQF---VPRYIQKSVRYDRRLADKM---
29 -----HDPSPVVCKSSPSACYENPRGAGGFY-----AYIK---
30 --KLGDKISEGQ----NDLLADIRAGVKPRPENICHPILRGIAKASYDPDRTQDLLSP
31 -----KEMSSI-LDHFY-----FYSG
32 CNEYLNNNPDKDIH-----
33 -----RVAVIPERGYKNRVVTAPPASIL-SMGEVVR--SSVFPYVKAHP
34 SC-----DIVKGDMPPVCFPAVD-----GAKLVSADLTAKATDGF--SHDV
35 IRSVGVGLRNAG----FPQHVVDFVDTLGAGDKPHYAEYDVHTILRRHSKNTVKYDRL
36 CGFLQSYGWDGHSSKVK-VPMLRGSPMGTCSFTLLCIINSWALDHCS-----
37 -----NGPKICGDD-MLAYITAKEFTKYKVRVAAIGSGVHPLKTFIS-
38 --PYA-GTFCENIYIKDPCDSEFP-----
39 -----SVVAFG
40 RAIKC-----PIKAMLYPQKGSNG
41 RIDWP-----
42 -----DEIVHQVFTGETRMSKEKVRLSSF--TDEP---EYHVQRKFEAG
43 PGDFG-Y-NDCPSVFR-LG-----RKVRRVLR-TLRRDQIRSAKK-
44 GRYPWLPA-CLGGLNYP-RRNPKGVMG-LS-GLCRARLFQFTHSDSLSELGKF---MS
45 TLDNAKR--PLPEHRELEFISD-MV----ADAEIND-AIVEDTLEK-----
46 -----QGECSFVDSYSLARHQGSVRSNFF-----
47 -----RSMGGKLDL-----
48 -----SAE-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5 >GAIR01012062.1
6
7
8 -----YTKNMKVRLSSHLLTRLEFKDSTAPS-----PRKSGGCKPLSKIRTKHWHSDKE
9 R-----
10
11 -----TPEQIVW-----EALWA
12 ALVVSDSYNDL-----DVSSMRGSFYLRD-----WVARSSR-----
13 -----RTVGYVVSVLKEELNRLRAISL-----G
14 SDPPIAHGFPKKLDLQLQVFPTL-----GLFAFSGVSRGFP-----TCVDVKVKVSET
15 LVDHMVNATPAP-----A--VPIN---FLRSIEQF---VPRYIQKSVRDRLADKM---
16 -----HDPSPVVCKSSPSACYENPRGAGGFY-----AYIK---
17 --KLGDKISEGQ-----NDLLADIRAGVKPRPENICHPILRGIAKASYDPDRTQDLL-S
18 P-----KEMSSI-LDHFY-----FYSG
19 CNEYLNPNPKDIH-----
20 -----RVAVIPERGYKNRVVTAPPASIL-SMGEVVR--SSVFPYVKAHP
21 SC-----DIVKDGDMPPVVCFCPAVD-----GAKLVSADLTAKATDGF--SHDV
22 IRSVGVLGRNAG-----FPQHVVDFFVDTLQAGDKPHYAEYDVHTILRRHSKNTVKYDRL
23 CGFLQSYGWDGHSSKVK-VPMLRGSPMGTCSFTLLCIINSWALDHCS-----
24 -----NGPKICGDD-MLAYITAKEFTKYKVRVAAIGSGVHPLKTFIS-
25 --PYA-GTFCENIYIKDPCDSEFP-----
26 -----SVVAFG
27 RAIKC-----PIKAMLYPQKGSNG
28 RIDWP-----
29 -----DEIVHQVFTGETRMSKEKVRLSSF--TDEP---EYHVQRKFEAG
30 PGDFG-Y-NDCPSVFR-LG-----RKVRRVLR-TLRRDQIRSALKK-
31 GRYPWLP-A-CLGGLNYP-RRNPKGVMG-LS-GLCRARLFQFTHSDLSVVELGKF---MS
32 TLDNAKR--PLPEHRELEFISD-MV-----ADAEIND-AIVEDTLEK-----
33 -----QGECSFVDSYSLARHQGSVRSNFF-----
34 -----RSMGGKLDL--S---ASSYSIFNAARIRWPAVK-----GEL
35 GTYVKRTPFSLI-----IQLDNLRAPRYRVL-----
36 -----SSFAEKLPGV-EH-----PKSNRWHLHG---GWGGGMMGY-
37
38
39
40
41
42 >GACI01002802.1
43
44 -----VRLSSHLLTRVDFKDLPTPL-----PVKGGGSKPRSTIISRDPRHR--
45
46
47
48 -----RPIQIVW-----EAWWA
49 ALIISDSYDGR-----DVSSLRGSFIFRR-----WVARSSY-----
50 -----RSVGVVVAILKLEMNRLRA-----ISLKS
51 SPVPGVPEGFPKRLDQFLDNISPF-----GLFTLSTMTRGFP-----VTTDKSDKAKA
52 LLEHLENISTPCA---P--VPTP---YLRSIEDF---VPRMFNESKFAKNARDFTK---
53 -----NVQFPAPDTIKLSSSACYEGPRSRRGGYL-----GHIK---
54 --KEGDDDAKDRANGYQEILKDIRGVRPPPEND-THPLLFNIAKRSYDPDNLISDL-S
55
56
57
58
59
60

1
2
3 P-----SRLAEI-MSYYY-----YWNN
4 CDKYMKG--HRDEQVH-----
5 -----RVSTIPERGYKNRIVTAPPSGVL-SVGEVIR--SIIFPFVRDHP
6 SC-----TLVRDGDLRGARFYDRT-----GCRLYSADLTAKATDGF--SHEA
7 IEAVGRGLLSYG----FPTSVVEHFLGTLGVGSK-HYAEYNVADLLQGHKEGSRKYNQV
8 CSYLMEEKGDGN-SKVMRVLMRGSPMGTCSFTLLCILNLWSLDRAC-----
9 -----NGPRICGDD-MLAYITPHDFAHYSRRISIIGSGVHVTKSYVS-
10 --DIA-GTFCEENLYVRSGTQD-----
11 -----GIVSFE
12 RGIRC-----PIKSVLFPQKGSNG
13 RIDWP-----
14 -----DPVVHQVFTGETRMEEPYRQVA---TVVPTEQGYVHTRQRELH
15 QGDFLF--EDLPLTVK-IG-----VRVARVLR-TLRRDTIRACLKK-
16 GRNPWAPA-CLGGLNYP--RKNAKGVKGFP-KWFTRANLYSFTHSDLSPPELSKVLA----
17 QIDCEKK--PDSLSKELEFVST-MV-----ADTEINESLVPEIHDR-----
18 -----VGDCAFVDMYALSRHYGGIKSNFF-----
19 -----RSMGGRLSV-----ASSSYS-----
20 -----
21 -----
22 -----
23 -----
24 -----
25 -----
26 >GAIR01012025.1
27 -----
28 -----SYMKNMKVRLSSHLTRVDFKDLPAPV----PVKGGARKPRSTIKSRDPANR--
29 -----
30 -----
31 -----
32 -----RPIQIVW-----EAVWA
33 ALVVADPSDGIR-----DVSSLRGSFHLKD----WVVRSSH-----
34 ---RSVKYTVEILKAEMTRLRAHSL-----L
35 STEIPVPHGFPRKLDRLRAAFAPM----GLFTLSTMTRGFP----VSTDPKQRSAE
36 LMGHMENVMSPPP---S--VPAH---YLRSIEEF---VPRMFKESSIYKRSDDLRI---
37 -----VNPAPVCNKTSSSSCFEGPRSKGGYF-----GYIR----
38 --TKGDKAGEGK-----DIAKEIREGTTLPPEKTCLPLFRGIAEKSYTMEFPA-P
39 L-----DEFADI-ISHYY----FLND
40 CEDYLNRDPSSQLVH-----
41 -----RIGVVERGYKNRIVTSPPSGVL-SAGEVVR--HVLFPFKSHP
42 SC-----EIVRDGDMVNASFYNRP-----YCSLYSADLTAKATDGF--SHEA
43 ILAVAQGLKTVG----FPTSIDIFLSTLGVSEI-HYAEYSIKDLMAGLRVGSRKYKNR
44 ASELMRGWDGFTDKLR-IAMRRGSPMGTCSFTLLCILNLWALDAAK-----
45 -----NGPKICGDD-MLAYITPSEFNSYSRRIAAIGSGVHPTKTFVS-
46 --RSA-GVFCENIYIRREDSD-----
47 -----GIACFN
48 RGVRI-----PIKSILFPQKGSNG
49 RIDWP-----
50 -----DAMVHQMFGETRLSESYSVVE---TLVPEEQGYKRIRHRELT
51 DGDRF-Y-EDLPHTMR-IG-----VRIRRVL-E-TLRRDAIRACLKR-
52 NRPWAPA-CLGGLN--PRKNPKGVKDFP-AWFRRDLYNFTHSDLGVGELSKV---IA
53
54
55
56
57
58
59
60

1
2
3 AIDCQKK--PSKGSEELGMLST-MV----AENDINE-AIRDDIGED-----
4 -----VGECVSVDVYTLARAYGSARARFF-----
5 -----TSMGGTLDT--K---ASSYSAFCVPKIDWPKGSG-----G
6 GCYDCKTPMELV----IKLDNLLRSPNYRIT-----
7 -----RDFSV-----
8 -----
9 -----
10 >GGCO01105932.1
11 -----
12 -----YMKNMKVRLSSHTRVDFKDLPAVP---PVKGGARKPRSTIKSRDPANR--
13 -----
14 -----
15 -----
16 -----RPIQIVW-----EAWWA
17 ALVVADPSDGIR-----DVSSLRGSFHLKD----WVVRSSH-----
18 ---RSVKFTVEILKAEMTRLRA-----LSLLS
19 TEVP-VPHGFPRKLDRLRAAFAPM-----GLFTLSTMTRGFP----VSTDPKQRSAE
20 LMGHMENVMSPPP---S--VPAH--YLRSLIEEF--VPRMFKESSIYKRS--DDP---
21 -----KIVNPAPVCNKTSSSSCFEGPRSKGYYF-----GYIR---
22 --TKGDKAGEGK-----DIAKEIREGTTLPPEKTCLPLFRGIAEKSYTMEFPA-P
23 L-----DEFADI-ISHYY-----FLND
24 CEDYLNRDSSSQLVH-----
25 -----RIGVVPERGYKNRIVTSPSGVL-SAGEVVR--HVLFPFKSHP
26 SC-----EIVRDGDMVNASFYSRS-----SCSLYSADLTAKATDGF--SHEA
27 ILAVAQQLKTVG----FPTSVIELFLSTLGVSEI-HYAEYSIKDLMAGLRVGSRKYNRR
28 ASELMRGWDGFTDKLR-IAMRRGSPMGTCSFTLLCILNLWALDEAK-----
29 -----NGPKICGDD-MLAYITPSEFDSYSRRIAIGSGVHPTKTFVS-
30 --RSA-GVFCENIYIRREDSR-----
31 -----GIACFD
32 RGVRI-----PIKSILFPQKGSNG
33 RIDWP-----
34 -----DTIVHQMFTEGESRMSESYSVVE---TLVPEEQGYKRIRHRELT
35 DGDRF-Y-EDLPHTKK-IG-----VRIRRVLR-TLRRDAIRACLKR-
36 NRPWAPA-CLGLLNYP-RKNPKGVRD-FP-AWFRRDLYNFTHSDLVGELSKV---IA
37 AIDCQKK--PSKGSEELDMLST-MV----AENDINE-AIRDDIGED-----
38 -----VGECVSVDIYTLARAYGSARARFF-----
39 -----TSMGGTLDT--K---SSSYSAFCVPKIDWPKGSG-----G
40 GCYDCKTPMELV----IKLDNLLRSPKYRIT-----
41 -----RDFSGKIQGAELHK-----SSRWHLRGAG--GWNGDLD-----
42 -----
43 -----
44 -----
45 -----
46 -----
47 >MH213236.1
48 -----
49 -----MTASLRVYRWTLSTDITAWPVTPGV-----GGVGSWISVIPGPATPCVEHM--
50 -----
51 -----
52 -----
53 -----TDRRAIW-----TAVAI
54 TLTAVNPK-----GYRNIIPPKLLAK----WVLR-----
55 ---APVQTIVKGLKSLAHEYRSKAV-----L
56
57
58
59
60

1
2
3 GPLHFVPEVPLVFQKWLGHVCHAT----QYALGRLGRLMP----KPWQKI-TSKA
4 VHEWIRVNTRRPDGPIHE--PSVDNERMSNDIYRF--AHEFATQRKDRLTTEVNFT--
5 -----PSESASTFMTSMSEGGRL-----KEMM---
6 --NDARDLIGEF-----AFDETIVNLAAREGG
7 Y-----KPPPRAQCGMQP-----VLDE
8 L--ILENQIQMADSKEFSP-----
9 -----RVIGVPEYGWKTRVLTLFPNYAL-TPGDLIR--QQLWPLVSEEE
10 WLDAD----RVPSWNLFNNMLSRSLDR-----VGTWVSCDLSNATDYV--PHLY
11 CKALWAGIYDALG--VSEGSWVRRYTEKMFSPMVI-AG-----T-----
12 -----KG-R-I--TFRGTQMGTPLSFMTLCLLHRFAVVRSGF-----
13 -----E-Y--CPHLIRGDD-LIGCFHKP--RIYMGMLEELGFKINKAKTFLS-
14 --TIG-GTFAERTFRAERKLKNV-----
15 -----RHGYFF
16 SHRKV-----
17 -----
18 -----TAGQWAYTRATLL---QDIPLR-GLIKVATTGPN
19 -----RYRSKLLM-LSEIVENVAMG--FRPKRRALIRR LAT-QGRGDILIAALKS-
20 GIPLSLPR-EVGGAGLPTVSEILRDRIYID-LETRRRLGWMISHVPKGQLVAK-----
21 LRRRASGGLIGFTERYLEGIKN-IS-----
22 -----LGLRVPVTEPERTAHYLAAKRRTI-----HL
23 KPWEVV-----HRPPCSFRE--W--GRTFRSGKVGPRWALGG-----
24 -----DRLSLV-----SDRLKVATACIVYVPI-----
25 -----YQGFRGPIRRISD-----GIWTEDPGQNE---GYRHQPSDCS
26 AMS-----
27 >GFJU01140648.1
28 -----
29 -----LAKPMRWFLPKNPLSWPAIPMVE---VGTASDVVKALTR-----
30 -----
31 -----
32 -----
33 -----EDVKTVL-----SFIE
34 PCRIVINSIHRR-----PWLDLRLPYRSFVN---WVCCMVH-----
35 -E-KQEQAIAATAKDLFAVRMWAI-----EGRFP
36 ETWKSRSAAHLPTHFVRWVKSFLCHTG----TRVQFARLGRC LP----VGSRAV-MDTA
37 RANFLAVVHTEPK---D--VSVD--IADELYEF--GRSFI--VSEGLKPVSWE---
38 -----PNESTATLNFHASQGGRM-----AELI---
39 --YDARLPLHLL-----
40 -----EEFTPI-TEPNE---LLDE
41 V--IVETAIIEKSRIEDWEP-----
42 -----ELLVVPERGYKARVLVKFPASAL-LVGDIVR--RQLWPQIEQAE
43 WLDLD----REVNSEKLTYFLRRCIDPNTG-----LGQCVSADLSNATDYI--PHLY
44 AQALWAGILEKVD---APDFVKGYVEKMFSKPMKI-----
45 -----EG-R--M--TKRGIGHMGTPLSFMTLCLMHRYALRKAGL-----
46 -----H-R--HPhiIRGDD-LIAIIPQP--DIYFSAMTSLGFKINKSKTIVS-
47 --KNG-GTFVERVFKFEREDVI-----
48 -----IPHPIL
49 RLPPQ-----SV-----
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 -----ITRALEV---VDIPLA-GMIWPSERGSK
5 L-----RTIGRWYN-QCQSRLTPRMK-----RTVSMLIY-MRHKDILAKATKL-
6 RLPISTPV-ELGGCGLPNRHGRNDFDAP--HWYRAIVGWAACHESMAKKFSG-----
7 RVTRLDG--KDPNPPYYIQNIR-ENQSRSTYEYDPTTTRYRGMKRR-----
8 -----FQIFRWAGALDEDRFKARVEREQS-----
9 -----RFISLPLNQ---W---YPEFKNLKVRPRWVPRR-----
10 ----RADTDRL---IRRIKMFNAEYVLPYT-----
11 -----SRVQFLGSDSY---KTGGDT-----
12 -----
13 -----
14 -----
15 >KX883548.1
16 -----
17 -----MTKSMRRFYLPHEPWRWPVPTPPE---VGRGSVPTLADRDMR-----
18 -----
19 -----
20 -----DFARRVW-----HSVVT
21 PIRMVTSGRLSR-----IAEFRSFAT---WMVRKVA-----
22 -A-GERSVVCRELKNLSFSFRVWAV-----E
23 DRWTFHSDQLPRLFVRAKGLAPCSE---ARAQFARIGRSLP---TGDGKV-VEAA
24 IRSYQRIVVGRPN----PSWD---VADELYAF--GKRWAEERAHLVRGRVNWA---
25 -----PNENTASLNFKASEGGRL-----AELI---
26 -----ADSAAVIEPL-----
27 -----LEFTPE-GEGSE---LLDE
28 V--LIETSIEKARTERP-----IP-----
29 -----KVIAVSEKGYKARVLVEFPASTL-LPGDIIR--RQLWPMLEEEE
30 WIDMD---VVPNEERLKAFLRHSDVH-----SATCISSDLSNATDYI--PHLY
31 AQALWRGVLDAAI-----APEWVGDYTDKMFGPWDL-LF-----P-----
32 -----DG--KTVT--SQRGIHMGTPLSFMTLCLFHKFAVERSGF-----
33 -----S--H--HPLHIRGDD-LIGLFDPD--ARYCRTMEDLGFKINKSKTIIS-
34 -----SKG-GVFVERTFRFTTGREPRRF-----
35 -----ERRSLA
36 SFLPD-----
37 -----
38 -----RMVIRETEL---KDVPVA-GLNPSVEGSL
39 L-----RSVGRWYQQLPDLS-----ARRRKRLHTVIW-SVFGDTLRKARCS-
40 GIPISTPV-QLGGCGIPNRHGRLDIDAS--FKMRRSMALLNKRSVARQFA-----
41 YAVSAIDNLPTHDFDRLFLKHH-EESQRSRSVYTYDP-----
42 -----QSTGYYQSASRRFKMYRLLGGGAE-----
43 -----SVKRRGLRE-----GKLHKPAGFRRWVRLFSNLKMSCPR
44 YAVGRRANSSRL---ETYLKISFGTYIPVAS-----
45 PGYSIFGPGNGEGGVA-----HQAPQLVVALP-----
46 -----
47 -----
48 -----
49 -----
50 -----
51 >KX883542.1
52 -----
53 -----
54 -----MTTSSRYFDLPSNILSWPVTGPGV---VGAGHWHQLPVSALTDRDGV---
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3
4-----RYSNSVW-----KAVVT
5NIVLAHNK-----KFRNVVPAKIFAG----WLLRCSR-----
6-NAGGRYYVAKQLKQLSFEYRQWSV-----T
7GRRPTPKSDVPRTFQKWLAGHVCHAE----ARAQIARIGRCLP----EGDTLV-QSKA
8IKAHRSIVTRPPC---D--VNME---LADDIYYW---ARKFSTDNAHLFKEEVSWA---
9-----PNENSASQGATILEGGRM-----EELI----
10--RDAKAESDFE-----S
11V-----KSLAEGDFDPWV-----SKAL
12DDAVYMTACTVPPKKEWIP-----
13-----DVMAIAEKGYKARVLTKFPAAAL-VVGDVIR--RQMWMAMVSDKP
14WADMD---RQPDDVKFLSTIRSLDR-----HGICVSSDLSNATDYI--PHIY
15AQALWAGLLEPHH---VMPYINSYCAIMFSSMKL-RY-----P-----
16-----DG-TVVT--SARGIHMGTPLSFLTLHRCVEKAGY-----
17-----G-S--YPHIIRGDD-LLGIFPRPEV--YFNVMQQVGFSINRAKTIIS-
18--RTG-GTFAERTVRFSHSLATQE-----
19-----LSNPLK
20RTLHQ-----
21-----
22-----FIPVNISSVKVL---QDLPVG-GVVRATPGKGS
23LV-----KALGRWFS-QTSNIVP-----RQRRKAYRAIG-LNHGDLVIRLSTT-
24-VPPHLPL-DLGGAGLPDRKGRVGLNG-VP-FAIRAAIGHAASHHETAVKLTG-----
25LIARTDG--VSRGFVDVFSKKD-FREQWSKSIWTTEP-----
26-----QETEFYQRTRRYWRYFGHRRDRPSR-----
27-----PISFRQWRS-----GLMTLPRVKARWVPRS-----
28-----NSDPSRL-----ANRIRSMTGTYIPYTQ-----
29-----PTTGSTY-----QWAAKCG-----
30-----
31>GGQW01011558.1
32-----
33-----AEHYE--RFHLPVNPLSWPVIPGV-----VRAGGWYNLPSEALT---GEGL
34-----
35-----
36-----KFKVSVW-----EAFVV
37---VFVASHKK-----TIKNLVPPKIFLR----WVMSCSK-----
38-NETS RDF CCE QLK KLA FIY RR WA V-----EDK--
39--YPTDHCSIPSSFLKWLKG NM SS CT-----QRAQIARIARCLP----KGTTVR-MERA
40IEGHKKIVSI-PA---R--VDID--LADEIYHF--AAKYTAERKEKLI EKV SFH---
41-----PRENTASLEFTEKAGGRL-----ADVM---
42--ANA-----SDWLRRR-P
43I-----ALSEHL-YTNKS-----YLDE
44A--IVEAAV---KQDLAYP-----YY-----
45-----RA---EVVAVAEPGNKVRVVCKFRSVPL-LISDIIR--RQLFPIFEDE
46NMDFD---QDIRSEKFGK MIRR CLHN-----EGVIVSSDLSNATDYI--PHHY
47AQALWHGILEQFD---APEWVLKHVLMLSPVEF-NY-----P-----
48-----DG--SVIM---SQRGIQMGTPLSFMTLSLLHKFAVERSRN-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----G-E--SSYLIRGDD-LIGVFKYP--RSYYSEmenLGLKINANKTLQS-
4 --HRG-GVFAEQTVKVSWKKSTRTS-----
5 YHRTLY
6 DYFPA-----AE-----
7 -----
8 -----TDEVSRIWVL---DDIPMK-GLLHVDPGVYG
9 SKL----RMIGRWYS-QWSSYYPICSR-----KRAYRAIG-IAHKELIKTARFL-
10 GLPSSTPL-ELGGCGLPNRKGALALDAS--FLHRGQIGYAASH-SDSRFS-----
11 HFVGKLD--GIVKDDGLLEQSI-ER-----SRRSLGSSVFETDPF-----
12 -----NTEFYLKARRARAYLSPDYLKE-----
13 -----RERCTPLSQ---W---KKSFRKLPTVKARWVPKR-----
14 --DKDSFGIDRL---IDLKIFNGRYVPFTP-----
15 -----
16 -----
17 -----
18 -----
19 >GBDZ01000351.1
20 -----S
21 SPDNTGAMTSQCETTYGLPISVLSWPVNPGV-----VGAASWIPNPRTALTGVERGRAL
22 -----
23 -----
24 -----
25 -----AYAQGVW-----YALTS
26 NIEAVHPR-----PFKSLIPAKILAG---WVVRCSA-----
27 -TAAGRTQCSGLLKRAFLYRLWAT-----T
28 DTKPGLVDGIPKCPQLWIRSLCSHAS----VRAQIARLGRMLP----QGDAVT-VRKA
29 LKKHRSVVTQSTR-----ANVF---VADEIYAF---AKEYMENREEHLREKVSFH---
30 -----PTEGTASRDYTIERGGRV-----QUELL-----
31 --DNAWAMMRKP-----
32 -----GYIDYQKYPDKS-----RIDE
33 F-IVQAAIDADKRLHHYRV-----
34 -----SVMAAAEPGNKARVLCKFPAVAL-VPGDIIR--RQLWPIVASDP
35 DLDFD----QDPRSEKFLQMIRRTVDR-----EGTIVSSDLSNATDYI--PHMY
36 AQALWRGLLESFD---APSWVQEHIINRMFSPMEM-TY-----P-----
37 -----DG--KKVW---SQRGIQMGTPLSFMTLSLLHKFAVQASGH-----
38 -----G-L--SAHIIRGDD-LLGVFTSP--ADYFRSMERVGFKINREKTIVS-
39 --RVG-GVFAEQTVRVRYTPRRQP-----
40 -----KKITLA
41 DFMGP-----PR-----
42 -----
43 -----EDVESVTVL---NDIPLK-GAVHVDTKGSV
44 L-----RGLGRWYA-QWSEQYPAKCG-----RTAHRAVM-RVHQALIRDAKRW-
45 RVPLHTPL-ELGGAGIPDRRGRGLVAD-LS-FDMRRKLGYACSH-FDRSYA-----
46 RAVRRLD--GGSTDPEEHEFEQ-ML-----KKAP-----
47 -----LGSCVYELYDPYATEYYAACRRKYY-----LTKV
48 TRPARK-----DLNPPPLRD---W---LKQFDNCPAVQPRWAWSK-----
49 -----HADVTLL---IHRLKMFSGAYVPFTT-----
50 -----PETSILARRKV--R-----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 >GBFE01004865.1
4 -----
5 ---GGAMTSQCETTYGLPISVLSWPVNPGV---VGAASWIPNPRTALTGVERGRAL
6 -----
7 -----
8 -----
9 -----AYAQGVW-----YALTS
10 NIEAVHPR-----PFKSLIPAKILAG----WVVRCSA----
11 -TAAGRQCSGLLKRAFLYRLWAT-----T
12 DTKPGLVDGIPKCPQLWIRSLCSHAS---VRAQIARLGRMLP---QGDAVT-VRKA
13 LKKHRSVVTQSTR-----ANVF---VADEIYAF---AKEYMENREEHLREKVSFH---
14 -----PTEGTASRDYTIERGGRV-----QELL----
15 --DNAWAMMRKP-----
16 -----GYIDYQKYPDKS-----RIDE
17 F--IVQAAIDADKRLHHYRV-----
18 -----GVMAAAEPGNKARVLCKFPAVAL-VPGDIIR--RQLWPIVASDP
19 DLDFD-----QDPRSEKFLQMIRRTVDR-----EGTIVSSDLSNATDYI--PHMY
20 AQALWRGLLESFD---APSWVQEHIRMFSPMEM-TY-----P-----
21 -----DG--KKVW--SQRGIQMGTPLSFMTLSLLHKFAVQASGH-----
22 -----G-L--SAHIIRGDD-LLGVFTSP--ADYFRSMERVGFKINREKTIVS-
23 --RVG-GVFAEQTVRVRYTPRRQP-----
24 -----KKITLA
25 DFMGP-----PR-----
26 -----EDVESVTVL---NDIPLK-GAVHVDTKGSV
27 L-----RGLGRWYA-QWSEQYPAKCG-----RTAHRAVM-RVHQALIRDAKRW-
28 RVPLHTPL-ELGGAGIPDRRGRLGVAD-LS-FDMRRKLGYACSH-FDRSYA-----
29 RAVRRLD--GGSTDPFEEHFEQ-ML-----KKAP-----
30 -----LGSCVYELYDPYATEYYAACRRKYY-----LTKV
31 TRPARK-----DLNPPPLRD---W---LKQFDNCPAVQPRWAWSK-----
32 -----HADVTLL---IHRLKMFSGAYVPFTT-----
33 -----PETSILARRKV--R-----
34 -----
35 >KX883517.1
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 -----
47 -MAVVIEAIHRR-----PFTNLVPAKMFAG----WLIRASR----
48 -GPGGQKYVCQQIKSLAFDYRVWSI-----T
49 AKKPNYSGDIPRAFRNWSSLCCHAE---TRAQFARVARMLP---VGNKLV-IDQA
50 LRKHRSVVTRRPK-----LDIG--LAHQIYGY--ARDYIHERKHSLSERVSFH---
51 -----PTEGTASRDVPASAGGRL-----RDLV----
52 --DNAHRVMSQR-----P
53 I-----VLSEYL-YSHKK----YFDD
54 A--IVEAAIYTDEKVSTYNV-----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----DVLAYPEPGFKARVLCKFPATAL-LAGDIIR--RQLWPIFESDP
4 NMDFD----QDIRSEKFQNVIRRSLNR-----DGTIVSSDLSNATDYI--PHEY
5 AKALWAGILDAFD---FPEWVENYLERMFAPIRM-SY-----P-----
6 -----DG--VTVT--SCRGIQMGTPLSFLTLSLLHKFCVHKSGH-----
7 -----E-R--SPYIIRGDD-LLGVFSSP--RQYLTVMEEIGFKINRDKTVIS-
8 --KDG-GTFAEQTVKVTWKAKERDPL-----
9 -----ARPTLY
10 DFIVT-----
11 -----
12 -----DQKVVSITNL---DDIPFK-GLIHLNDKGGR
13 L---RQVGRWYA-QWSPYYPPRKG-----KVAYRAIR-RTIGNVLRARIARSL-
14 RIPITCPM-ELGGCGIPNKRGTMQLDAN--FQHRSRVGYAASHESHNFQ-----
15 -----LAVRKLDIGNASD-LLEDYRQHVADLPKGKSVMFDPY-----
16 -----SSEDFSGYRRKRMLLNNNTGNYSAY-----
17 -----HTKPVPLHR----WLA VFSRCKENRARWVPSR-----
18 --RCSGSDLRRL---IDRIKMLSPIYVPYTT-----
19 -----PNAQILSIRI-----
20 -----
21 -----
22 -----
23 -----
24 >GFLP01304447.1
25 -----
26 -----FLALGLSSGLYLTSLD---SKHFSTLAFELDNRKRS-----
27 -----
28 -----
29 -----RTFHDTI-----GDFRP
30 RKLFAKLITYQM-----QNPEILNEEKLRA-----
31 -----VMPEVSISTNRSRS-----
32 -----SKFPVTLGSLP----HPSETV-VEET
33 TRKYLEVVTG-PKSSIDS--PPLD---LLTSIEEF--GRKYVVDRLLHRVKD--KVS---
34 -----WVPNLTNTASLHVTAREGGRM-----TELL-----
35 -----LGADTVLKLL-----
36 -----LEYTDP-SEPNA-----DLEE
37 A-IVQT SIEIAKLKALDG-----NF-----
38 -----EA---EVMAVGEAGYKARILCKYSAHIL-LVGDIIR--RQLWPLVSEEP
39 WFDDEKDYSFSEKLERAFHDLPFWTGYH-----EPEVVSTDLSATDFI--PFEV
40 ARALWSGILSGLG--VGFDSWIFQHVITWAGPTKV-KVGD-----
41 -----SVHT---TNRGIPMGTPLSFMTLCLLHHYAVKSTEK-----
42 -----ETGM-E--GPYVIRGDD-FIGVLPNKEV--YFRKMEELGFVVNHSKTFIS-
43 --KVP-GIPVRPSLRS DTSNEDTVRSSACGVFAETVVQFNF-----
44 -----REHPVT
45 SGKRK-----PF-----
46 -----
47 -----TWQPIRLSDVTVI---PDVPFA-GLVREEKSSIK
48 RSQL---VRYGEWYRQIFRRPDGTLLDRGRWVKASRVVFTAIR-RLSKDLIMRAVES-
49 GVPIHTPL-EIGGAGIPDARGRLTPRD---FVTRSKITYLL-----
50 -----NHTGNSRVR SARA-TL-----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 RSAAAEELGG KIWSS
4 FDKV
5 -----
6 -----
7 >NC_033474.1
8 -----
9
10 ---MTVTSVPVVAQVRIPTLRLAAAVLQRSP---VLRKGSLRESLLRLRSLSLKAEE
11 RLDPIERKF-----
12 -----
13 REELLRW VDAAL
14 YMSLPMYNYRTIS GSHQVEQYLLALW KLLHYVY
15 -K-NSIPDVVKTLKDLFFRYLDAGLQS DREFRRRRF
16 TYLG-KRHPLKGFIAVKGLTRE SARTIGLAGRCIP S-LPLSKERRLEE
17 YLNLLKERLTT-PM PTIP ACRGCGHL PPEHPDKRLAAGPGCQRLSRDH
18 IVESEVIKFAEVFARPVKHLPMSYPADNMVYIPGPLHHGETR PWLM
19 --GVTKPAVEAISDAFDSFDEWKEAVNSPKLPYQLGIWSVHPEAIENPPERDPAVGLTRV
20 ISEGESLKTPKRRANWYDRKDLRVWYSERHAPSVLHPKDPEAERRDDHR ILPP
21 GYGLRALKRSPAPILN DV VLSGEDQCLVP-----
22 -----VH---KVLFPIEAGAKVRVATVPSVVS LQTAARQIN--GIVLRALKHCP
23 TF---HLGLKNSKCFSSKTFAMPD---GAKFYSVDLSAASDRLSQPLSL
24 S-VLWPFIKREA EK--AGLEN-AFEVARLSCGPQRI-ILEDPEVMELLADPKY-----
25 -----SAHLPFFE--TEAGLLMGTPLAWPALSLLHLFAALEAGI-----
26 -----D-I--HRVFIVGDDGG LIATAPEYQKYYHIMTNILGLRVNREKTHES-
27 --PMM-GVVAQKVLKIRPAPCPH-----
28 -----AGTPRG
29 KLGIG PL
30 -TVWD-----
31 -----FVRHELNLPNESPVYWGKVELDQSHCYVP--AALTARSRRRR
32 DPHAP--WNPPEWSD-VA-----RWKRWRERLPKALR-LVYPSLVSQSRYR-
33 GLRMEAPA-ILGGCGPLTEGAVRKGLTPVQ-RAGAYAIAREIMHGTRKDLSIDVGPDLTR
34 TIPRAVI--LAANEVFNELPRS-ET----RGTGLPKEAALRSLTSF-----
35 -----FVRVAWE GGFRDANSY LGRLSGH R--VVRHPGRLLEAKGRIF
36 LSRGK-----RAVPKPLTV-----VQAFKRLPKPVDPVSPGVLKGLNRLIHD
37 LDLGGRQRKRLDYVLSAGR NDDAEGTSSSVQ-----V
38 PEAMSSLNLRLRGWG-----KWWGQIQSLGS--GGKG NLTSTH
39 HEA-----
40 >GBTA01004534.1
41 -----
42 -----LVPEIKEFTL-----
43 -----
44 -----
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----KIVSKMY QDWKS
51 KTEKFNEQCEET GRKKINLDLSVPE W-----
52 -----SPNLHASACIE-----
53 --NPIKNGGQQAYMH-----KLFQVAHTG-----VGEARN-ARAF
54 AMRIRNVLQQEMI-----WLKSMGDW-----
55 -----
56
57
58
59
60

1
2
3 -----QVGNELPSTTDRKEGGVE-----
4 --EIPKDDFRFN-----
5 -----TSMPEG-IPD-----YVSG
6 S--YLPVG-----
7 -----RASVVYEKGKARVVTVHAAFV-ALSCALN--RTLINLLKCYT
8 PS-----KASLLGRDIELKTLKRD-----GLEGFSADLSTATDLI--PFEV
9 ARAVVDGICEALN---WGPHIKEIMYYATGPFRI-EYDRAPRSKDS---I-----
10 -----VP-DYII---SRRGILMGLTSWPILSILNSFCAEYNCS-----
11 -----KEG-V--RNYRICGDD-LAALWDKNDIGNYIGNLESVGLKLNDKKTVTS-
12 -----DQG-LVFTEELYSIKSRTHINR-----
13 -----
14 -----
15 -----
16 -----
17 -----
18 -----KGERSPDQKDL YLSRFKRAIVSVIVDAKHNSTTKFT
19 NPE-----ERLPRYLT-IG-DCSHAIDE--FCPLWQKSRAQKALL-LMNKKCIREIFRA-
20 GLDPFLPK-YLGGAGVKTVGNEKRL-----HPLMKHLSALIQLPYEDIQLVCR-----
21 RLQRPWL--SSSVNKETSAIFS-KTLNLFEKLPRAEH-----
22 -----GSAIPYKEAVRKFTGAVLARY-----
23 -----SLLNFGSCQ-----TNKVCTLKFIRKQVNKAR-----
24 KTLFSKAPRRYH---YMSLLKFSNALEKLEP-----
25 -----
26 -----
27 -----
28 >GEWH01003494.1
29 -----
30 -----
31 -----TLNSSVAEKSETGK-----RPRSRMEKLLRRCTERDMTRQL
32 LETLGLRL-----
33 -----
34 -----DSLQLG-----QDLVQ
35 AFSTLLDVERGN-----KYFERRWFIRLAQ-----WLLRAIDSFTDK
36 VKAEGISSAMTWIKDALTHLRACAL-----
37 DAGKVTDYPFPSKAPRVERRR-----MLVQASRCRSRLR----AATGDK-LEQA
38 AQSHKEILTS-PF---K--TPEP---ELLSFREY---VRAKTRASTVQVGE-----
39 -----VSASASLDTKGSGGGSL-----EEVR-----
40 -----GLVDSLRSRPFTYPDLLELDRTFPHEMLGGHVLLNRPKVQERFRIGQVGHSSSGTLLG
41 VPGDQVLFLY-----DQSCEISLDEWE-----FLRE
42 QLFAIVACVLDVDLDSRPAC-----
43 -----RQCVIAERGWKTRVTPISAASN-YLLTVAN--KGLLSIIGEHR
44 EL-----DNSQRGRPCDNLDWSFGKR-----YALTFSADLKSATDYM--PHDL
45 CLVAAEELTRG-----WPAYLRRLVLSVGSYDL-RC-----L-----
46 -----DG-GTVE---TRRGILMGVPTSWPILSLYSGWLHYRSSS-----
47 -----D--GWFAVCGDD-YLGCHTYASYHKYLRARERTGAKGSPGKDILS-
48 -----TESTGIFAEELVT-----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----VGRCRVL---RTVPIR-PLTGIPK-----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----GDTPSW-A-MG-EAISVLLS---EAGFSEERSGRLVS-RRFGSEMDKLRRRA-
4 GIDPKGPR-WALAGGFPGKPG-PRVLETAR-RVLSQGEDTIL---GWLT-----
5 ALGIAWS--DAPSRDLIQGMVR-DK-----
6 -----LGGL-----
7 -----
8 -----KGQTWRAE-----
9 -----
10 -----
11 -----
12 >GDUK01007882.1
13 -----
14 -----KVEKWLPFSVFGGSSMASRSR---PLHVGRIIDLRLVEMQSQHFPGAS
15 MTAD-----
16 -----
17 -----
18 -----QIAQQFW-----NVRS
19 HLLAAFPPPLVSGLPVAKRS-----GRNARFGLVLLQY-----HLLRRRH-----
20 -L-GGVASLVRWLKALAFSCRDSVV-----
21 -----SDKKLPPREFSRYFHGVLNPSAL---SLTTLAKLGRALPECDAKMGGGSV-IRKA
22 LRQHRADMTSARVKQFGFNPMDR---DLAVAKET---ASFLLRGRLGS prm-----
23 -----HDLSSASADVSSREGGLA-----AGLRPSWA
24 RMEEIGPVPWALGHALAE EWQAALLRGARNTAWATAEEERGIVLSNGSFLPGPLVLLA-
25 S
26 P-----DGLWVT-RRSEYIRS VLSFHEG
27 LGPVRTGRTTVQVDHLTQVF WHVAGEPD LKRFFASCARD-----LP-----
28 -----KA---KVALVRERGWKV R VVTSLETDQV-IAGHALR--DLIWPVLQDDP
29 CF-----DLTGDPRIEGMPV DQ-----GEVVL SADLSRATDV V-SIGL
30 ALAVFDQWNRIDP---RFDRQWHDLARATVSGYRI-QY-----P-----
31 -----DG--RYVR--SSAGWLMGPSTW ALLCLVHK AIAEMAGF-----
32 -----KKYRIRGDD-LLAVGSRESVEEYFTLISKY-FEVNVKKSFVS-
33 -----PNG-GTFAEDTFVVRNGQLRLIA-----
34 -----
35 -----
36 -----
37 -----
38 -----
39 -----
40 -----HAAPPR-GWLAGEKASLA
41 AAYTQ-F-GRLPRKWR-----MAYRHSLL-SSGKHQIAEFRKH-
42 RIPLFLPR-VLGGVGVPHPRGVGGAVA-----GSVLVSKMVTGAG-----
43 SVSTAWV--NRMAYRQREALYS-EAREVTRMLHQDGSLGSSTVLTEE-----
44 -----VINEITLKSNIAMGCTWP-----EGWDKT
45 LLGELR-----DPSYASIGR-----DWALFRACKCAGIRPPP HK-----
46 --VGTGWSEQRL-----ADAIRKSSLGGMYWPY-----
47 -----RPDPIRLS-----
48 -----
49 -----
50 >GEUE01057748.1
51 -----
52 -----QRMKQPAKPGKFYLPNLVESEVPKVTL DLILGGKGGA VIRDLPTLSRV RP PARPMPS
53 AKGKSAD EDNCNVLGNDNNNGSKRAKTRR DESWERSD-----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----KDANVFW-----GAWL
4 LLLDSISSAYERGLDLKNPS-----ERRSAREALSVRG-----WIRKSVH-----
5 -F-NGVDFVQKEFKLAAHYRHRAV-----GD
6 VRDPPAPGLLDRLSGIPRSVLKGPIEGLRACAFIGRLGRCPG-----PPSKKV-EERA
7 LKEHRRVLCRRFK-----TPKV---QLASIRKS---TKAWCKDLNEVGPSVVAFS---
8 -----SAASVEFSRLKGQT-----AELF---
9 --DHAKGAIQMY-----QTLLYGEPE
10 I-----EGSNSTGLRDED-----IRIT
11 VQVALDSAIEKAKSRSPRLV-----
12 -----KATVVPETGGKARVVTAGPADMV-LAGNALR--RAVWPIVMGDD
13 SIDTEEAQAQEEVSAGDKELLARLLGG-----GCSWYSGDLTAASDWL--PRDV
14 CLTIWDAIAEAAG--FAKSPEYLLGRKLLGNVVI-EYA-----
15 -----DG--TEVT--TGAGALMGFPLTWLVLCAYNRALVSMVSK-----
16 -----KA-I-QRTVFRGDD-MVSQMTPEEGARYEELVRLTGGQPNTAKSFRS-
17 --VTG-WVFTEATYQLTTWHTTLTRADLAE-----
18 -----RGVTLE
19 HGLGR-----
20 -----
21 -----AGTLRAITVRRV---PDCSLR-HLLRQRVASQP
22 YLA----IAGPAVSGAVAPLLASRLAD---VRAVGNRLVRRWMR--VNARMVESIQRH-
23 NLGLFVPR-ELGGFGLPHPKGWARGVTSA--RWAHKRLAVFFTATRGAAVKIGC-----
24 LGYDPVVDPNTGRLWGEVSKR-LASREKSCVLPPERAEPADR-----TPESW
25 EQYVPNYGGYYLAVVPGAGRGAQPVTEAVAQEMALESAWR---
26 RGLGLDPVVDPKQEK
27 RKSGTSDGCKSKAKKPSRSALTIISK-----NHRAMRKTLGKSRFPEKF-----
28 LPKRALRPEKLQ---MALMRHKMYVQHGSLFGG-----
29 ---PSPLHIVFNSEKGRGPPRRP-----GGWKYYHDPKQ--GRSKTSSDPV
30 PRSPAPPLVGGVPLDPAWENSWPSRKSGNTS-----
31 >GEUE01035234.1
32 -----
33 -----
34 -----
35 -----
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----
51 -----KATVVPETGGKARVVTAGPADMV-LAGNALR--RAVWPIVMGDD
52 SIDTEEAQAQEEVSAGDKELLARLLGG-----GCSWYSGDLTAASDWL--PRDV
53 CLTIWDAIAEAAG--FAKSPEYLLGRKLLGNVVI-EYA-----
54 -----DG--TEVT--TGAGALMGFPLTWLVLCAYNRALVSMVSK-----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----KA-I--QRTVFRGDD-MVSQMTPPEEGARYEELVRLTGGQPNTAKSFRS-
4 --VTG-WVFTEATYQLTTWHTTLTRADLAE-----
5 -----RGVTLE
6 HGLGR-----
7 -----
8 -----AGTLRAITVRRV---PDCSLR-HLLRQRVASQP
9 YLA---IAGPAVSGAVAPLLASRLAD---VRAVGNRLVRRWMR--VNARMVESIQRH-
10 NLGLFVPR-ELGGFGLPHPKGWARGVTSA--RWAHKRLAVFFTATRGAAVKIGC-----
11 LGYDPVVDPNTGRLWGEVSKR-LASREKSCLVLLPPERVAEPTADR-----TPESW
12 EQYVPNYGGYYLAVVPGAGRGAQPVTEAVAQEMALESAWR---
13 RGLGLDPVVDPKQEK
14 RKGTSRGCKSKAKKPSRSALTIISK-----NHRAMRKTLGKSRFPEKF-----
15 LPKRALRPEKLQ---MALMRHKMYVQHGSLFGG-----
16 ---PSPLHIVFNSEKGRGPPRRP-----GGWKYYHDPKQ---GRSKTSSDPV
17 PRSPAPPL-----
18 >GBYB01012090.1
19 ASARLGRSGKWKLPLNLEPVATAVTPAAAARLLSNLRKDAFPRGVGTNYLSRLEAAGI
20 P
21 ACSVSRHPRGSEPHPLRKYTPGRGDPACSPSLPHTPLPTSPSRPGRARGGPAGERPRK
22 RGHSRDPPGGCPLGCGAACADC-----
23 -----
24 -----RDLLRIW-----DSFRL
25 CILSLIPRLQRSGLDLSNRR-----DYRMAVSILKVR-----WLSINWR-----
26 -S-SGRNETIRQIKAYAEFCRARAL-----DAEQS
27 RRGAPRGFPVRAFEHGLGSLLRGASSGRT--ALAQLARLGRSMP----PADQAI-LTAA
28 LEQHRRVLTADP---G--VPLT---LLDGLFH-----ASSWSGSRPASLRNSVVAS---
29 -----FSSSASEGMSRAKGGQR-----AELE---
30 --QLARAEIDAI-----RADIAGHPDTDPGFFDVDEGD
31 Y-----AGLYFV-EWDLQ-----TVTD
32 S-AMARVREEAADTRYAIPV-----
33 -----RATALPELGKARVVTAPPAHWG-IIGDAMR--KVLWPLLETDP
34 RIDLSGRRSLDGAASFHDQVVKSLRGAA-----GQWMYSADLTAATDLM--PENV
35 ILALWHGVLHGLG--IPEESFFARAGDKILGCVDV-SYPDLAA-----
36 -----PG--EKPIVVRSMRGCMGMLNSWFLLNLYNLAIVDIACLGGIAALTDPQD
37 GGLDEAEVRRIVG-L--APAIVRGDD-LAAALTERQATAYEELIAATGGEANRAKSYRS-
38 --HSA-FVLAEKSFVLVDREVRPLSGGQKRGFISFPAGLA-----
39 -----VDAPML
40 ARDPR-----ALEALGFGGPLTG-
41 -----
42 -----KDLGGATWTDAAVEVVSCLTAL---QDIPVR-HLIPAPSK---
43 -----EGLPVYVS-LPAAAADVLYAAEEEGSPLFPGMCRAVL-SVNSEYVQRYRDY-
44 RIPLLLPR-EVGGAGFPHPGGFSKALASGG-RGHWLRATLRVTTYGVEQRARRH-----
45 LDEDVWR--VDGSNSDRRAAAR-LIRAREAAAMNSGQ-----
46 -----LGNMVPVPLEDEVTFREVAFESLWR-----DLFCPADRKER
47 CTRRDR-----GVRLATISK-----RLSKADMAARRLAFNDRF-----
48 --LVRSLVRAEY-----QADVAMARGAESVYVPLSERDPRLMPLSDDHPSLVGLHRRR
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 YRAGRASSHGRGNGCGGSQVGGVP-----DNWESRLDDDFVYGGGLDDNEAGIS
4 GHCLPKRSTLERAVRSRGQGA VAVRTLRSGMVLADSIHYTESRSKRRGKPPTS-----
5 >KF298275.1
6 -----
7 -
8 GVMIRSRRKRNHQRWRLPSTLVAPEGRITPRAHSQPNEAGGAQGAAQSNVAQVTAER
9 A
10 ANKPTQANSQRTRHARAPKDPKE-----
11 -----
12 ----- QMGQATW-----DAFRL
13 ----- AIWALIPSLQKKGLALKGPR-----NLATVRGIQSVHK-----WISVSMA-----
14 ----- R-SGPEWTAKQVKEYNYCRARSLR-----DARTP
15 ----- CP---RGFPVKAFEVGLIKAAISDPNGGRP-ALAQLARLGRAMP----VATDRV-GRQA
16 ----- LEKHRETLSRATD-----VPPA---LIARFGAW---AEQWARQQADKTGL--TKE---
17 ----- GSRLFQAVETKFSNSAAEAFSAASGGQQ-----AELR---
18 ----- ENYQRTMVDV-----QAEYLLATEGFVWDPEV--
19 ----- PTLGAE-DSDLM-----NITL
20 ----- C-AREEAE---DYYLRRH-----RWRGVD-----GM-----
21 ----- PV---RATIVREQGMKARVVTACPAWAV-VCGDACR--KTLWPLLAADP
22 ----- RVDLSGARPVAEDLDRFNDQIALSLVGSI-----TPEFFSADLTAATDLM--PFEV
23 ----- SNALWAGLCRGLD--LLPTSQISRIGHALLGPVEV-SYPDLAA-----
24 ----- KG--EPIPRILSKQGCMCMMGLPLSWTILNLYNLAADVADFAVV-----
25 ----- APSVVRVG-V--APVIARGDD-LVAAYLPEEADRYTQLLRESGGEVNVLKSFRS-
26 ----- SDS-FVLAERTFKVTLRDAVPLPVHRRGFELRARN-----
27 ----- THAPLL
28 ----- TQVPQ-----PSGTSRSDPNVLG
29 ----- GPA-----
30 ----- AESILGKTRETRGKVVVSIKMF---DDCPLR-TLVKGPGYAA
31 ----- G---AVIPAYIS-VPSAASASLAE--FEGKRFYPALAQGLL-SVHKSLVGEFRRS-
32 ----- AIPVFYPR-ELGGGGFPHPRGFHAVCSAG-LLGRKRAAFAMTCFTHKARKKHG-----
33 ----- LGDDPW--SRHRLAELERARQ-QL-SAEEARAQTRAASEVPGAPRV-----IP
34 ----- DDWEPKVWDQIKDLEPPRGARIPVPMEDA VVQRAAHGETWA-----SMILQEG
35 ----- VGQRDR-----YPSLGTIAK-----RLQAIRKVTAERKFSNRF-----
36 ----- LPKRGKLSRERL-----LERVAMARGGETRYVP-----
37 ----- --YVDFGGGGMVLVCRGFDTYAAARYEPESQPERYPERRQKLRP----GTEGEEMVEV
38 ----- PVAVRTTRAGLGLADVAFVRHGRVKRVKRRLPSQFPR-----
39 ----- >KP642119.1
40 -----
41 ----- -MVLSGTCAQPNPSGVASAVKAVASPSVGASA---KLRKGPVVGTD RIPKSSLSSPRG
42 ----- HVGRKPRETVKPKPDPLQ-----
43 -----
44 ----- GQAQRAW-----DVWRL
45 ----- ALFALMPSLQKKGLDVRKAKPAKKPVTA--QHATQRAIQVLYK----WLASSFV-----
46 ----- -R-SGPEWTCKQIKEYANYCRSRSL-----GDE
47 ----- RTDAPRGFPVKGF EAGFIKACIMDPNAGGT-ALAQLARVGRAMP----MGTT RV-AVAA
48 ----- LRKHRDTLTSTTV-----VPEE---TQVNLR LW---AERWTRDRLAAGARAREGT---
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
 2
 3 -----ATSFsRAASASVSAADGGQL-----AELR----
 4 --QLPRFVDHRQ-----MLRDLLEEQGFDLDSDDIF
 5 L-----SATCDE-GDDLQ-----LLAD
 6 C-AIETARDRCMEGSPIPM-----
 7 -----VATTVLELGMKARVVTKPPAWAV-VAGDACR--QSVWPLLESDR
 8 RIDLSGARPNTESLDRFHDLNLAHSLVGAA-----TPQFFSADLTAATDDA--PLRL
 9 -----LGRYLLGPVTV-GYPTLAAIEEHDRLYQ-----
 10 -----AGEPTEIV---SVRGCMGLPLSWTLLNLYNLAVALDMACS-----
 11 -----PVGVVQVG-V--APAIARGDD-LVAAIPPGEADRYEELIALTGGEANRLKSFRS-
 12 -----ATA-FVLAERTFRVETQEIPDRVRLPGRWRWRVSRCFGDP-----
 13 -----ALPPLL
 14 TREHQ-----PG-----
 15 -----
 16 -----LRGQPSRNLEGAREVVGVRMS---VDLPLR-ALIAGSASFAG
 17 G-----AAVPTYIS-LPPAATSClSE--FEGTRFYPAVARGLL-SVHRDLVSEMRRS-
 18 AVPLFYPR-ELGGGGFPHPGGFAGLASAG-LLGLQRAGLALTSHGWAARKKAA-----
 19 LLEDPWV--PRRQNAQLNRARE-QLLASERRAWAQSVLKRGETVASL--EAAGLAERAK
 20 ELKLALATPVPEYERVPRGARIAVPLEDETIRQAARAGLWE-----DAFLSPTVD
 21 DGGRSKTK-----YPSLGDIAR-----RLKIREVTSKCKFDPRF-----
 22 VPKDRVQRDSF-----FEKLRMLRGSETVLVP-----VDSRPVRVTVWD
 23 PARHRADPELDNDDCGY-----PGRRRLLRALD---REEGGDPELV
 24 PDPVPLLQRGAGATLGDLVVFRPAKSRSRGTGARREKRSRRTPQHP-----
 25 >KP642120.1-----MLA
 26 TGVMACPQRPKVPNRWFLPKPSVGASAALKR-----GPVVGTD RIPKSSLSSPRGHVGR
 27 KPRKTVKPKPDPLQ-----
 28 -----
 29 -----GQAQRAW-----DVWRL
 30 ALFALMPSLQKKGLDARKAKLAKKPVTA--QRATQRQAIQVLYK-----WLASSFV----
 31 -R-SGPEWTCKQIKEYANYCRSRSL-----GDE
 32 RTDAPRGFPVKGF EAGFIKACIMDPNAGGT-ALAQLARVGRAMP----MGTRV-AVAA
 33 LRKHRTDTLTSTTV-----VPEG---TQVNRLW---AERWTRDRLAAGARAREGT---
 34 -----TTSFsRAASASVSAADGGQL-----AELR----
 35 --QLPRFVDHRQ-----MLRDLLEEQGFDLDGDDI-F
 36 L-----SATCDE-GDDLQ-----LLAD
 37 C-AIETARDRCMEGSPIPM-----
 38 -----VATTVLELGMKARVVTKPPAWAV-VAGDACR--QSVWPLLESDR
 39 RIDLSGARPNTESLDRFHDLNLAHSLVGAA-----TPQFFSADLTAATDLM--PFSV
 40 SWSLWDGLCDGLG--AAADAPLRLG RYLLGPVTV-GYPTLAA MEEHDRLYQ-----
 41 -----AGEPTEIV---SVRGCMGLPLSWTLLNLYNLAVALDMACS-----
 42 -----PLGVVQVGV---APAIARGDD-LVAAIPPGEADRYEELIALTGGEANRLKSFRS-
 43 -----ATA-FVLAERTFRVETQEIPDRVRLPGRWRWRVSRCFGDP-----
 44 -----ALPPLL
 45 TREHQ-----PG-----
 46 -----
 47 -----LKGQPSRNLEGAREVVGIRMS---VDLPLR-ALIAGSASFAG

1
 2
 3 G-----AAVPTYIS-LPPAATSCLSE--FEGTRFYPAVARGLL-SVHRNLVSEMRRS-
 4 AVPLFYPR-ELGGGGFPHPGGFGAGLASAG-LLGLQRAGLALTSHGWAARKKAA----
 5 LLEDPWM--PRRQNAQLNRARE-QLLASERRAWAQSVLKRGETVASL--EAAGLAERAK
 6 ELKLALATPVPEYERVPRGARVAVPLEDETIRQAARAGLWE-----DAFLSPAVD
 7 DGGRSKTK-----YPSLГDIAR-----RLKRIREVTSKCKFDPRF-----
 8 VPKDRAQRDSF-----FEKLRMLRGSETVLVP-----VDSRPVRVTWWD
 9 PARHRADPELDNDDCGY-----PGRRRLRALD---REEGGDPELV
 10 PDPVPLLQRGAGATLGDLVVFRPAKSRHGPKSRAGTGARREKRSRRTPQHP-----
 11 >KF298284.1
 12
 13 -----G
 14 VMAKPQRLFRSQQRRWYLPAVLSLEERSGVTM----PLGTFPVKTGKVAKDGPCHPVPQ
 15 GSAKRANRHPNSPPTRRPKPERRARTQRTPTQPGVELQ-----
 16
 17 -----
 18 -----RQAQHIW-----DTFRL
 19 TIWVLLPSAQKKGIAVS-----AKATIRSIKGVYD----WIRNSMI----
 20 -R-SGVEWTSKAIKEYANYCRARAL-----GDD
 21 RVGRPRGFVVRTFERGYIAAAVKCPGQGRL-HLAQIARIGRAMP----KGTRNV-GISS
 22 LRKHGVLSQ-PM---E-TDKD---LLENLRSW---ATQWAQERVSERGTVVNPS---
 23 -----VSFSRSASATVSAMSGGQL-----AELS---
 24 --QLESVKAHHR-----LIRELLLEEQQFEWNPAGTSDDSIG
 25 L-----LTVADE-VADLQ----ILAD
 26 C-AISDALDAAREQIPIPM-----
 27 -----TATVISELGMKARVVTKPPAWAV-VAGNACR--KTVWPLLEADR
 28 RIDLSGLRPTAEVLDEFHDNLAHSLVGAR-----TTQFYSADLTAATDLM--PFDV
 29 SRALWNGLCDGLG--ASPSAPLRLSGSYLLGPVRV-SYPDLAALKERDRLFK-----
 30 -----AGEETEIL---SSRGCMGLPVSWTVLNLYNLAMADMACT-----
 31 ---PAGVPTLLG-V--APA VARGDD-LVAAIPEFAADRYENLIAATGGEANRLKSFRS-
 32 --TFA-FVLAERTFEVGVKSVKDLTALTRRGFKVRQVSAP-----
 33 -----RTAPLL
 34 CRFDE-----
 35
 36 -----
 37 -----LGPQQIPAFIQEMDRPVEVVALRMS---CDLPIR-SLLGGMAGFAG
 38 G-----EAVPNYVA-IPSAAAACLAE--FEGTRLYGAICHGLL-SVHRDLVKEMRRS-
 39 AIPLFYPR-VLGGGGFPHPRGFAAAVASAG-PLGLQRAGLYLTASGYKARRKVG----
 40 IGHDPWV--PLADSRRMDVARD-RLVATDRRLWAKSVLDRLPDLLLD--EDPYLDKRLS
 41 ERAEALSVPCVGTFETVPRGARVAVPLDDAVISEAATIAWE-----DAVLGPC
 42 PSAERR-----YPTLGGIAK-----RLKAIRQKTVEARFPMKF-----
 43 VPKDRAQRDSF-----FEKLRMLRGSETVLVP-----VDSRPVRVTWWD
 44 TVQEPEDEGRVEGDLGL-----PIRRRELRERH---GGGVRVLRDG
 45 ATLGDVAVIRHVPQKRRRRRTPGPSRNVL-----
 46 >KF298276.1
 47
 48 -----G
 49 VMAKPQRLFRSQQRRWYLPAVLSLEGRSGVTM----PLGTLPAKAGRVAKVGPCHPDPQ
 50 GSAKRANRHPNSPSTRRPNPERRARAQRTPTQPGVELQ-----
 51
 52 -----
 53 -----RQAQHIW-----DTFRL
 54
 55
 56
 57
 58
 59
 60

1
2
3 TIWVLLPSAQKKGIAVS-----AKATIRSIKGVYE----WIRNSMI----
4 -R-SGVEWTSKAIKEYANYCRARAL-----GDD
5 RVGRPRGFPTGLFERGYIAAALKCPGQGRL-HLAQLARVGRAMP----KGTRNV-GISS
6 LRKH RGVL SQ-PM---E-TDKD---LLENLRSW---ATQWARERVSERGTVVNPS---
7 -----VSFSRSASATVSAMSGGQL-----AELS---
8 --QLESVKAHHR-----LIRELLLEEQGFEWNPAGTSDDSIG
9 L-----LTVADE-VADLQ-----ILAD
10 C--AISDALDAAREQIPIPM-----
11 -----TATVISELGMKARVVTKPPAWAV-VAGNACR--KTVWPLLEADR
12 RIDLSGLRPTAEVLDEFHDNLAHSLVGAR-----TTQFYSADLTAATDLM--PFDV
13 SRALWNGLCDGLG--ASPSAPLRSLSGYLLGPVRV-SYPDLAALKERDRLFK-----
14 -----AGEETEIL--SSRGCMGMLPVSVTVLNLYNLAMADM ACT-----
15 ---PMGVPTLLG-V--APA VARGDD-LVA AIPEFAADRYENLIAATGGEANRLKSFRS-
16 --TFA-FVLAERTFEVGVKSVKD LTALTRRGFKVRQVSRAP-----
17 -----RTAPLL
18 CRFDE-----
19 -----
20 -----LGPQQIPAFIQEMDRPVEVVALRMS---CDLPIR-SLLGGMAGFAG
21 G-----EAVPNYVA-IPSAAAACLAE--FEGTRLYGAICHGLL-SVH RDLV KEMRRS-
22 A I PLFY PR-VLG GGGFPHPRGFAAAVASAG-PLGLQRAGLYLTASGYKARRKVG-----
23 IGHDPWV--PLADSRRMDVARD-RLVATDRRLWAKSVLDRLPDLLL--EDPYLDKRLS
24 ERAEALSVPVCVG FETVPRGARVAVPLDDAVISEAATIAAWE-----DAV LGPC
25 PSAERR-----YPTLGGIAE-----RLKAIRQKTVEARFPMKF-----
26 VPNMKRMGREGF-----FNRV KLLLASE TVLVP-----YSNRNTARVTVFDPERHC
27 TVQEPEDEGRVEGDLGL-----PIRRREL RERH--GGGVRVLRDG
28 AT LGDV AVIRHVPQKRRR HRRTPGPSRNVL-----
29 >MF176344.1
30 -----MSL
31 PERVGGTI P RTPA GSREPASA QV GRPQP SPK GL--ARKTNHTQIPPTIVDPKSSPRRA
32 RPPKKACAPGKATTKPPASAQQ-----
33 -----
34 -----KQAQHIW-----DAFRL
35 SLWVLLPRLQKHGLSTK-----AVATIEAIKALYS-----WIGNSVA-----
36 -R-SGVEFTAKQVKEYANYCRARAL-----KAQ
37 QCEPPRGPPIRAFERGYIRAMLACPGQGRL-HLAQLARVGRAMP----IGTAKV-AIAS
38 LRKH REVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRERLGAGGRTLTNP---
39 -----AAHFSRSASATVSALKGGQL-----TEL R-----
40 --QLPAVAEHYA-----LIQELLGDQLDPVTGEPFVWDPYTFADGGAI-G
41 L-----LTSASE-ADDLQ-----ILAD
42 C-ALRTATEHAADNTPIPM-----
43 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
44 RIDLSGVRPTAEVLDTFHDNLAHSLVGAR-----TTQFYSADLTAATDLM--PFDV
45 SRAMWDGLCDGLG--ATATAPLRTLGRYLLGPVQV-SYPDLRALPASSKLYV-----
46 -----AG--ERVE-CLSKRGCMGMLPVSVTVLNLYNLAMADLACT-----
47 ---PEGSPV LVN-V--APAIARGDD-LVA AI PAGE ATRY E D LIAATGGEANRLKSFRS-
48 --ADA-FVLAERTFEVGVLRRPNV-ELRQRGYVVR--R-TY-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----RTAPLL
4 AQFDS-----AE-----
5 -----
6 -----LHGGTGDARRLGGAPEVVAIRMA---CDVPIR-SLLGGGPRTAG
7 A-----NPVPDYVS-IPPAACLAE--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
8 AVPLFYPR-ELGGAGFPHPKGFAAVASAG-ELGLKRALTTAGHKAAQQKVG-----
9 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSLARHPDLPRG---GSAWTDRAAG
10 ELRKALSIIPARGFETVPRGARIAVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPPM
11 Q
12 EPSRKRSARRKEESSGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
13 LPKRAVQERARF-----FDRKLALRELETVLVP-----AQSAGARLVVFDPVTHR
14 VSESAGEIDGGEGRMGF-----PLWRHTLQERK--GKRGAAQGQEA
15 AVPVPVATTRSGSLGDVMAVRRL-----PRRPRGVRLPNPGGAHSRNVR
16 >MF176385.1
17 -----MSL
18 PERVGGTIPRTPAGSHEPASAQVGRPQSPSKGL---ARKTNHTQIPPTIVDPKSSPRRV
19 HPPKKVCAPGKATTKPPASAPQ-----
20 -----
21 -----KQAQHIW-----DAFRL
22 SLWVLLPRLQKHGLSTK-----AVTIEAIKALYS----WIGNSVA----
23 -R-SGVEFTAKQVKEYANYCRARAL-----KAQ
24 QCEPPRGFPIRAFERGYIHAMLACPGQGRL-HLAQLARVGRAMP---IGTTKV-AIAS
25 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW--AEIWTRERLGAGGRTLTNP---
26 -----AAHFSRSASATVSALKGGQL-----TELRL---
27 --QLPAVAEHYA-----LIQELLGDQLDPVTGEFPVWDPYTFADGGAI-G
28 L-----LTSASE-ADDLQ--ILAD
29 C-ALRTATEHAADDTPIPM-----
30 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
31 RIDLSGVRPTAEVLDTFHDNLAHSLVGAR-----TTQFYSADLTAATDLM--PFDV
32 SRAIWWDGLCDGLG--ATATAPLRTLGRYLLGPVQV-SYPDLRALPASSKLYA-----
33 -----AG--ERVE-CLSKRGCMGLPVSVTLNLYNLAMADLACT-----
34 ---PEGSPVVLN-V--APAIARGDD-LVAAIPAGEATTRYEDLIAATGGEANRLKSFRS-
35 --ADA-FVLAERTFEVGVLRRPNV-ELKQRGYVVR--R-TY-----
36 -----RTAPLL
37 AQFDS-----AE-----
38 -----
39 -----LHGGTGDARRLGGAPEVVAIRMA---CDVPIR-SLLGGGPRTAG
40 A-----NPVPDYVS-IPPAACLAE--FEGTRLYRSVAEGLL-SVHRGLVADLRRS-
41 AVPLFYPR-ELGGAGFPHPKGFAAVASAG-ELGLKRALTTAGHKAAQQKVG-----
42 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSLARHPDLPRG---GSAWTDRAAG
43 ELRKALSIIPARGFETVPRGARVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPPM
44 MQ
45 EPSRKRSARRKEESSGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
46 LPKRAVQERARF-----FDRKLALRELETVLVP-----AQSAGARLVVFDPVTHR
47 VSESAGEIDGGEGRMGF-----PLWRHTLQERK--GKRGAAQGQEA
48 AVPVPVATTRSGSLGDVMTVRRL-----PRRPRGVRLPNPGGAHSRN--
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 >MF176278.1
4

5 -----MSL
6 PERVGRTIPRTPAGSHEPASAQVGRQPSPKGL---ARKTNHTQIPPTIVDPKSSPRRA
7 RPPKKVCAPGKATTKPPASAQQ-----
8

9 -----KQAQHIW-----DAFRL
10 SLWVLLPRLQKHGLSTK-----AVATIEAIKALYS----WIGNSVA----
11 -R-SGVEFTAKQVKEYANYCRARAL-----KAQ
12 QCEPPRGFPIRAFERGYIHAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
13 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRELRGAGGRTLTNP---
14 -----AAHFSRSASATVSALKGGQL-----TELRL---
15 --QLPAVAEHYA-----LIQELLDQLDPVTGEPFVWDPYTFADGGAI-G
16 L-----LTSASE-ADDLQ-----ILAD
17 C-ALRTATEHAADNTPIM-----
18 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
19 RIDLSGVRPTAEVLDTFHDLNAHSLVGAR-----TTQFYSADLTAATDLM--PFDV
20 SRAMWDGLCDGLG--ATATAPLRTLGRYLLGPVQV-SYPDLRALRASSKLYV-----
21 -----AG--ERVE-CLSKRGCMGLPISWTVLNLYNLAMADLACT-----
22 ---PEGSPVLVN-V--APAIARGDD-LVAAIPAGEATRYEDLIAATGGEANRLKSFRS-
23 --ADA-FVLAERTFEVGVLRRPNV-ELRQRGYVVR--R-TY-----
24 -----RTAPLL
25 AQFDS-----AE-----
26

27 -----LHGGTGDARRLGGAPEVVAIRMA---CDVPIR-SLLGGGPRTAG
28 A-----NPVPDYVS-IPPAAAAACLAE--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
29 AVPLFYPR-ELGGAGFPHPKGFAAAVASAG-ELGLKRAGLALTAGHKAAQQKVG-----
30 LLSCPWS--PKLTDERRAKEARR-RLIASDQRAWANSVLARHPDLPRG---GSAWTDRAAG
31 ELRKALSIPARGFETVPRGARVAVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPP
32 MQ
33 EPSRKRSARRKEESSGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
34 LPKRAVQERARF-----FDRLKALRELETVLVP-----AQSAGARLVVFDPVTHR
35 VSESAGEIDGGEGRMGF-----PLWRHTLQERK--GKRGAAQGQEA
36 AVPVPVATTRSGLSLGDVMTVRRL-----PRRPRGVRLPNPGGAHSRNVR
37 >MF176306.1
38 -----MSL
39 PERVGRTIPRTPAGSHEPASAQVGRQPSPKGL---ARKTNHTQIPPTIVDPKSSPRRA
40 HPPKKVCAPGKATTKPPASAQQ-----
41

42 -----KQAQHIW-----DAFRL
43 SLWVLLPRLQKHGLSTK-----AVATIEAIKALYS----WIGNSVA----
44 -R-SGVEFTAKQVKEYANYCRARAL-----KAQ
45 QCEPPRGFPIRAFERGYIHAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
46 LRKHREVLSQ-PM---D-TEPT---LVEKLRAW---AEIWTRELRGAGGRTLTNP---
47 -----AAHFSRSASATVSALKGGQL-----TELRL---
48 --QLPAVAEHYA-----LIQELLDQLDPVTGEPFVWDPYTFADGGAI-G
49 L-----LTSASE-ADDLQ-----ILAD
50
51
52
53
54
55
56
57
58
59
60

1
2
3 C--ALRTATEHAADNTPIPM-----
4 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
5 RIDLSGVRPTAEVLDTFHDNLASHLVGAR-----TTQFYSADLTAATDLM--PFDV
6 SRAMWDGLCDGLG--ATATAPLRTLGRYLLGPVQV-SYPDLRALPASSKLYV-----
7 -----AG--ERVE-CLSKRGCMGLPISWTVLNLYNLAMADLACT-----
8 ---PEGSPVLVN-V--APAIARGDD-LVAAI PAGEA TRYEDLIAATGGEANRLKSFRS-
9 --ADA-FVLAERTFEVGVLRRPNV-ELRQRGYVVR---R-TY-----
10 -----RTAPLL
11 AQFDS-----AE-----
12 -----
13 -----
14 -----LHGGTGDARRLGGAPEVVAIRMA---CDVPIR-SLLGGGPR TAG
15 A----NPVPDYVS-IPPAAAAACLA E--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
16 AVPLFYPR-ELGGAGFPHPKGFAAAVASAG-ELGLKRAGLALT TAGHKAQQKVG-----
17 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSVLARHPDLP RG---GSAWTDRAAG
18 ELRKALSIPARGFETVPRGARVAVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPP
19 MQ
20 EPSRKRSARRKEESSGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
21 LPKRAVQERARF-----FDRKLALRELETVLVP-----AQSAGARL VVFDPVTHR
22 VSESAGEIDGGEGRMGF-----PLWRHTLQERK--GKRGQAQGQEA
23 AVPVPVATTRSGLSLGDVMTVRRL-----PRRPRGVRLPNPSGAHSRNVR
24 >MF176257.1
25 -----MSL
26 PERVGGTI PPTPAGSHEPASAQVGRPQSPKG L---ARKTNHTQI PPTIVDPKSSPR RV
27 HPPKKVCAPGKATTKPPASARQ-----
28 -----
29 -----KQAQHIW-----DAFRL
30 SLWVLLPRLQKHGLSTK-----AVATIEAIKALYS----WIGNSVA-----
31 -R-SGVEFTAKQVKEYANYCRARAL-----KAQ
32 QCEPPRGPPIRAFERGYIHAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
33 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRERLGAGGRTLTNP---
34 -----AAHFSRSASATVSALKGGQL-----TEL R-----
35 --QLPAVAEHYA-----LIQELLGDQLDPVTGEPFVWDPYTFADGGAI-G
36 L-----LTSASE-ADDLQ----ILAD
37 C--ALRTATEHAADNTPIPM-----
38 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
39 RIDLSGVRPTAEVLDTFHDNLASHLVGAR-----TTQFYSADLTAATDLM--PFDV
40 SRAMWDGLCDGLG--ATATAPLRTLGRYLLGPVQV-SYPDLRALPASSKLYV-----
41 -----AG--ERVE-CLSKRGCMGLPVSWTVLNLYNLAMADLACT-----
42 ---PEGSPVLVN-V--APAIARGDD-LVAAI PAGEA TRYEDLIAATGGEANRLKSFRS-
43 --ADA-FVLAERTFEVGVLRRPNV-ELRQRGYVVR---R-TY-----
44 -----RTAPLL
45 AQFDS-----AE-----
46 -----
47 -----
48 -----LHGGTGDARRLGGAPEVVAIRMA---CDVPIR-SLLGGGPR TAG
49 A----NPVPDYVS-IPPAAAAACLA E--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
50 AVPLFYPR-ELGGAGFPHPKGFAAAVASAG-ELGLKRAGLALT TAGHKAQQKVG-----
51
52
53
54
55
56
57
58
59
60

1
2
3 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSVLARHPDLPRG---GSAWTDRAAG
4 ELRKALSI~~PARGFETVPRGARVA~~VALEDAIREVAAAAAWEDAFLGPGARGWSGVPPP
5 MQ
6

7 EPSRKRSARRKEESSGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
8 LPKRAVQERARF-----FDRLKALRELETVLVP-----AQSAGARLVVFDPVTHR
9 VSESAGEIDGGEGRMGF-----PLWRHTLQERK--GKRGAAQGQEA
10 AVPVPVATTRSGSLGDVMTVRRL-----PRRPRGVRLPNPGGAHSRNVR
11 >MF176365.1
12 -----MSL
13

14 PERVSGTIPRTPAGSHEPASAQVGRQPSPKG---ARKTNHTQIPPTIVDPKSSPRRV
15 HPPKKVCAPGKATTKPPASAQQ-----
16 -----

17 -----KQAQHIW-----DAFRL
18 SLWVLLPRLQKHGLSTK-----AVATIEAIKALYS----WIGNSVA-----
19 -R-SGVEFTAKQVKEYANYCRARAL-----KAQ
20 QCEPPRGFPIRAFERGYIHAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
21 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRERLGAGGRTLTNP---
22 -----AAHFSRSASATVSALKGGQL-----TELRL---
23 --QLPAVAEHYA-----LIQELLGQLDPVTGEPFVWDPYTFADGGAI-G
24 L-----LTSASE-ADDLQ----ILAD
25 C-ALRTATEHAADNTPIM-----
26 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
27 RIDLSGVRPTAEVLDTFHDLNAHSLVGAR-----TTQFYSADLTAATDLM--PFDV
28 SRAMWDGLCDGLG--ATATAPLRTLGRYLLGPVQV-SYPDLRALPASSKLYV-----
29 -----AG--ERVE-CLSKRGCMGLPVSWTVLNLYNLAMADLACT-----
30 ---PEGSPVLVN-V--APAIARGDD-LVAAIPAGEATRYEDLIAATGGEANRLKSFRS-
31 ---ADA-FVLAERTFEVGVLRRPNV-ELRQRGYVVR--R-TY-----
32 -----RTAPLL
33 AQFDS-----AE-----
34 -----
35 -----
36 -----
37 -----LHGGTGDAARRLGGAPEVVAIRMA---CDVPIR-SLLGGGPRTAG
38 A-----NPVPDYVS-IPPAAAACLAE--FEGTRLYRSVAEGLM-SVHRLVADLRRS-
39 AVPLFYPR-ELGGAGFPHPKGFAAVASAG-ELGLKRAGLALTAGHKAAQQKVG-----
40 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSVLARHPDLPRG---GSAWTDRAAG
41 ELRRALSI~~PARGFETVPRGARVA~~VALEDAIREVAAAAAWEDAFLGPGARGWSGVPPP
42 MQ
43

44 EPSRKRSARRREESSGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
45 LPKRAVQERARF-----FDRLKALRELETVLVP-----AQSAGARLVVFDPVTHR
46 VSESAGEIDGGEGRMGF-----PLWRHTLQERK--GKRGAAQGQEA
47 AVPVPVATTRSGSLGDVMTVRRL-----PRRPRGVRLPNPSGAHSRNVR
48 >KX883461.1
49 -----MSL
50

51 PEGVSGAIPRTPAGSHEPASAQVGRQPQGPPEGL---ARKTNHTQIPPTIVDPKSSPRRA
52 HPPKKARAPGKATTKPPASVKQ-----
53 -----

54 -----KQAQHIW-----DAFRL
55
56
57
58
59
60

1
2
3 SLWVLLPRLQKHGLSTK-----AVATIEAIKALYS----WIGNSVA----
4 -R-SGVEFTAKQVKEYANYCRARAL-----RAQ
5 QCEPPRGFPIRAFERGYILAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
6 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRERLGAGGRTLTNP---
7 -----AAHFSRSASATVSALKGGQL-----TELRL---
8 --QLPAVAEHYA-----LIQELLDQLDPVTGEFPFWDPYTFADGGAI-G
9 L-----LTSASE-ADDLQ----ILAD
10 C--ALRTATEHAADNTPIPM-----
11 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
12 RIDLSGVRPTAEVLDTFHDLAHSVLGAR-----STQFYSADLTAATDLM--PFDV
13 SRAMWNGLCDGLG--ATATAPLRTLGLYLLGPVQV-SYPDLRALPASSKLYV-----
14 -----AG--ERVE-CLSKRGCMGLPVSVTLNLYNLAMADLACT-----
15 ---PEGSPVLVN-V--APAIARGDD-LVAAIPAGEATTRYEDLIAATGGEANRLKSFRS-
16 --ADA-FVLAERTFEVGVLRRSNVELKQRGYVTRRTF-----
17 -----RTAPLL
18 AQFDS-----
19 -----
20 -----AELHGGTGDARRLGDAPEVVAIRMA---CDVPIR-SLLGGGPRTVG
21 A-----NPVPDYVS-IPPAAAAACLAE--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
22 APILFYPR-ELGGAGFPHPKGFAAAVASAG-ELGLKRAGLALTGKAQQKVG-----
23 LLSCPWS--PKLTDERRAREARR-RLIASDQRRAWANSVLARHPDLPNG---GSAWTDRAAG
24 ELRKALSVPAREFETVPRGARVAVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPP
25 TQ
26 EPSRKRSARQKEESLGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
27 LPKRAVQERARF-----FDRLKALRELETVLVP-----AQAAAGARLVVFDPVTHR
28 VPESAGEIDGGEGRMGF-----PLWRHTLQERK---GKRGAAQGQEA
29 AVPVPVATTRSGSLGDVMTVRRLP RPRGVRRLPNRSGAHSRNVR-----
30 >KX883537.1
31 -----MSL
32 PEGVSGAIPPRTPAGSHEPASAQVGRPQPSPKGL---ARKTNHTQIPPTIVDPKSSPRRA
33 HPPKKARAPGKATTKPPASVQQ-----
34 -----
35 -----KQAQHIW-----DAFRL
36 SLWVLLPRLQKRLSTK-----AVATIEAIKALYS----WIGNSVA----
37 -R-SGVEFTAKQVKEYANYCRARAL-----RAQ
38 QCEPPRGFPIRAFERGYILAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
39 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRERLGAGGRTLTNP---
40 -----AAHFSRSASATVSALKGGQL-----TELRL---
41 --QLPAVAEHYA-----LIQELLDQLDPVTGEFPFWDPYTFADGGAI-G
42 L-----LTSASE-ADDLQ----ILAD
43 C--ALRTATEHAADNTPIPM-----
44 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
45 RIDLSGVRPTAEVLDAFHDLAHSVLGAR-----STQFYSADLTAATDLM--PFDV
46 SRAMWNGLCDGLG--ATATAPLRKLGLYLLGPVQV-SYPDL SALPASSKLYV-----
47 -----AG--ERVE-CLSERGCMGLPISVTNLNLAMADLACT-----
48 ---PEGSPVLVN-V--APAIARGDD-LVAAIPAEATTRYEDLIAATGGEANRLKSFRS-
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 --ADA-FVLAERTFEVGVLRRPNVELKQRGYVTRRTY-----
4 -----RTAPLL
5 AQFDS-----
6 -----
7 -----AELHGGTGDARRLGDAPEVVAIRMA---CDVPIR-SLLGGGPRTVG
8 A----NPVPDYVS-IPPAACLAE--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
9 AIPLFYPR-ELGGAGFPHPKGFAAAVASAG-ELGLKRAGLALTTAGHKAAQQKVG-----
10 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSVLARHPDLPRG---GSAWTNRAAG
11 ELRKALSVPSRGFETVPRGARVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPP
12 MQ
13 EPSRKRSARQKEESLGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
14 LPKRAVQERARF-----FDRLKALRELETVLVP-----AQAAAGARLUVFDPVTHR
15 VSESAGEVDGGEGRMGF-----PLWRYTLQERK---GKRGAAQGQEA
16 AVPVPVATTRSGSLGDVMTVRRLP RPRGVRRLPNRSGAHSRNVR-----
17 >KX883538.1
18 -----MSL
19 PEGVSGAIPRTPAGSHEPASAQVGRPQSPKGL---ARKTNHTQIPPTIVDPKSSPRRA
20 HPPKKARAPGKATTKPPASVQQ-----
21 -----
22 -----KQAQHIW-----DAFRL
23 SLWVLLPRLQKRLSTK-----AVATIEAIKALYG----WIGNSVA-----
24 -C-SGVEFTAKQVKEYANYCRARAL-----RAQ
25 QCEPPRGFPIRAFERGYIHAMLACPGQGRL-HLAQLARVGRAMP----IGTTKV-AIAS
26 LRKHREVLSQ-PM---D-TEPA---LVEKLRAW---AEIWTRERLGAGGRTLTNP---
27 -----AAHFSRSASATVSALKGGQL-----TEL R-----
28 --QLPAVAEHYA-----LIQELLGDQLDPVTGEPFEWDPYTFADGGAI-G
29 L-----LTSASE-ADDLQ-----ILAD
30 C--ALRTATEHAADNTPIM-----
31 -----TATVISELGMKARVVTKPPAWAV-VAGDACR--KTVWPLLEGDR
32 RIDLSGVRPTAEVLDCAFHDNLASHLVGAR-----STQFYSADLTAATDLM--PFDV
33 SRAMWNGLCDGLG--ATATAPLRKLGLYLLGPVQV-SYPDLSALPASSKLYV-----
34 -----AG--ERVE-CLSERGCMMGLPVSWTVLNLYNLAMADLACT-----
35 ---PEGSPVLVN-V--APAIARGDD-LVAAIIPAEETRYEDLIAATGGEANRLKSFRS-
36 --ADA-FVLAERTFEVGVLRRPNVELKQRGYVTRRTY-----
37 -----RTAPLL
38 AQFDS-----
39 -----
40 -----AELHGGTGDTRRLGDAPEVVAIRMA---CDVPIR-SLLGGGPRTVG
41 A----NPVPDYVS-IPPAACLAE--FEGTRLYRSVAEGLM-SVHRGLVADLRRS-
42 AIPLFYPR-ELGGAGFPHPKGFAAAVASAG-ELGLKRAGLALTTAGHKAAQQKVG-----
43 LLSCPWS--PKLTDRRAKEARR-RLIASDQRAWANSVLARHPDLPRG---GSAWTNRAAG
44 ELRKALSVPSRGFETVPRGARVALEDAVIREVAAAAAWEDAFLGPGARGWSGVPPP
45 MQ
46 EPSRKRSARQKEESLGMYPSLGEVAR-----RLKTIRSQTQAARFNRTF-----
47 LPKRAVQERARF-----FDRLKALRELETVLVP-----AQAAAGARLUVFDPVTHR
48 VSESAGEIDGGEGRMGF-----PLWRHTLQERK---GKRGAAQGQEA
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 AVPVPVATTRSGLSLGDVMTVRRLPRPRGVRRLPNRSGAHSRNVR-----
4 >KX883539.1
5 -----
6 -----
7 -----
8 -----
9 -----
10 -----
11 -----
12 -----MGREETVKRVKQYAEYCRARSI-----DARQA
13 EAGPPRGFPFHALEGGSFKALLRCPVQGPA-ALAQVARLGRSMP----LPTHKQ-QGRA
14 LIEHCNELTAGPH--EE--IPVE--LLDAFYYW--TRTWGAGKGAALRDEVTVN---
15 -----FSSAAAESVSXRARGGQR-----EELR---
16 --LLAAPHFEAI-----REALAIDDIDFRVEEGD
17 Y-----AGIYSE-EDDLE-----IITQ
18 A-AIERVRADIRAGRRPDV-----
19 -----RAAVVPELGAKSRIVTAGEAHWV-VIGDAIR--KCLWPALAEAEP
20 RIDLSGRRDVAGTA AHLHDEVAKALTGAE-----AVEFYSADLTAATDLM--PHDL
21 IQAIWRGLMEGLG--VSDDDLMFTAGLALLGPVNL-HHPDLDLDR-----
22 -----PGKYAYGV--THKGCMGFPLSWYILNLYNLACADLALS-----AER
23 GRFLAEVAPRLVG-A--APAVVVRGDD-LCSAHTPAEADRYERII RATGGGRANLSKSYRS-
24 --RKG-FILAERTFLVTTARTKCSAARSRLGTVVVRPGVH-----
25 -----PRAPLL
26 TRAPD-----AFESGALG-----
27 -----
28 -----YRSIDPRSRRRGAATVVGLTML---SDVPVR-HLMPSPDS---
29 -----EGLPAYVT-LPPAASAVVDE--IKDIKRREAICYGIL-SVNAPLVRQFNRF-
30 KIPIFYPR-ELGGAGFPHPKGFGSAIRSAG-GMGHLRATLAMTTFTRRARQAYQ-----
31 LDRDPWI---CDRQVEFERSR-AQTGLIAEERRAAESGQ-----
32 -----LGLATSVPLEDAVAARAAEQLWA-----EVMLPQSR
33 ELASARRGRR-----GPRLSSVRK----GLSDVFNALAKVHAKFDPKF-----
34 -LANPRIDKETW----MDRVAMAKGSSTVYRV-----VRPSGRRVVMHGIRD
35 HREIRAEREMLKGGWT-----FGGQRDLRQNTAAGLEGGEIAVPV
36 PTQRAGATLWDVLSVSCKRARKRKKGNPRL-----
37 >GENC01006608.1
38 -----
39 -----RVS RQGSDKAWFLPDNVRWTQIVGNLR---PPGESAGTAEMVESLPASQKSKV
40 GRKKEKFPGRATK-----
41 -----
42 -----AAGQRIW-----NAFRL
43 AFWAIAPSAQRKGLNPRSQR-----QARSIQAFRTLEK----WIAISLL----
44 -K-CGTEWTCKAVKDYSEYCRARAL-----RAK
45 QCVRPREFPVTGFEIGLVRGAINDPLAGAE-ALAQLARLGRAMP----QGTDKQ-VIRA
46 LEKH RD TLLGETR-----VDEE--TLDRYHAW--AERWSSDRRRFAPERVAVK---
47 -----FSDSAAESVSRADGGQK-----ADLL---
48 --ENARPHFEAL-----RAALAESQGWDAPD-D
49 Y-----IGLYSE-EDDVE-----TIAQ
50
51
52
53
54
55
56
57
58
59
60

1
2
3 M--ALETVRAAVRSGEPVPA-----
4 -----RASAYPEYGSKVRVVTACPGPFV-VAGDACR--KIVWGLLEADP
5 RVDLSGRRPLGDSISQFQDEVARSLVGAV-----TTEYYASDLTAATDLM--PFQL
6 SAAIWTGLCDGLG--ESQDSDLRSIGLHLLGPVQV-RYPDLK----EGSVE-----
7 -----EG--FRVPTIISQQGCMMGLPLSWTVLNVFNLATADLAIG-----
8 --APEDRVTDLIG-R--APA VARGDD-LAAAFLPEEADRYEALISATGGRVNLSKSFRS-
9 --SIG-FVLAERTFMVKRSVDRYQRRAPRQAVLIQPHRPEGIP-----
10 -----QNAPIL
11 AQDPS-----
12 -----
13 -----APTMSGVGRDVRGHRTESLRML---PDCSLK-GLLGGAST--
14 -----DGAPRYVT-LPVSAASSLYE--FEGTRFYGALCRGLL-SVHKDLLGTREF-
15 GVPKFYPR-ELGGGGFPHPRGYL RAMESAG-GIGHLRATLNITRFTESSRRAAA-----
16 LHMDPWR--TERDSRTLERARS-QLIAEESRAREAGL-----
17 -----LGSFVSKFQEEAEVTRAQGSLWE-----TTLLSQ
18 KKRDEK-----FPRLKDV RH-----RMLEITEKTKAIKFNR RF-----
19 -LLRRQLHRDEY-----LQIASHARGAPTRYVRMERAPGLVALRDIVPGQQAPNSFQTG
20 TVPQHGNARGNRGERAEMSLVPTLRSRASLGQVARILPARASRKRG---ARRGPNPRPS
21 DSSRGE-----
22 >GEND01011317.1
23 -----
24 -----DLSRQGSDKAWFLPDNRWTQIVGNLR----PPGESAGTAEMVESLPASQSKV
25 GRKKEKFPGRATK-----
26 -----
27 -----AAGQRIW-----NAFRL
28 AFWAIAPS AQRKG LNP RSQR-----QAR SIQAFRTLEK-----WIAISLL-----
29 -K-CGTEWTCKAVKD YSEYCRARALR-----AKQCV
30 --RP-REFPVTGFEI GLVXGAINDPLAGAE-ALAQLARLGRAMP----QGTDKQ-VIRA
31 LEKH RD TLLG-ET----R--VDEE---TLDRYHAW---AERWSSDR RRF APE--RVA---
32 -----V-----KFSDSAAEVS RADGGQK-----ADLL-----
33 --ENARPHFEAL-----RAALA-ESQGWDAPD-D
34 Y-----IGLYSE-EDDVE-----TIAQ
35 M--ALETVR----AAVRSGE-----PV-----
36 -----PA----RASAYPEYGSKVRVVTACPGPFV-VAGDACR--KVVWGLLEADP
37 RVDLSGRRPLGDSISQFQDEVARSLVGAV-----TTEYYASDLTAATDLM--PFQL
38 SAAIWTGLCDGLG--ESQDSDLRSIGLHLLGPVQV-RYPDLK----EGSVE-----
39 -----EG--FRVPTIISQQGCMMGLPLSWTVLNVFNLATADLAIG-----
40 --APEDRVTDLIG-R--APA VARGDD-LAAAFLPEEADRYEALISATGGRVNLSKSFRS-
41 --SIG-FVLAERTFMVKRSVDRYQRRAXRQAVLIQPHRPEGIP-----
42 -----QNAPIL
43 AQDPS-----
44 -----
45 -----APTMSGVGRDVRGHRTESLRML---PDCSLK-GLLGGAST--
46 -----DGAPRYVT-LPVSAASSLYE--FEGTRFYGALCRGLL-SVHKDLLGTREF-
47 GVPKFYPR-ELGGGGFPHPRGYL RAMESAG-GIGHLRATLNITRFTESSRRAAA-----
48 LHMDPWR--TERDSRTLERARS-QLIAEESRAREAG-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 >KX883556.1

10 -----MQEVNISTLTRQGVSNDRP
11 LSATEKVDADCMKILHESSGTDPTAARGAPH----PSTEVEGEESRQRGRSKAAPAPA
12 REGSGADHRSKNPA-----
13

14 -----ADNVSVWGPGRALLSDSLLS
15 SVLSLGPSSTRLITGLRGVTLAIFTAEQDEVLARSLAYRTCR----HLARRAL----
16 -H-EGVEAVLSVMGRWFSSARE-----SFLNG
17 AKNPSLNPVMRRFSPGTRKDRV-----CVAQFSYKGKRALP----VASDRV-AQES
18 LRVHERRLTT-PL----S--TPDP---VLADFRRF---CRAWAEGHLDPNPLPLGVV---
19 -----GLSQAAASYTHSVRKGGHM-----ARCR---
20 --EILEEMNASN-----YCEPPTRP
21 P-----GLLPQE-WWDGW----VKER
22 -----LLTHCRERYSTPGRE-----PP-----
23 -----KG---TAVIVYERGLKARIVTKIETEAC-ILGHQAR--MRLVSAMRKMP
24 EL-----SALSGNHKEFLTAFNGQ-----EGWILSSDLTAASDLL--PLDL
25 VSAGVEGICDSGR---LLPDEELGLRVCTGPFDL-EW-----V-----
26 -----RG-SPRR---SSTGILMGAPPTWCLLSLIHLFWLEQARR-----
27 -----VYGNSSGG-Q--LPARIFGDD-LVAAFDKRQKLAYERSVTVCHGLLSKGKHSFH-
28 --DSH-GVFLERLFVASRHIVEVDHT-----
29 -----NRGPLQ
30 VMGRL-----
31

32 -----TTRHNKIKSISLL---RTLPLR-PLTVLQVSVA
33 GRRHV-SKGTLPTPLA-VGGVSDGLMQH----GFPPELIRRCQI-ALWPGLPAFYRKL-
34 GIPAFLPT-LLGGGGGLILPEGWEAPITRFS-KLVRASTTALVTGNGRLSYV-----
35 PLARL---PPPVLQMALEEAD-QLLPAHPHRVSHVTPKPWLGHIAW-----
36 -----HDMGVWNDFRITAAMVAQEGYLLQ-----MPME
37 CLGRPL-----NITVGRWKD-----QVWKTRLALCKKAWWASV-----
38 -----TRARYTI-----ARAVRKCTEWPBVWLP-----
39 -GSPDPDHGPVYGIPMYIDSKPV-----MEAKAAIALRQ--GRPTSR---
40

41 >KX883562.1
42

43 -----MKILHESSGTDPTAARGAPHPSTE---VEGEESRQRGRSKAAPAPAREGS
44 GADHRSKNPA-----
45

46 -----ADNVSVWGPGRALLSDSLLS
47 SVLSLGPSSTRLITGLRGVTLAIFTAEQDEVLARSLAYRTCR----HLARRAL----
48 -H-EGVEAVLSVMGRWFSSARESFL-----NG
49 AKNPSLNPVMRRFSPGTRKDRV-----CVAQFSYKGKRALP----VASDRV-AQES
50

1
2
3 LRVHERRLTT-PL---S--TPDP---VLADFRRF---CRAWAEGHLDPNPLPLGVV---
4 -----GLSQAASYTHSVRKGGHM-----ARCR---
5 --EILEEMNASN-----YCEPPTRP
6 P-----GLLPQE-WWDGW----VKER
7 ---LLTHCRERYSTPGREPP-----
8 -----KG---TAVIVYERGLKARIVTKIETEAC-ILGHQAR--MRLVSAMRKMP
9 EL-----SALSGNHKEFLTAFNGQ-----EGWILSSDLTAASDLL--PLDL
10 VSAGVEGICDSGR---LLPDEELGLRVCTGPFDL-EW-----V-----
11 -----RG-SPRR---SSTGILMGAPPTWCLLSLIHLFWLEQARR-----
12 -----VYGNNSGGQ--LPARIFGDD-LVAAFDKRQKLAYERSVTVCHGLLSKGKHSFH-
13 ---DSH-GVFLERLFVASRHIVEVDHT-----
14 -----NRGPLQ
15 VMGRL-----
16 -----
17 -----TTRHNKIKSISLL---RTLPLR-PLTVLQVSVA
18 GRRHV-SKGTLPTPLA-VGGVSDGLMQH---GFPPELIRRCQI-ALWPGLPAFYRKL-
19 GIPAFLPT-LLGGGGYLPEGWEAPITRFS-KLVRASTTALVTGNGRLSYV-----
20 PLARL---PPPVLQMALEEAD-QLLPAHPHRVSHVTPKPWLGHIAW-----
21 -----HDMGVWNDFRITAAMVAQEGLLQ-----MPME
22 CLGRPL-----NITVGRWKD-----QVWKTRLALCKKAWWASV-----
23 -----TRARYTI-----ARAVRKCTEWPBVWLP-----
24 -GSPDPDHGPVYGIPMYIDSKPV-----MEAAKAAIALRQ--GR-----
25 -----
26 >GENC01041260.1
27 -----
28 -----LRSKVDQWFDQPS-----
29 -----
30 -----
31 -----
32 -----EFGDTLF-----
33 -----R-----DFPSFRDDTRFTN-----
34 -----QGDENAAQYNEY-----
35 -----AMSLVAADENS LP-----EAKL-PSRG
36 VTEYVVVDVTAD VH-----LERER--ISKIVRDI--SLRWLA KLTGPPRA-----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 -----SVVALRERGFKARI VTKSPVELV-ECGHLLR--SLVWPMLRNDP
47 HV-----HGSLSGARLESFINDLKERPLVNPAWVGRPSIISADLTAATDGL--YRWA
48 ISAVWEGVCDGAG---FGEDVRKMGRVLLGPMKM-SYPAS-----A-----
49 -----GV--DEFT--SKRGCLMGLPLSWFVLNVINIWAAKSAVE-----
50 -----KSCRGRVGKPA-D--FRFAICGDD-LVGIMPARAQR LYQRNIEVVGSGLSEGKHLES-
51 ---DRL-AFFTEQAATFTRVT-----
52 -----RSAPRG
53 TLAEK-----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----AGLREEYLEPVEMV---GAIPVK-SIVHPSHFAGL
4 RQGP---AATPSWAA-AGPALSSLAVP--WVHDSTLKVVSRGLGR-YLRPE-WRLLASS-
5 GIPPELPR-QLGGGGFPARIPIRGFLAAPK-QWRSALTCAALLGDHKWLS-----
6 RLRNSWLTTGVYGDVMAEALSM-AEEEFQDYPLAELEIPLSEVRTTR-----
7 -----FTGGGLAITEADLRTRVATVWAPA-----LTIA
8 SPFIEG-----RXYCERFHS-----FRKRLSLMVKRKRLPGHL-----
9 --QPAQFNDTEA-----VSRLISLSEGRSFLVP-----
10 ---RESLPPGLNIDF-----IGTRRPCGDGE---DRAIEVLT--
11
12 -----
13 >LA534138.1
14 -----
15 -----
16 -----
17 -----
18 -----
19 -----
20 -----
21 -----
22 --RGV-----DRAR-
23 -----ELLFQSSFARALP----PADDLV-CKEA
24 LRDHREILNT-PV---S-TPVA--VVSELRSY--ATRWGQRMARFANT--SVA---
25 -----KL-----SASSSATLDYSRRLLGGMR-----SDLK---
26 --GMVDGWFEEP-----ASDVTEFHYPPTFSDPTRFTATGEMNRAMHPAPTRGSV-E
27 Y-----VVNVLE-DPDLE-----RSRV
28 ARIIRDVSLRRFRAREG-----PL-----
29 -----PC---KATTVKERGYKCRVVTKSPADV--EVGHLVR--SVVWPMLEHDP
30 RV-----QASLDGGRLEEVFKGFSDRQIECPASLGSLLLVSADLTKATDGF--SRES
31 IYAVWDGVCEGAD---LPSDVHALGRRILGPMSP-EYEDD-----E-----
32 -----LE-DLES---SSRGCLMGLPLSWFVLNIINLWACESSIR-----EC
33 CHRLSLPSEVCKD-L--LRFAICGDD-LASVFPAEAHDGYERRISDVGSGLSDGKHLVS-
34 --LHV-LLFTEQMCWFEEFERPA-----
35 -----PPYTLL
36 AWLKP-----
37 -----
38 -----
39 -----
40 -----GKAVPEGFRATYTCFTAVERMI---DYVPVR-SLVHPGHFAIK
41 RVMGPIS-FELPSWAT-SGPAITSAIPE--WIGVEKRRRIHQVVK-VLRPE-ARALASV-
42 GIPPVPR-EIGGGGFPPAR-PGRVLRDCP-RSYRLFLSLLVGDQKKAETVAR-----
43 RIINLWRTCGVAGDLLSDALAEAEFEVEQKPLWKSPSDGSLAERLST-----
44 -----VIDDHLTMTTEDDALLLASVWAPA-----LGIAGF
45 QQGRPY-----CKKFGAFAK-----QYRALVGAGSARITGRG-----
46 -NPLRNEDEAAV-----ALRLKEICAGPVVFVP-----
47 ---RDRIPPGLVVVQGM-----CSRKRAVGLHRKIRSNRMESA---
48
49 -----
50 >LA710658.1
51 -----
52 -----PPQKKTLHDNPRVDVWWNACEAV-----LGSRGAIKPGRGATRQLPRST
53 RQRAAFLSVSLPSAAHT-----
54
55 -----
56
57
58
59
60

1
2
3 -----KLLSLVW-----EALLG
4 VAAAYRPLPVWGSRF-----GNKRCLELQRLAL----WLVRTVD----
5 -G-QGIGATIAWLKSAASGARN-----TCVTG
6 A---ALLPSTRFLVRKILIGDTRLE---SLDQLSFLSRSLP---EGDSVV-ESRT
7 LIAHRDNMTT-SF---R--TSPN---LISNLRGF---AMAFATRHLKKEDLREEVV---
8 -----PTPSASLDSSRKEGGCR-----EEVR---
9 --RMFMRWLIEC-----PTHVVPHDAWNQAIFSEWLDPS-E
10 V-----YGLRRAQATSHV-----ARAA
11 A--YKTAAA---ETLEH-----
12 -----RVLTVPERGWKRRVVSAPPAT-VAGTVLN--RAMLRGVRRHG
13 PC-----AMFLRGDRKGAVEQAVSHASV-----GCRFVSTDLTAAATDRL--PLDL
14 VRAVVDGLCDGWS---GLPPVWAEALYALTGTQTL-RY-----P-----
15 -----WG--QTIQ---SESGVLMGLGPSWPIMSIHAWWVELAAS-----
16 ---RVGIPRSVA-R--RTTAIGGDD-LLGAWPPRLEESYRRLVLETNGKPSKGKDFSS-
17 --DTS-GNFTEMTFWVVGEA-----
18 -----DGTPQI
19 RWS-----
20 -----
21 -----AAIPTK-GLVGTSID---
22 -----ELGAAYES-LGSEPGRCLR-----RRVLK-ALRPHAWRVCREA-
23 GSVVNAPR-LLGGAGLPPLRGSLARVD-FK-KWHALALGKFLYG-SGQDQIP-----
24 FSPPSWV--EAADPAVWEARQV-AE---QRLRAEAEIGIVSFDT-----
25 -----PLCGTPGSKHVVQSISDQMAWW-----
26 AGARVFSD-----TPFPPVATE-----MVSLLKYHRLIKRWCSSRTKEGIPSSL
27 AVKSGRNSRFAL-----TAKARRNRDRWVIR-----
28 ---NALFDNGIPVD-----PRHRPII-----
29 -----
30 >U90136.1
31 -----
32 -----MHHKVNVKTQREVHFPM DLLQACGASAPRV----ARVS RATDLD RRYRCVLSLPEER
33 ARSVGCKWSSTRAALRRGLE-----
34 -----
35 -----ELGSREFRRRL-----RLADD CW-----RAICA
36 AVC-TGR-KFPS-FSVDTRP-----ARARLA KVYRMGR-----RLLVGVV----
37 -C-RGE-SVVS D LKQECADLRRVIF-----EGSTR
38 IPSS-SLWGLVGVL GW-----TSPE----RAMQLTFIGRALP----YGSPDV-ERRA
39 LASHAATLSI-PA---E-CHPN---YLVA AEQF---AKSWADDNLPRKFR--IYP---
40 -----I-----AVQE S SCMEY SRAQG GLL-----QSFR---
41 --KGFVGY-----DP-A
42 A-----PSADPD-DLELA-----KERG
43 F-SRIRASWYSTFRYRGEL-----KSTNQS-----LE-----
44 -----A-----RVA VVPERGF KARI VTTH SASRV-TFGHQFR--RYLLQGIRRHP
45 AL-----VDVIGGDHRR A VETMDGDFG LLRP---DGRL LSADL TSAS DRI--PHDL
46 VKAILRGIFSDPDR-RPPG TSLAD VF DVL LGPYHL-HY-----P-----
47 -----DG--SEVT---VRQGILMGLPTT WPLL CLIHL FWV ELS DW-----
48 APARP NHSRGF VL-G--ESFRICGDD-LIAWW R PERIALYN QIAVDCGAQFSAGKHLES-
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 --KTW-GIFTEKVFTVKPVKMKVRVRSEPSLKGYVFSRSSAFSCRMGGKGITGI-----
4 -----RAARLY
5 TIGAM-----PRWSRRIRDVY---
6 -----
7 -----PGSLEHRTASQRYGEPEVTYRFGRWSSAIPLR-WAVRAPTRTVG
8 NPV-----QLPDWFT-VGPAASSVAAD----SNAFGAVSRVLR-RMFPGLPRKLASA-
9 GIPPYLPR-VFGGGGLVKSTGLTTKIGAVASRRWMSRIGHDLYRSRERKS-----
10 TLGRVWT--LSTSPAYAASLHE-VEKFMDRPDIILTRKCRNPML-----
11 -----KHARELGLFEEVFESRVGGGILWA-----
12 -----SLNGKALVE-SHSPSILQVSRNLLRSLACPSGGF-----
13 -----LRPSAPIGKL-----VQRHTLPRGT-VWFLE-----
14 -----SSATDSARQGGMGL-----PPPPPPPPLGGG--GMAGPPPPP
15 MGLRPESSVPTSVFTPSMFSERLALESLFGRPPPS-----
16 >MF176258.1
17 -----
18 -----MPLVARPT-----SMRVATRSEVSRRV-----
19 -----
20 -----
21 -----
22 -----
23 -----
24 -----NLAAETW-----KAICA
25 AVATRKWPKNLDFQT-----DRRRFCTFQALAR-----HLLSAL-----
26 -A-RGPVETIKRVSNTAQIRE-----DVFL-
27 -----HGRCPPSAFRGLVGALGWTDAN---SHLQLSYIGRALP----YGDATI-CEHA
28 LKSHADVLGSLGV-----TPPS---LLGAARGW--AKEWAIGNRLSDL--IRE---
29 -----P-----PIQESSCLEYSRREGGYA-----RAIK---
30 -----EHTARCLTLP-----P
31 S-----SVRHSM-KEGTR-----ASNP
32 GRVLKDLAF---ADLSNTA-----
33 -----RVVAVPERGWKARIVTAHAAARV-AFLHSLR--HGLFAALRNDR
34 RT-----KLVAEGLHREAVEALFATGVDP----RCTIVSADLSAASDTL--HGDL
35 LQAIVEGLGDALG---CTSKFMDEFTNAAVGSYTL-SY-----P-----
36 -----DG--STVT---TQRGALMGLPTTWPLLCLTHLFWCDQSLV-----QLHG
37 RGGCREKETPGSS-R--EAEVICGDD-LAAAWRPDRVAAYEEIAVLCGAVFSPGKHLKS-
38 --KVY-GIFTEDIYSVQVSCVPTRVTREVSKWTGPVDKVP-----NLVNELGQ-----
39 -----LLRWRA
40 QRYKV-----GKTGAFQFPPR---
41 -----
42 -----NVTLSYVSKETGLKMNPVSVLREWSTAIPR-WAVRAPKHAPG
43 MR-----GDLPPWVT-VPLAAHAVASA---HPGRWSKVCRVVK-LCYPGMAKFFASH-
44 GIPPYLPR-VLGGGGLPTPLGDRVKIGRVASRRIRKALGGALYRSTDPT-----
45 TFGSIWT--TALSPAYALALAD-AQDLDAMHPDAFKVRQTRSLRPP-----
46 -----FKDNGNYADFVEMRAGK GSLWA-----LRQGAQLVGR
47 SFTP-S-----P--ARVSN-----ALRTKVRKALAKRGYLR-----
48 ---SSAPVSKL-----IERASVKPSDHMWIRE-----
49 ---RPVLSFGRFGGRMSGKKEEPVS-----FQWKKRESVKT--GS-----
50 -----
51 >GDUK01014407.1
52
53
54
55
56
57
58
59
60

1
2
3 SCMRWLR
4 MLKTGTSPAADAAGSVGHSGPTESWTRSGSAV----ARAPLGSDVLASGMSLEKRKALA
5 SRIKRANLRL-----
6
7 -----RSLLGTW-----DAGLT
8 AFFGRQIRRGEV-----ADSVVVWYDHVRS----WLVRVAA----
9 -T-QGVEVAAAELKKFSSDARRAWV-----
10 -----EGSPPRHHFWRSCPELRRKQG---LWAQLSYLGRALP----LGVDFH-EKKA
11 LDQHYADLTS-SF---S-CSEE---VLASARAF---AERWARSHLSRAPGPEDWL---
12 -----EFPSGSSATFGRKRREGGFA-----TDCQ---
13 --ILLESAAEVD-----
14 -----VTPPDH-VPVGP----WSGD
15 VTAFLVTAALVEAHGRRDT-----LP-----
16 -----KG---RVAVVPERGHKVRVVSAMERHAL-ILGELAR--KRLFKGLRKWG
17 PL-----AAVRGDSRSGLVRLLGV-----PGEFVSADLRAASDLI--PLDL
18 ASAIVDGLEASGR---LMPAEIWGLRLGTGPQEL-TW-----P-----
19 -----SG-KTAV---TSRGILMGLPTTWALLNLIHGWAWSQAQK-----AVPVQ
20 SLPRASAARGGRP-A--LSCAICGDD-LVGVAPPVIDAYEAAMRSIGAEMSAGKHFRS-
21 --TRR-AVFLEELWEFRGSAVSVV-----
22 -----DGNPIY
23 RTIRE-----KG-----
24 -----KRVRRQVNSTRLLRADRLVRA---TAIPLK-GMVSAPGVPG
25 -----IDPPDWV-AGLAESALADR---GFSRAKISSLVR-SLRPDLPVRFRAS-
26 GIPPFLPR-ALGGAGLIAPHTRCEAP---RLHRKAIATLIWGYDPDSTL-----
27 AFDRVWS--DSVPGDWRELAID-MV----DAEFTPEVTRVLPVGSR-----
28 -----PTGSWVAIGHPDQVREKAVADNASA-----LSSSLGPSPD
29 VVG-----YPTIPVLAR-----RLRKVREDLASRWASAK-----
30 ---PTKRPVDEL-----LLLWRQRADLRVWVP-----
31 ---LTDQTSRWNDWLYVSATSPYK-----KSWRRLVISSV--SREDLPAVTE
32 WWLPGGRIPSSTMPSPEWR-----
33 >GFKT011160020.1
34 -----PMKEPVDCRLSTPAGFSGTVPP----PGRTKAARQGTIPVRRSRGSASA
35 LPGKIYGWSRRQR-----
36 -----DRFAMLL-----SSFDA
37 ALAAAYSJVVS-----GTRSLPPSLRLFRTMTRKWLSVTAR-----
38 -G-NGVEFAIASAKEFSAACRAGWI-----
39 -----SGTVPDHFFMKWLPEVRRKSG---LWAQLSFIGRSLP----EGGDRH-EIEA
40 LANHKAAALSS-SF---E--VPAD---VLTSLRNY---SEDWARRHLAADPDPSSL
41 C-----EP-----CTGNSATFERTRREGGFA-----QSIT-----
42 --DLVSSSPTDN-----LL-P
43 L-----ESMPFG-PTQQG-----ALPV
44 H--VLEVSL---SRYRNGS-----DP-----
45 -----KG---RVSVVVERGHKVRVVSAMETHEL-VLGHAAR--RRLFKGLRER
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 RL-----KDTLKGDFEATTKAFVGC-----AGTVISSLMDKSASDLI--PLSV
4 ASAIVDGLEASGR---LLPVEVAGLRACTGPQHL-VY-----P-----
5 -----DG--SEIT---TRRGILMGLPTTWAILNLMHLWCWDSADR-----
6 -QYRLEGHPFRATVR--SDCRVCVGDD-LIGVGPDSLLRSYDRNLGLVGMILSPGKHFRS-
7 --NRR-GVFLERLLEFQTRKTVY-----
8 -----EHAVIY
9 RKVGH-----
10 -----
11 -----RRVPVDRSHIPVVTRVTVL---NTIPLK-GLVRASVLGR-
12 -----DDPPVWWA-AAVAESSLLSD----YPRKKIFAAAR-TLRPGLSRQFRL-
13 GIPPFLPR-ELGGAGLVGSSDRVDAP----AFHRKAISLNVWGSDATAAY-----
14 SFIRMWQ--GFECHPWKTAASQ-ETDTWFADYKVTRPGKMYPDR-----
15 -----YGFLDGESLRTKSTMLNSAVYETF-----LGPDPDE
16 ATH-----YPSLRIVAS-----RLAKVRKDLVSRWPSVR-----
17 ---PVGKDLGTI----LEAFEESKLCTLWVTP-----
18 ---YDASGYFDDSLLMDESVYQ-----RRFRQLVIAGLMREDRMDLLFPS
19 WLPPSTVVS-----
20 >M64034.1
21 -----
22 -----MKEPVDCRLSTPAGFSGTVP-----PGRTKAARP GTIPVRRSRGSASA
23 LPGKIYGWSRRQR-----
24 -----
25 -----DRFAMLL-----SSFDA
26 ALAAAYSGVVVSR-----GTRSLPPSLRLFRAMTRKWLSVTAR-----
27 -G-NGVEFAIASAKEFSAACRAGWI-----
28 -----SGTVPDHFFMKWLPEPVRRKSG---LWAQLSFIGRSLP-----EGGDRH-EIEA
29 LANHKAALSS-SF---E--VPAD---VLTSLRNY---SEDWARRHLAADPDPSLLC---
30 -----EP-----CTGNSATFERTRREGGFA-----QSIT-----
31 --DLVSSSPTDN-----LP-P
32 L-----ESMPFG-PTQQ-----ALPV
33 H--VLEVSL---SRYHNGS-----DP-----
34 -----KG---RVSVVRERGHKVRVVSAMETHEL-VLGHAAR--RRLFKGLRER
35 RL-----RDTLKGDFEATTKAFVGC-----AGTVISSLMDKSASDLI--PLSV
36 ASAIVDGLEASGR---LLPVEIAGLRACTGPQHL-VY-----P-----
37 -----DG--SEIT---TRRGILMGLPTTWAILNLMHLWCWDSADR-----
38 -QYRLEGHPFRATVR--SDCRVCVGDD-LIGVGPDSLLRSYDRNLGLVGMILSPGKHFRS-
39 --NRR-GVFLERLLEFQTRKTVY-----
40 -----EHAVIY
41 RKVGH-----
42 -----
43 -----RRVPVDRSHIPVVTRVTVL---NTIPLK-GLVRASVLGR-
44 -----DDPPVWWA-AAVAESSLLSD----YPRKKIFAAAR-TLRPGLSRQFRL-
45 GIPPFLPR-ELGGAGLVGSSDRVDAP----AFHRKAISLNVWGSDATAAY-----
46 SFIRMWQ--GFECHPWKTAASQ-ETDTWFADYKVTRPGKMYPDR-----
47 -----YGFLDGESLRTKSTMLNSAVYETF-----LGPDPD
48 ATH-----YPSLRIVAS-----RLAKVRKDLVNRWPSVK-----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 ---PVGKDLGTI----LEAFEESKLCTLWVTP-----
4 -----YDASGYFDDSLLLMDESYQ-----RRFRQLVIAGLMREGRMGDLLFPN
5 WLPPSTVVV-----
6 >AF039063.1
7 -----
8 -----MKEPVDCRLSTPAGFSGTVP-----PGRTKAARPGTIPVRRSRGSASA
9 LPGKIYGWSRRQR-----
10 -----
11 -----DRFAMLL-----SSFDA
12 ALAAAYSGVVVSR-----GTRSLPPSLRLFRAMTRKWLSVTAR-----
13 -G-NGVEFAIASAKEFSAACRAGWI-----
14 -----SGTVPDHFFMKWLPEPVRKSG---LWAQLSFIGRSLP-----EGGDRH-EIEA
15 LANHKAALSS-SF---E-VPAD---VLTSLRNY---SEDWARRHLAADPDPSLLC---
16 -----EP-----CTGNSATFERTRREGGFA-----QSIT-----
17 -----DLVSSSPTDN-----LP-P
18 L-----ESMPFG-PTQQG-----ALPV
19 H--VLEVSL---SRYHNNGS-----DP-----
20 -----KG---RVSVVVERGHKVRVVSAMETHEL-VLGHAAR--RRLFGLRRER
21 RL-----RDTLKGDFEATTKAFVGC-----AGTVIISDMKSASDLI--PLSV
22 ASAIVDGLEASGR---LLPVEIAGLRACTGPQHL-VY-----P-----
23 -----DG-SEIT---TRRGILMGLPTTWAILNLMLWCWDSADR-----
24 -QYRLEGHPFRATVR--SDCRVCGDD-LIGVGPDSSLRSYDRNLGLVGMILSPGKHFRS-
25 --NRR-GVFLERLLEFQTRKTVY-----
26 -----EHAVIY
27 RKGH-----
28 -----
29 -----RRVPVDRSHIPVVTRVTVL---NTIPLK-GLVRASVLGR-
30 -----DDPPVWWA-AAVAESSLLSD-----YPRKKIFAAAR-TLRPGLSRQFRL-
31 GIPPFLPR-ELGGAGLVGPSDRVDAp-----AFHRKAISLWVGSDATAAY-----
32 SFIRMWQ--GFECHPWKTAASQ-ETDTWFADYKVTRPGKMYPDR-----
33 -----YGFLDGESLRTKSTMLNSAVYETF-----LGPDPD
34 ATH-----YPSLRIVAS-----RLAKVRKDLVNRWPSVK-----
35 ---PVGKDLGTI----LEAFEESKLCTLWVTP-----
36 -----YDASGYFDDSLLLMDESYQ-----RRFRQLVIAGLMREGRMGDLLFPN
37 WLPPSTVVSGFP-----
38 >MF190030.1
39 -----MGNK
40 VGRFSSRSCDRYLKRAPLPGDRKGRVRGRVIDKLEVPRTRSRSRKSKGSRMRGARPPMAES
41 Y
42 KIKPVGWRKGFMTPG-----
43 -----
44 -----REIEGLW-----VALIA
45 TLTSISPRICRFNMRCRA-----DRDRHGAMVRLAV-----WFTRRV-----
46 -----RSSYDLLKSAKKLTAGFREASI-----LRRPY
47 NWVD-APPGWAKRFRPILKGIRLSEA-----VLDQLSYMGRALP-----KGGPRV-LQEAE
48 LKDHKDCMLKPSP-----LPEE---VFGAARDF---ACRWADRYVASGGRLSHPG---
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 -----TTESSVLGSTRRDGGIS-----SSL-----
4 --GLTYSVKPLP-----MDVQDEVRALVSSLK
5 C-----DTPSEC-DVDA-----LRTT
6 S--IREWAIDESSPDVIPEN-----
7 -----KVLALPERGYKTRVATIPDLPRQ-ILGGFVR--SYLIPALRRTP
8 EC-----ADTLRGDDERAAKRIGSSQRSFS---TDMIVSSDLTAASDTL--SREL
9 LLALTAGLVDSGK---FPSWFGPALRQLVGPMLL-RY-----P-----
10 -----DG--SSIT--SQRGTLMGLGHTWSMLSLSHLFWLDYSNK-----
11 -----TVHKTIARSV-E--RTAAVCGDD-LIVRAPWRWINSYHRIAKMCGAQFSSGKHFIS-
12 --RER-YCFLEQLWTLKSKEIVPAYTRALDK-----
13 -----RIPKGG
14 DQFPP-----LG-----
15 -----
16 -----TDPRKSVRRVMKFHIPSKCVRS---TAIPLA-GLTSGAGSDPE
17 KPLV---VGTPTWAK-LGLAVENTSRT----PVAARKVRAVSR-ALYPRRLPRWMRKH-
18 GIEPYLPR-AYGGGGLLPP-GNVSHLGALA-PQVRKLVTVTAYGRCPKTAA-----
19 RLSRIYL--PMNRTATNFTEAQ-EH----TDIRFKRVGAILCRGK-----
20 -----VPFNGHAVKLGLTLTDAPERAAALAIARE-----LSMVLFPPPR
21 NRK-----KSTISPFQV-----GRAFHSA MRKA IKKWP SAH-----
22 -PCSAKVTRRVL----LQKARALEEEFVWWSP-----
23 ---RRRIPEEGFGGWGLA-----PAAQRVLKDSL--RWSSDC-----
24 -----
25 >KX883516.1
26 -----
27 -----MKPPRPTEVHKDFRLPVLLRACRVGGTWS---RDVSRYLVPALQDGVGVRPGE GG
28 NPQTRPRKPAKSVDLERTRR-----
29 -----
30 -----YDLTVWRKEFRSRG-----RLVEGAW-----AAVHA
31 ALV-AAHPDLGR-ADMSTRA-----GRDRIWAWRRLAV----WMVRWID-----
32 ---RDPRDMVRGWKQLAATFREASI-----MQKPY
33 DWAA-APAGFPKRLRPVLAGVRLTPP---VLDQLSYIGRALP----QGDAWV-QRKA
34 LDDHRECLSSAPR---P-LANH---VLA AARDF---ARS WASR HVRPSG---RIS---
35 -----HPKVLESATIQEPRRSGGVA-----AHLC-----
36 --KQGFDL-----PPFEGLTDEL RDQIRAM-G
37 E-----SAPSDC-DLEAI-----GRST
38 A--FRDWALGQLCCPDAS-----PP-----
39 -----RS---RVEAPIERGFKARVVTVPDLPES-ILGSALR--CYLLPALRRTP
40 EC-----ADVLSGKEELAAMRIASSRRPIE----GDWIVSSDLTAASDTL--SFEL
41 LQALVDGLEESGT---LPPWAIRALKVLVGPQEL-VY-----P-----
42 -----DG--SVLV--TRRGTLMGLGHTWSI SLSLSHLFWLDWAHR-----
43 ---SVDPRVADV V-R--RTAAVCGDD-LVVRAPLKWWQSYHQIAEASGARFSTGKH FVA-
44 --KRR-YVFLERIWSLKSSLQRVRLPRERTR-----
45 -----RAGR WV
46 REFPP-----
47 -----
48 -----LGARPVKVRPTVVRSIPSRSVPS---TAIPLA-GLTAGEVDCHF
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 GPVRN---LVLPNWVR-LGMAVETLLHS----PEAARKVRKVVR-ALYPGLPRWMRRH-
4 GVQPYLPR-FLGGGGLLPPGRNVTRIGRLP-RPVAKAVTTWVYGRSPRENL-----
5 SLAKIYL--PTTRRITNYAVGE-QH----TAVRFQQVSAVLCRAGK-----
6 -----KPFDGHGRRRLGTLSEAPEAAALAIARE-----LSMVLLPEPS
7 EERAYRR-----LLRPSQVGQ-----TFRRLVLRANRRWPGSK-----
8 -PICPRAPRRAL-----LARCKAVEEGFVWWAN-----
9 ---PRLLTQFGFGGWGLE-----PAAQRVVKDAL---GWSTFRPTPG
10 RGTH-----
11 >GEND01007370.1
12 -----LDHEPPT
13 MGSSKKTARETHVSNVSLHALLRNPRVGGWS----KDVSLYLAPARRSGVQGESGERG
14 LPSRTGLATHANLEHRGASRRGQGHAKRGPEPRRKAPKGGYTRQRCWSREFKSSG-----
15 -----
16 -----RKLEAGW-----NALTA
17 ALAACHPAALRFDSLVR-----DRDRVGVALKRLAV----ACVRRLD-----
18 -----RGVGELMAELKSLATTFRASI-----RQVP
19 YSWDTCPAGLRRHLEAALRGFRLTTE----TLDQLSFLGRALP----PGDSTV-EERA
20 LRVHRDVITSDPA---P-IPPG---ILKAAEAF---ACTWALTYLHRGGRLAHPK---
21 -----TMESSTITSRREGGVA-----ADLC---
22 --MKSMDLDPLP-----ETVQAEVRRQVLELEG
23 -----PLPSEC-DLSAA-----FRTT
24 A--LRDWAVTTLTEPGYVPTH-----
25 -----KVLPIPEAGWKRIATVPELPAS-VAGGLCR--SYLLPALRRTP
26 EC-----ADVLRGEEDKAARRVAASRRPIE----GDWIVSSDLTAASDTL--AFGL
27 LQALVRGLALSGR---LPPWVRSLEVLVGPQRL-TY-----P-----
28 -----DG--ETVI---TARGTLMGLGHTWSLLSLSHLFWLEAAHA-----
29 ---SVPRPVSEQV-R--LTAAVCGDD-LVVRAPLGWIRRYYEEVASLTGAQFSQTKHFA-
30 --KRR-FVFLEELYTLRGTQQAASTVLGPTS-----
31 -----RGGRLV
32 RAWPA-----PG-----
33 -----
34 -----TRPPKVRVTRLRTYASGTVK-----TAIPLA-GLVAGDVMTHQ
35 VMA----DSVPLWSK-IGLAVESLAST----RERALRVRRVTR-HLYPALPSWLLRH-
36 GIRPTLPR-ALGGGGLLPNNSNCMRLGRLP-RRVAKMVTAVTVGRCPRTAA-----
37 RLSRIYL--PLSRRATNFIAEK-VH----TDIRFKRKGAVLSRSGR-----
38 -----EPWNGKGVRGLGTLTEAPEQAALAIARE-----LSMVLLPEPK
39 KGGSSH-----NAALSPFQV-----GKQFLALTAKRRWPGAK-----
40 -PVSARAPRKSV----SDRTRAVEEAFFVWWAP-----
41 ---SHRIPEEGFGGWGLA-----PAAQRVLKDSL---GWSAPSP---
42 -----
43 >GENB01008321.1
44 -----LDHEPPT
45 MGSSKKTARETHVSNVSLHALLRNPRVGGWS----KDVSLYLAPARRSGVQGESGERG
46 LPSRTGLATHANLEHRGASR-----
47 -----RGQGHAKRGPEPRRKAPKGGYTRQR-----
48 -----VCWSREFKSSG-----RKLEAGW-----NALTA
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 ALA-ACHPAALR-FDLSVRA-----DRDRVGALKRLAV----ACVRRLD----
4 ---RGVGELMAELKSLATTFRASI-----RQVPY
5 SWDT-CPAGLRRHLEAALRGFRLTTE----TLDQLSFLGRALP----PGDSTV-EERA
6 LRVHRDVITSDPA---P--IPPG---ILKAAEAF---ACTWALTYLHRGGR--LAH---
7 -----P-----KTMESSTITSRREGGVA-----ADLC----
8 --MKSMDL-----DPLPETV----QAEV----RRQVLEL-E
9 G-----PLPSEC-DLSAA-----FRTT
10 A--LRDWAV---TTLTEP-----GYV-----PT-----
11 -----H-----KVLPIPEAGWKRIATVPELPAS-VAGGLCR--SYLLPALRRTP
12 EC-----ADVLRGEEDEKAARRVAASRRPIE----GDWIVSSDLTAASDTL--AFGL
13 LQALVRGLALSGR---LPPWVIRSLLEVLPQRL-XY-----P-----
14 -----DG--ETVI--TARGTLMGLGHTWSLLSLSHLFWLEAAHA-----
15 ---SVPRPVSEQV-R--LTAAVCGDD-LVVRAPLGWIRRYYEEVASLTGAQFSQTKHFA-
16 --KRR-FVFLEELYTLRGTQQAASTVLGPTS-----
17 -----RGGR LV
18 RAWPA-----PG-----
19 -----
20 -----TRPPKVRVTRLRTYASGTVK-----TAIPLA-GLVAGDVMTHQ
21 VMA----DSVPLWSK-IGLAVESLAST----RERALRVRRVTR-HLYPALPSWLLRH-
22 GIRPTLPR-ALGGGGLLPNNSGNCMRLGRLP-RRVAKMVTAVTVGRCPTAA-----
23 RLSRIYL--PLSRRATNFAEAK-VH----TDIRFKRKGAVLSRSGR-----
24 -----EPWNGKGVRGLTLEAPEQAALAIARE-----LSMVLLPEPK
25 KGGSSH-----NAALSPFQV-----GKQFLALTAKALRRWPGAK-----
26 -PVSARAPRKSV-----SDRTRAVEEAFFVWWAP-----
27 ---SHRIPEEGFGGWGLA-----PAAQRVLKDSL--GWSAPSP---
28 -----
29 >GBWZ01004682.1
30 -----
31 -----GS RTPWRKSLPQMNRGSR-----
32 -----
33 -----
34 -----QKV RATW-----SVIRA
35 ALAASGFVRCDSLTLA-----ARDSANGLCRLAW-----WLWRTTA-----
36 -H-SGIEYTLKLLKTWLSETRRWVV-----EE
37 VPKPKRGGGMARFLHGNFRSMVDSNA----KALQLSYAGR ALP----VGGPSV-ILRS
38 LRGHQDAFASNPV-----TPPE--LLTSARIF--AREWAKRHLSSSPSFARAG---
39 -----LNESSSLESTRAKGGV-----AQTR---
40 --ERIYAI PDEA-----LRERKSSLQSREE
41 F-----ADLSSC-DLDIV-----ARES
42 L--LRDSA EELNRSLENGNPQA-----
43 -----EVRYCLERGLKCRVVTKSPWALV-YLAHSVR--HELLNALRNDK
44 RT-----KHVLKGDHQEAVEDVFSTPRRID----NRQIVSADLTSASDLL--PLDL
45 VGALVDGLRDTAR--LWWPEWMLKVLGIVTGPMDL-WYPA-----
46 -----ELGEDELTIVGTRCGILMGVPTTWSILSLVQLFWAEEAWR-----
47 SRFPNPAWGHPSM-T--PETAICGDD-LVGVWPKSVIKTYEAVAKRCGAQFSAGKHFHS-
48 --RTH-AVFTEEIYRLRSISEETPRCPDWESLPKLQQRWVRSGTKRR-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 PRKRLV
4 NFLRT-----
5 -----
6 ARLGGPVGVVKY---PSIPLA-GLARPSVLPGS
7 Y-----QTVPWIA-IGPTYEAVAAQ----TGNPSAARRVLR-TLHPNAWRWMRER-
8 GIFTPELPR-ELGGGGLPLRGSADVLK-LP-PFLRFGVAKLLYGTKFRDLR-----
9 SPGALWS--TTFAPEWRVMAAD-HAEVKFKTSAVNCHPGRVPFHKG-----
10 -----FELGGVKALEDVTTRLAGELTLMMEHP-----K
11 DGRQAF-----SVVPRFIAR----SVRKFYSARGKVARRSKMKG-----
12 -IDPSGVVRTAL----LKRWASLGQTRVWFSE-----
13 ---PAAAGEGFVLGIP-----GQVKRRIAPAL--GWARP-----
14 -----
15 >GBPT01024369.1
16 -----L
17 RMAVRKPSTQAARERVTTSLVEAALIGSRL-----VPLAGIVRPARARALRSDRGATL
18 NPKRRGVGQPLEPIWAQGST-----AVPIPNSSSPHKACLKPISRQQDGKPPKA
19 SAMKKGHKIKDHFRRFRLSPRSSEKVRASANPPKEDHDVDWSSVR-----
20 -----RLYPEWGRISKRGHRQLKTVW-----AAFRA
21 ALAAVGSTRVSLRTNS-----GREALRKWCKLAH---WLVRTTG-----
22 -L-SGITHVLATLKAWTATVRRWWV-----EGVRE
23 QD---RQGGLAKFLAGSLRFLVTRDD---AAVQLSYVGRALP----QGTLSV-QRKA
24 LVSHKKALGE-DY---R--TPRH---ILREAREF--AAGWGDRYLRSPVPL--MNC---
25 -----GFSQSACLEVPTSKGGTL-----EATL---
26 --ERTNALQTDV-----NWSKAVADWRDTLKS-S
27 I-----PTLPDC-DREIL----ARGT
28 A--LRDRSIVELREAQSQSG-----HP-----
29 -----RA---AVA VPERGSKARVVTKSPWALV-QLAHPIR--GWLLSGLRRDP
30 RI-----AKVLEGDHGAAATGLVDVPRLLAD---GREFLSADLTSASDLI--PLDL
31 AAALIEGLMEGTG--DTLPDWAREVRLATGPMLV-DYSTAWPPSGGRRPLP-----
32 -----GP--CEFI--SRRGALMGLPTTWSLLCILQLFWADRAWR-----L
33 TFPNIAQRDTPGNPGMTPATIICGDD-LAAWWPKQVSDAYEQTVVSCGGIFSVGKHFRS-
34 --RAV-VVFTELLFHAEKSTVRALAGAGPDKISAPAIASWER-----
35 -----EGRPFL
36 MKVRR-----PHAKSTARRRSER
37 GRRPMIPKM-----
38 -----TRGIGPHDLSDHRRVVTRLVPF---KSIPLK-GLLEAEGGPL
39 ADNH---VRQPRWVN-IGPVADRISKM----TGNPYAVRRVLR-TLFPHYWGWFQKH-
40 GFVPTLPR-FLGGPGLPPRTGDLRLLA-LP-RDLLLKVRTLVHGTKWHLLR-----
41 SPKAPWT--QMGAGFYRQMVG-HAEARLQTHARCARVGRVPYNGKG-----
42 -----VLLGGVDVYEDLATRFENEATVGIFGC-----RDD
43 SRRSAF-----KIRPGQVSQ---ALRKFWKGVKVPTGEKHRLG--GEERRR
44 GRLLNKTSARNL-----ERRWIELTSSRAWHAD-----
45 ---PMAAQEGFVLGLN-----RRVKLDLAIKLGQTGRAGGETESPG
46 PSAEEKSPAGDAPAASGSPTAS-----
47 >GBPU01010966.1
48 -----MPSNK
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 PNTAGVSLEKTLVEAYLLGSALVPLSDSLPS----RSGGGAVTPRPNVGPRNPVGPRR
4 STLHLPLSGTRSSGANPIKIPPLPKFPARSSGKMRLIDLIKTSKSkrkTPVLRSK
5 KNGKAKRKGRSSAFSHKWVRDNRGRVRWDWTKLRALAPELASLD-----
6 -----RKARGLC-----EQAWT
7 AVRAALAADGRRRATLMTSS-----GRTAVRSYCRALAK----WLVHTTA----
8 -F-SGITHALKTLKDGVSLARQ-----YVVEG
9 VRPKDRKLGLARFFRGDFRHVYDSDK----ALLQLSYIGRALP----LGGPSV-VRKA
10 LIKHKEVYSTEH-----TPAH---LLEAARRF---ASEFAQDRLGVVPL--LSS---
11 -----GFSESACLERPAAKGGSA-----AHTH---
12 --AVYNSLVSAP-----EWEEAKR---EWLAEIAKED
13 L-----N-IPDC-DRDIA----VRVA
14 I--LRDFSLRKLKESLEETSG-----IP-----
15 -----KA---EVTVVTERGLKARVVTKSPWALV-QLGHPVR--AWLLSALRKDP
16 RT-----SGVLEGDHNAAKGVVDAPRLAD----TRRFLSADLTAASDRL--PRDL
17 VSALVEGLIEGSG--ETLPEWARWVMRVSVPMLL-DYSKAWNPKLG-PPLP-----
18 -----GD--DKII---SRRGILMGLPLTWVLLSLTQMFWASQGWK-----A
19 ILAHMTRSAGPSA-T--PSTVICGDD-LLGWWPQVVIDQYQSVALACDAQFSEGKHFVS-
20 --TRR-AVFTELMFEAIRPASNVPRIALWER-----
21 -----RRFPRL
22 TFAPT-----KKPKWRKVLL-----
23 -----
24 -----GPQPQCPGQRDTRRVVTGLRRT---PLLTLR-GMVEPEAGLDS
25 SY-----VRQPRWIN-LGLVYQEVAAQ---SGRPAARRVLR-TLYPGYWSWYRQK-
26 GFQPTLPR-FLGGPGLPPLTGDVNRLH-LP-KPLRLGVAFLYSTPWRDIR-----
27 CPSSLWS---TGTAKWVREMAA-DS----AEWKFHKAFAKSSRVGGK-----
28 -----VPYGGLVLLGGPQAVERETTRLANEL-----SLILD
29 PKESSRW-----KKKPREVAK--ALNGFYARVLKRKPGVRNWVRKG-----
30 --VSKGVPVGAL-----EARWRRVQNAYAWYSH-----
31 ---PTAAAEGFCLALPQGVKSNLSQLRQSGVPARVERKSKMAREPSLDGPDDLQLSQL
32 AEFPLPLSAGASRRR-----
33 >GBWH01015246.1
34 -----
35 -----
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 -----
47 -----GRALP----VGTQVV-IRRS
48 LEAHKVNLLGEYH-----VPEP---ILERARAW---AKCFAKDKLRSVPL--ITC---
49 -----GFTQSACLENPSSKGGSF-----AITL---
50 --SRCATLETET-----EEQRGGRTQHEWKKSIDKEV
51 L-----DAVPDC-DAHIA----TRTA
52 A--LRELALLEEIRKQVEQHG-----YP-----
53 -----MA---EVAALPERGMKARIVKSPWSLV-TLSHHIR--VWLLHGLRSDV
54 RT-----SGVLDGKEKEAVEGLVDVPRLAD----SRKFLSADLTASDLL--PLDL
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 VAAIVDGLVEGSE--GTLPAWAQEVFRLSTGPMVL-DYSKAWDGSVPL-----
4 -----KGP-TQMI---SRRGILMGLPTTWTLLSLVQLFWSDTAWR-----N
5 GFPEIMARRGFSAPS--PATVICGDD-LAAWWPEKVIEEYEALARTCGALFSAGKHYS-
6 --PKN-LVFTEVIFRSVVSTAQDHPPKRSFWERERLPFALIGKTVRSSNVGAKGQVEGR
7 NLTKKERRAKLRQSE-----RTITLV
8 RRDGV-----
9 -----
10 -----RRVSSLSSKDDRRVVRKLA VM---PAVPLR-GLVEPELGLNE
11 SF----IRQPRWVN-LGPAVDCAIQI---SGLPHAVRRILK-ILYPGYWAWFRQQ-
12 GFEPTLPR-FLGGPGLPPLTGSPARLR-LP-RRLRVAIGILLYATPWKEIR-----
13 SPAAVWT--CLGAGSVRESARE-HT----EWKFSKSAVACRKGKV-----
14 -----PFNGNGAVLGGHSAVEDVTTRMYNEL-----ALILDPE
15 DRRSRF-----KVKPREVSS----AVRKFYDRLLPRDPKQRVKT---HPYIT
16 RGIAPSVPVSAL----RRWEGILGQRTWYST-----
17 ---PGASSEGFCGLA-----AGVKHELSALGHSGRRTRRSPS-
18 -----
19 >KX883605.1
20 -----
21 -----MVARLVEDRRRDNRSGAAAKQPA---SRLGGFMGHRRGDRQRVAALFPW
22 RRRLGKSEW-----
23 -----
24 -----TRLSEAW-----RVIQA
25 AIAAAGFLTGVLEDPA-----SREFLRLGLCRLAW----WMVKTGH----
26 -Y-TGVERVLSDLKG FAGKARLWVM-----EE
27 TPPTLRRGSFARWFRGIGRSLVDRDE----RAMQLSFIGRALP----AGGPRV-VAQN
28 LKTFREVVTA-PY----E-TAPH---LLSEMEAF--AKAWASRYLQPAEIPPPWM---
29 -----GEGATFERTRTEGGFA-----SQSV---
30 --EMAGLFDQAY-----RVDADWDELRIRAAE
31 V-----AGMTDC-DVDIA----CREA
32 Q--TREVAMGTLSRLRRPIAV-----
33 -----EALAIPERGLKARVVTKCPWALV-YLGHFLR--SWLLQGLRRDP
34 RI-----SAVLEGEHQRAVLDLLKSETAEVT---DSILLSADLTAAATDRF--PHDL
35 VQACTRGILEGAR--CPLNSDWTQVFADLTGPLLV-----VPP-----
36 -----RGEGSDYL---TARGIMMGLPTTWCFNLNFWSERAWS--LCPRVNPVI
37 HRASLFRPGHHAR-T--PATVICGDD-LAAWWPHDVCDQYERIAVACGASF SKGKHFRS-
38 --ARY-LLFTEEPYEVLVSETERRHG-----
39 -----RAVERP
40 QEWT-----AD-----Y
41 LFPSR-----
42 -----TRSALEPSGRGSGWHYTGFRMM---KSVPLR-GLVRPTHKPF
43 H-----HTIPPWAA-LPFAVQASVRA---SGD GARVRRILK-VLHPGIWAWARNN-
44 GFSPTSPT-ELGGYGLPPVHGRRGEGYLLP-PWHRRGISVLLRH-TWKD LI-----
45 SPSTAWT--LIESGSHWRGMGE-EH----TRHKLAV-SAVACKRGK-----
46 -----VPFEGKGSELGTVDCVEQLTRMARELRMLLGPE-----R
47 QPRDEF-----RVAPRQIAV--RV---RKMFRGLEANKGVPGGRL-----PSV
48 SLVPASGSVQRL-----LSEYESRLTERCYFSA-----
49
50
51
52
53
54
55
56
57
58
59
60

----PGPAREGFVLWLP-----GSARRDLAVAL---GWTRSSAHAD
D-----
>KX883602.1
-----MPRNKA
RHRWVCLYELTLSVATLLPARPGTGPYLGGR-----PNGPDTGRFLPRFPRHPNIRCPK
ERGNAATGKTNGNVNGIL-----
-----SKVLRRLPLGNQRWTRLEEAW-----RTILA
AIAACGFVRLDSSRRA-----HREALRGLSRLST-----WIVKTAA-----
-L-SGTEEVLRELKTWSARLRA-----HVVNS
LPPHKRRRGFARYFQGVRLSLIDKDS-----RALQMSYIGRALP-----VGSFRE-TEGA
IRKFREVTQV-PF---E--THPD---LLDSIRDW---GAHWAARFLDTPAM--PPL---
-----WRSSSGCVDFSRLQGGLA-----
--EAAVRAADFA-----AAEVESREDWDDLNTAAE
V-----PEMTDC-DVDIA-----VREA
Q--LRDVAL---TVA AHLPD-----LP-----
-----PV---EVCTVPERGMKCRIVTKAPWCLI-HIGHFLR--SWLFGGLRKDR
RV-----QGVLAGDHAGALRGQTERVYSQKPTEE-ENICLSADLTAATDLF--PQDL
VAALVDGLLSGAR--KPPTAEIQEVFRKLTGPLNA-RLP-----
-----DG--DGFA---IRR GIMMGFPTTWTLNLVNLFWSETAWS-----ASGSL
PFQDVALSAQPAAS--PRTIVCGDD-LASVWPPKVADRYEEVAEACGAKFSVGKHYRS-
--RDY-ILFTEEIFKVSWEKQEYSLMDSAKP-----
-----RRTTLA
DFFPV-----PR-----
-----APRKSTMKR YRVQAKLAPV---RCVPLR-GLVRPLHAPKD
R-----RLRPSWAA-LP-AAVEAAMD-----LGAPVAVRRVLR-VLHPGIWCWARKR-
GFSPTLPL-VVGGYGLPPLRGSVRRVS-LP-KWLRYGLAAWLLN-SWKDLE-----
GPSRCWE--GVCRPVSWQTMAR-EH----ADVRIAS-STFAVRKGR-----
-----TPPPGF---HLLGGLEAIQNLSVKFEGEELCVLL-----GTEEG
SLKRQF-----GLAPRRVAN--SL---RKFFARMVS--DHP SRR-----
-PFPSLLGRTAL-----VQRWRSRAEERALYST-----
---PRSAVEGLVLSLP-----VRAKRLLAEAL--GWV-----

>IABX01132835.1

---PLPKQQHRWVHL SRVSLVGAASVGARRG----APLSIFFEERA STD RQKGRLNPE
SAADYCLAEIARVSRLAWT-----
-----PDDTGIPKGPLYKVFVSRPM P R ALR-----
-----SSLVETW-----RVVLS
SLAAAGFIRLDVSKWA-----GREALHGLCDLAW-----WIIRTS-----
-L-TGIDLMFSQLKGYCSALRL-----AVVSE
IPPAERRRGFARFFRGVLRSCVDRDA-----RALQISYIGRALP-----KASPAQ-CTAA
LENFRAVLKR-PH---T--TSPE---ILD RAYQF---AVEWAGIHLTSPEI--PNP---
-----FLGAGATLETTRAQGGFS-----EETR-----
--RKAGOFDWOL-----GAPDOEMWDERRMLLNE

1
2
3 I-----GSLTDC-DVDIA-----VREA
4 M--TRDVVL---DRIHRMGD-----FP-----
5 -----EV---SAIALAERGRKTRVVTAKPWEIV-YLGHFLR--SWLLDGLRRDS
6 RV-----STVLMGDHAEAVKRLSREALYCR---GDEILSADLTAASDLL--PLDL
7 VRAITDGLWEGTRKGCRPEERFRLVLDRLTGPLRV-EMPDLY-----
8 -----DAFI--SERGIMMGLPTTWCLLNLVQLFWCEFAWM----SGQQY
9 RHMHDPRLRGSPRS-S--PNTAICGDD-LVAYWPSDIIGRYEQVVKESGGMFSVGKHYS-
10 --GRY-AVFTEQFYDCKVTRFSAPGPALRK-----
11 -----GERTLA
12 DFLFP-----CM-----
13 -----
14 -----PSKGGEITPNQVGEVRNV---PTLPLR-GLVRPSHGPGD
15 R----HLMPSWAA-LPLAVEHTMMY---RPENVKALRRRALR-ILHPGIWCWARSR-
16 GFAVGLPT-SLGGYGLPTLSGCPSCR-LP-KWVQVGVAALLFK-VPWRDII-----
17 SPARYWQ--MVRQNLPHSTMAA-DH----ARYKFET-SAVACRPGK-----
18 -----SPFNGNGHGLGLGAPDEVAVRLANELFMILEDL-----
19 PPRRPEL-----RVRPRTIAS--GV--RKFYQRSLTK--RYPVGI-----R
20 GVKRVGKTWMDL-----QRQWDSSRQNERLWWTN-----
21 ---PRSASEGFVLGLP-----RSAKRRLASAL--GWV-----
22 -----
23 >KX883500.1
24 -----
25 -----
26 -----
27 -----
28 -----
29 -----
30 -----
31 -----
32 -----
33 -----
34 -----MEAVP
35 VTSR-RRCGFTKYFRGSLRSIHDRDC---RALQFSYVGRALP----EASPGQ-CDAA
36 LTKYREELSQ-PA---T--TPDS---ILDSAEEW---ARRWASRFLSGPV---VPP---
37 -----T-----WLSGGACKEVPRAKGGLG-----TRSV---
38 --DVARRATDAI-----CEPSEPEWE----SLRE-----RLN-A
39 V-----EGLTDC-DVDIV----CRER
40 L--VSNQAL---VELSEL-----PRP-----IA-----
41 -----V---DAMALPERGGKVIRTKCPWALV-YLGHFLR--VWLLEGLRRDE
42 RT-----KSVLGDGDHAGSVGGLVGGAGVRLL---EHLAVSADLTAASDLL--PHDL
43 CQAIVRGVLGSR--TRPAPHLDEVWADLIGPLTI-QM-----P-----
44 -----DG--SEFV--NQRGIMMGLPTTWILSVHLYWAEWAWA-----S
45 TVPRHISLGLPSAA--PRTAICGDD-LAAVWPTSVVQRYESIVKSCGGQFSAGKHYKS-
46 --RTY-LMFTEEAFLRVEMRDYDTRRVGDGRRP-----
47 -----GEATLA
48 DFLPS-----
49 -----
50 -----WMVEEPSRDSRWCTGVDH----PLVPLR-GLVRPLHLPKD
51 R-----QPLPVWVA-VPHCVEAACTQ----SGDAGAVRRILR-VLHPGIHRWARSR-
52 GFVNAPT-SLGGYGLPPVRGDPKRAR-LP-KWFRWGISSLRATPWKDLI-----
53
54
55
56
57
58
59
60

1
2
3 NPARAWA--PVMRDIPHDMGA-DHASEKLKVAVAARKGRVPFMGR-----
4 -----GADLGYDAEDRLTVRLSNELLFI-----LDRD
5 TLPSSF-----FLQGPRKVS-----GAVRKMFQHLRRFPIPE-----G
6 SVVASETPMASL-----RRWEEVRTARSWWSS-----
7 ---PRGAAEGFVLGIS-----SSAKRAVASAL---GWASMEPAHS
8 GGPVSD-----
9
10 >GFNE01046070.1
11 -----SDLSGVRRPQTRKVLLS
12 RITYVEATRLQAARCRSVLPSVSGRRRTPAVC----RSRGGVVTRRKCRPVAPASNAF
13 LSSRTGGGLSLHVADSTLSSLSQPPSISLLLSSKAFSTGLAHSPTLIPLLMQDPTWPSD
14 RPSPAAVPAKAVTTAKHADPRPPGRAPAPQPPPNNHPPGPKGRVPARKVGPRPPAGAGG
15 G
16 PGAGDAARKHQPPRAPFGTQREHAVVPGLVGTPRTGQAFRTLCGAW-----DALRA
17 CIAAGFVEFDLRERS-----SRDAVGGLAQLSG----WLTTATR----
18 -G-SGLGYTIKTLKTWSAQFRS-----WVFNR
19 TPTHKRRGGLARFLRGKLRSLVDRDE----RALQMSYLGRALP----VGSDGV-CEKS
20 LKNHRDALTR-PH---P-TDPD--LLARAYRF--ARLWADRHLPEVPNPSHVVR---
21 -----LGHGACKEATRATGGFA-----RAIV----
22 --HLAGRDLPE-----TEARQTIMDRLRQEEDL-G
23 P-----ESITEC-DLDSFVAESELIRAV
24 VKPLLDAPR-----PP-----
25 -----RH---SAIPVAERGFCKRVVTKSPAALT-AGLHHIR--TWLMAGLRSDK
26 TS-----STTISGDHRAAVEAILAGDPVPRE---GYCFLSADLTAASDLL--PRDL
27 LIALVDGLSDADH---SAPQWYFDLRLACGPQWV-SYP-----
28 -----WG--DTIL---TERGTLMGLPHTWCMLSLVHLFWIDQAVW-----
29 -TTPATRIPPALL-L--KRSAICGDD-LAAHWPDREVIGRYEDVATRCGAESVGKHYS-
30 --RDY-VCFTESLAKARCNPVGNCNPTCSGGGRVPYRRTRPRPAPQTPIKRPyGSASP
31 IPSPSLASAVAADPHPALLAKSLGLTPPASAPPSSRISPTPTSKPKG-----RSRPLL
32 PNLPPKGRRRGGRPVALRAPDPSPRDPFPYPPPTVPGRPDPGPRWPRGPKRWPRLGAGS
33 GAKLP-----
34 -----RRRRPPRSNRRIMRRVVGVRWS---NAV PVR-GLIRPNVVPFG
35 REALGDTMGPAPA WA V-IGPAFWAIVGE--NPPPPRLNAARRALK-
36 ALHPHLD A WGRHNF
37 GCLATVPR-ELGGGLPKSKGDALKVGKLP-KWLRQALAVDLYGTRAGESP-----
38 PPIGTWS--GMKPGHHR SMAEE-HSRARLQRMTIRCAGKVPFKGAP-----
39 -----TPATLLGPEVDVEELSTIAMSRSLT YI-----LGAE
40 LLGDRL-----SVNPRRLAD-----RIRRSMVRSTRWAGSR-----
41 -GICQSAPWAKL-----IARHRNLVQNWVLWLK-----
42 ---PGA DEGFAGALID-----KA AKRSLAGTL--GWAIPRASAG
43 GSPARPTS QELSGP-----
44 >KX883526.1
45 -----
46 -----MP-----PRKSNGRKG-----
47 -----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----GRAPRTC-----VALEA
56
57
58
59
60

1
2
3 VLSAGAAAGLTS-----AEVRTLPIGVLAS-----
4 ---ED-NPDLGALKALSAAYRAHWIA-----
5 ----TGDP-----LAGQLSFVGRALP----PAPASA-RLAA
6 LVQHREDLTS-TF---R--TPPS---VLALAREF---ATRWARRHLARAADV--PLE---
7 -----E---ASWPTPSSCYERSASKGGLL-----RHLL---
8 ---ECTSFEPAPP-----G
9 F-----DGLPDGPVTQLW-----VQSA
10 R--FREYGLRGLERCRSAAG-----PP-----
11 -----TH---RVAVLSERGLKTRVVTVGPAWLQ-VLGHCVR--QRLLRALRSNR
12 GT-----YAPLAGARDDEIWASLEGGH-----SEAVVSTDLTRASDLL--PLDL
13 LSAIIDGLQESGK---MSTLEVEILRISAGPQRL-EYPHGW-----
14 -----G---TYS---SSRGALMGLPTTWCLLSLVHLFWMDQVRL-----
15 ---TSRAGPMRA-G--HRFSVCVGDD-ALLATTEAGAQRY SALVRLCGGQPSEGKH VSG
16 GVRCR-GVFLEKLLEFDT-----
17 -----EGGRLT
18 GMRRRL-----
19 -----
20 -----AAMPVK-GVTSVSLPRDF
21 TGALPI--QCRSKGLR-LLVTLDSVLED--LGPAGLVPVRNWVR-TRASWLGRFAREQA
22 LLTPGAPL-RQGGYAFGDPNDRTPEEE---RRFRAISLLPRPSIGAAALR-----
23 SVDPLWR--LASRETLAGEDLP-----LSSGKVVEIGRHPPLSGE-----
24 -----GPLLPTQWRQSELDRLAEG-----
25 -----YPAVAIQSQ-----EERHTSACVALFAALK-----
26 --SLMGPPLLEVY-----TFRVSQLRAAVLKATA-----
27 -----TGEALRESL--GPTA-----
28 -----
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 >HAGX01098760.1
4 -----
5 -----
6 -----
7 -----
8 -----
9 -----
10 -----APPARLPL-----EF
11 LEGVRLAQPK-----SDGHLSLAASATVEYPRSKGGRAA-----
12 AYFTW----AKARMKELGLK-AP--ETVAET---ATHWL-----
13 -----
14 -----YSQEAAQEYYSECLGQ---AEAWLE-----
15 -----SGEPVPHRLAVVGEKG-KARVITVAPTWFVQVLI
16 TMQKILMHSLRFIPEVSAGVFSRDDAYK-----AYAEIPDVSTG--DY-----
17 HFFMSDLKDATNGPPKSGIVDTVLWLFQDQVGA---ESPIPQRDLEIALSLL---MSER
18 KI--TVKN-DYGIPREF-ITRRGILMGDPMTKVILTIAY-----YLIVKAVCAD-----
19 -----VEREHPGCHTWFRIK----G-----DDVVILVVGVSPEA
20 SSTEKHFLRRISDAQFRVSELDTFVSH-AAFYCEGLFL-APRGLE-----IEKISRARR
21 FG---HAYADPVKGRYMVPFWK---VDPGGKDPLVLPLSKLESIVKRLSWFVPDS---
22 VQWRRTYSCAMTATKSFRLERDPSL-----FLPPKQGGIGLLPDAGIRELSTADWNQF
23 IWSWAVDLRATVLRDKIPEFSDPGVPRFVARSSNHAOSHHLRVDPQVIPAEWEET---
24 -FLHLGDVDERLSTRALAEYLGRSPNYSSDISVASYAESVDRFSTVHLGRPL----IKRL
25 KTQGSYLGSLKEVGSI---PPPTDLIEAVCQYRSNEARRRIDLRFMRSDLIG-----
26 -----PRLVVFKPEDPEVIDLEAIKQAS
27 SPGEARTLVQAGDPFMDPDAQILRALYLAQFSVYPIVTNDRR--LCFEARRTLSSRWPG
28 RDFVVARSRTPLLDSRSGLEMLSRHLRTLVEGSDSHLWWDPRLTVELFPSLSVTDGER
29 F
30 LRRRTSDSDWSESLMDITNVQYELVKSQLAPLIQRGDQHWEELAQYETGVVFPESSVQ
31 D
32 LELTPG-----PFQLSWRPRRLREPRRLKVNHLQKIWRLRP-----
33 >GFJG01044014.1
34 -----IRWKPPIPSIGRDGPNVLLS-GDET
35 PTQKLEI--WSLYGDVCRSMFEKRIKPTALRRAGIPWTNKR-----PNTRRLANLTAD
36 QRSHFILR-----QLYP--RMSGAFKASLKKQSLG-----QIKSWFHTLDGLILP
37 LLLEGAED-----LRLIDRVTKHVMENCAQ---NYRAFLDRLKSLRKFRKRWV
38 LSH-----VNTPYKGFPGECHDLAPI---DLWWISLQPH--QGDRRKFIQVT
39 CMCTQ-TRATGLVPARMMQESVKKFEATVSKEG---PTI-----QL
40 DGEVLMETVAACRLVNP-----ETAKLSGGPAACFEKPRSKGQTQAILEICKQ
41 TINYHYEPEGSTLRRVETPTH-E---VNLPDD---ILDYC-----
42 -----I-EQLRMSP-----
43 -----
44 -----ESMKTRVSVVNEPS-KARVITAASVSYQVYMS
45 MVAHMLQPVIKS-PRIISGLRSSRHLWNFLRKTMHQPQNPEWDRLAGQGAI-----
46 FALSTDLEEATDFGNIS-VADQIWKAIVETVPE----S-VKPYFILAHEAY---TRER
47 DL--LMPD-G----RWVKKKRGWLMGDPCFKAILSLAQ-----EYIFRVA---
48 -----TKRWGKVFMSGSLVG-----DDCIML----SRY
49 KEALEDYLQEYTRCGFKISEKDTFITTNAMEYYCEEASL-VPSDSMVDNL-T-VQIKRGMKD
50
51
52
53
54
55
56
57
58
59
60

1
2
3 ----AILYIDVPRLLRMVATR----PDQP-SRHYSQDTGRTELLGRESNWAYR-W----
4 SEKLYLPFRLGTLQRALVPRSVDT---ISPFLPISMGGDGFYA---DPQLTRDLIER
5 FAWSIGEVRYRI---RRVLSGQRNP-YYLRTSNPIGGTDKFHIIMD-MASIEELVPRECR
6 IPVDDMEHCSMLMSFKGTFIMEPIRTAYRALQSRYWESILIEKRIPS-----DISI
7 PHRVTKTGHTEGAVPLSA--QEIAFLK-----LIEDRPKMQEWPFVLVDMR
8 KAYIMDYMHGWKRERDPTPQEEDLYLSQIDLMGASRALYNFITRGCTDSGEPIEIP
9 K
10 VIRDRNLNFFSTDPPVKLEVRRMELK---PESHLILVSRDYR--LAFEIQKMKPETYIT
11 L-----VDPIIHLGFNPWTDD-----GDKTRVIEDSGAIL-WAD
12 LNLFKDG--YPVNRRVELVYSACMNCEPTRYGYSRTADWRLHLNSRIGTYFAY----
13 --RHGG-----WILQKVPRERRPSHEAE-----
14 >GBKL01052498.1
15 -----QKKMSETIRWKPPLSIGRNGPTVVVS-KAST
16 LTEKLKV--YTLETGICVSFFEKRIPRSQVLRLNISWRETE----PESRRLSSSLSPD
17 QKTAVILS---NIWP--KMSPHFITLVRAAPVT-----AIRNWFQTLDGLLIP
18 AIL---FGLDNE-----YGLIDRITKSCLERMAQ---NYSFLREIKEVRKRFRLFF
19 FESD-----QRFGTLRYPGPGSYGLTSIL---EFWRRILTPL--IGNQREFLRAV
20 VTIGQ-TRATGLVDGATLQKSLDKFVRTVTPPE---ETP-----CE
21 VDEV--AVLRLGPLFAR-----ADPSRAKLSGGPASCFSVRSQGGQTGAIVE----
22 ICKRSINF SYGIDLVRTDIPE-HR--ICTSAD----LFSWA-----
23 -----VEELK-----
24 -----I
25 N-----PETSFAVRAVVVNEPG-KGRVVTAASIAYLLVMG
26 VMGHIFQPTIWT- TSSGLKKTRHLWTFLKEGLHTEGLAWEKLGTpv-----
27 LALSTDLEEECTDFGHPR-LSTAVMQVMAKFCPE----V-LRPLVEFATTVY---HSER
28 PI--HLPD-G----SIRT KRRGWLMDPFTKFMLSAMQ-----EYVYQRM----
29 -----LIRWGPQFVTTVVG-----DDLVVL----SSN
30 KPLLEDYLSQYKVAGFKVSILDTMISDQIMYYCEEASQ-IPQKASKALT--VSLKRGEGL
31 -----EFCYLDYPRFRLLIP-----THPDVDRHSYTDTGRVELLNREFSWVK-W----
32 KTGGEEAFERA VLLQKII VPRSSET---SHPFLPVWLGGDGA YHP---DPGYVFEYVRS
33 HSWNVGEYAWRL---QRYLEDTVEV-AYTRTRS RSGPSVHRYHNIMA-LAEIDGLVP-EEL
34 RFPVVDEEREMWLAFKGTHILT PPEAAFKFLKARYWLT VLEGEIPE-----ETKI
35 SKVIPLGNVAPELTR---KSISELLD-----RMTERKPRL EEWKWLVDMS
36 NAICADYMTLSWERLKPNPGE-----EWKLFTEGLLDMHSAGMSEIYHAITHGDIDIPQ
37 GILSRMNVFFSSDPLIYRLRKDVPD----EENIALISRDFK--LACKIQQMFPRKGVA
38 LVDPL-----LYVLGMIPEEIDSRFK-----IIEDPGA IY-WTD
39 INCFQDG-----APREQWSDTVEDVYSGRLAQR
40 WHPKPGHRILNRRIGVYFAWTSKRRQGGGY-----
41 >GBKL01051391.1
42 -----FFFLKKKMSETIRWKPPLSIGRNGPTVVVS-KAST
43 LTEKLKV--YTLETGICVSFFEKRIPRSQVLRLNISWRETE----PESRRLSSSLSPD
44 QKTAVI-----LSNIWPKMSPHFITLVRAAPVT-----AIRNWFQTLDGLLIP
45 AIL---FGLDNE-----YGLIDRITKSCLERMAQ---NYSFLREIKEVRKRFRLFF
46 FESD-----QRFGTLRYPGPGSYGLTSIL---EFWRRILTPL--IGNQREFLRAV
47 VTIGQ-TRATGLVDGATLQKSLDKFVRTVTPPE---ETP-----CE
48 VDEV--AVLRLGPLFAR-----ADPSRAKLSGGPASCFSVRSQGGQTGAIVE----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 ICKRSINF SYGIDLVRTDIPE-HR--ICTSAD----LFSWA-----
4 -----VEELK-----
5 -----I
6 N-----PETSFAVRAVVVNEPG-KGRVVTAASIAYLLVMG
7 VMGHIFQPTIWT ER-TSSGLKKTRHLWTFLKEGLHTEGLAWEKLGT PV-----
8 LALSTDLEE CTDFGHPR-LSTAVMQVMAKFCPE-----VLRPLVEFATTVY---HSER
9 PI--HLPD-G----SIRT KRRGWLMGD PFTKFMLSAMQ-----EYVYQRM---
10 -----LIRWGPQFVTTVVG-----DDL VVL-----SSN
11 KPLLEDYLSQYKVAGFKVSILD TMISDQIMYYCEEASQ-IPQKASKALT--VSLKRGE G-
12 ---EFCYLDYPRF RLLIP-----THPDVDRHSYTDTGRVELLNREFSWVVK-W---
13 KTGGEEAFERA VLLQKII VPRSSET---SHPFLPVWLGGDGA YHP---DPGYVFEYVRS
14 HSWNVGEYAWRL--QR YLEDTVEV-AYTRTSRGPSVHRYHNIMA-LAEIDGLVP-EEL
15 RFPVVDEEREMWLAFKG THILTPPEAAFKFLKARYWLT VLEGEIPE-----ETKI
16 SKVIPLGNVAPELTR---KSISELLD-----RMTERKPRLEEWKWLVDMS
17 NAICADYMTLSWERLKPNPGE-----EWKLFTEGLLDMHSAGMSEIYHAITHGDIDIPQ
18 GILSRMN VFFSDPLIYRLRKDV PD---EENIALISRDFK--LACKIQQM FPRKGVA
19 LVDPL-----LYVLGM IPEEIDS RFK-----IIEDPGAIY-WTD
20 INCF QDG-----APREQWS DTVEDVYSGRLAQR
21 WHPKPGHRILNRRIGVYFAWTSKRRQGGGY-----
22 >NC_032404.1
23 -----MTNSCNNIPINMLLKEYNLQV-----
24 -----GPCVSIFTKRVPIWWLKR SNIRFYH-----DKVPFIKNLQPD
25 DRVRYLLSSIWT PRDRPLSFAFR AKLKE LSTKGSIN-----TIESW WNTVQGCCLP
26 IMVGTNLWTDQI-----PNWCNRITKWA FESCSR--NYSLFQSHIKAERKLMRQYF
27 METG-----ISTLEG YKSSRREVVTYVSVSQEV RLNLSTV-----REREVM
28 IGLLTQT RATGLADQR MINASLEKFLSVVSHPCA--HIT-----LN
29 QKVMDSMVPRISGYIMG-----YTKLSSGPSAVFSHPRT RGQTAIAE---
30 LSQRYITYSYDLHT LEKTTVERTR--IDNARK----VLDYC-----
31 -----IHRAIHRR-----
32 -----TEVLKRVSTVVEPG-KARVITVGTLEYVQILS
33 VIAHIFKESANSQR-TRSGMKASRHLWNFLWKDLDPRNSIWETISGETI-----
34 YGLSTDLETATDYGNMS-MSGQIW DLLIQAATR-VP-GV-PIKLLHLGKELY---QS DR
35 PL--FSAG-K----LIATKKRGWLMGDPMTKILFTLIQ-----EVISR SV---
36 -----GRTY-RIAC-SIVG-----DDL VVL-----SRE
37 RSGLKEYLRLLETLDLKISWKDT CISDQYI YYCEELAR-IPQSAGVSL--RSLARKNK D
38 YI----GYIDVPKLRTLIP-----TRGESDAFSNTLGRFSLQGKSSEYTSRVN---
39 -TGYYPVNER ALLQCMFIAEDTSC---IHPHIPIEMGGGGSLVN---DPFYVQRSIS
40 LCKYQREITYRI---RSILSHNGWGFRFVHGDRPDRVTHKYH RFGKIIQDPRVFPE-GCI
41 IKPRNKAQSELLASVRSSKII PPVDAFMRLQRMYYFRGLFS--GTGEVPDPDEYLKSLEI
42 PDGITKPSIGYTKLNP---LTLDEVSTFLSL-----WVKRGLA FRDDIPFWINRE
43 KLDINHPLVIELFKKPDM PWT-----TEDLIESEKLSLRLA---ENPWEIPD
44 TLTPALSLVLDQDAILINILERKIKK-ITNIGR-LVVVTLDVR--LLIRMNRILRKRGKS
45 LDRIILIDPFLNLLGMAPDFNPDIYLEDQGGITYVEDMYFQDGMPVPEVEEHMFSD-PVY
46 FKLCKPD--LPI-----YIAASVIHTK
47 RFPSPK-----YRACGANLLMS-----

1
2
3 >GEMJ01010140.1
4
5 -----ENILTDXWRPPLPLTGRFGTRKII--ENLN
6 DRSNSYLLKQYFLSSTSCTNIFQKRIPVFWLKWNNGIPFKH-----TYVPYLRLDLPKY
7 YQSKFLIESLY---PRMSPRNFCNRQLQKLCYEGKLH-----IVEQWFNTAEGCVLP
8 LFL---LPDLWKDVN---LKLIDRVTKWSLEWCAC---NYAHFISILKSFKKAIRKYF
9 ALHG-----NLDGFTHPEDRTMVPYL---DLVRSLLKDD---TLERKLQVI
10 LTCTQ-TRATGLSNHQMMMNSLKKFKATVTPPVH--IEM-----DQ
11 LLLSEILNKRVLRFSEP-----LSRISAGPSAVQGYSRKSGGQTAKALAE---
12 LSTKYIDYDYLRTLTIKEER-VQRRISSAAD---VLNYS-----
13
14 -----
15 IHQALTER-----
16 -----KRVSTVTVTLVNEPS-KARTITVATLESVTLLN
17 VAMHIFAPTCNSQP-VRSGMRANRHLWQLSWKDLHPSNTLWDKLDHKHNDT-GY---
18 NPL
19 WALSSDLESATDYGNPS-VAKQVWRKLIDILSV-FP-DA-PIGLLELCCEIH---TGPR
20 DV--LYEG-E----PFCVTQRGWLMDPMTKVILTIQ-----EYVLQF---
21 -----QQLIP-----AVG--SVVG-----DDLIIL---SYS
22 RAHLEMYPSTMLEALDFKVSMEDTFVSKRLMFFCEEGLSI-
23 VPQTVRDTIVCRVKRPTWKKG
24 YKPISCGYIDYPRIRLLIP-----VRPDQDRLSYTLGKFSLLGREVIWAHSIN---
25 -EPLFPVMERAVVLQNIILPQDSDT---LSPFLPIEIGGNGSYCP---DPDFLKSVEK
26 RSRSASQCRYRM---CQLLNHE-YGFRYLRSERRNKVLHKNVVVARLVEDESVFPP-DSV
27 IRAESDEHKTILSSL--KAYISPEEFMKLEYQKYFTQVLS--GTGDIPS-----NFEY
28 QRDHSLPT---NLGGVDVEPNYSKFLKI-----WHKPGFTFKNSWKFVVNPK
29 KLGHNPDPLSVDLKFKPDLQVK-----PDPGDL
30 LGRTLEDFLNRVQAwdMYFPDHILLER-LE----LVVESDSLL--VALLKARLQRGETAQ
31 T-----LCLVSNDKLLTRLVTIWDQLRPGECPLEVVLVHPDVYLTGMYS-WQE
32 YTMTPPI--
33 MVLKDPMGSFWESNNFNNAACQYWEVCWEVDKYQPLPGVATFTTLFYLPGPR
34 RHPQDD-----PVQEITPGRVFPHMGIQDVETSGADTPQEHIPEYVPPHLRK
35 >NC_035120.1
36 -----MPPLPLIGRLGPRMAVSSDLLTLGDEKISSVPL-EVLL
37 ERYFTSF-----GGCANLFYKKVPISFLRKNSVRFHH-----EEVPYLSNLSRA
38 MQFRYLRKSVWP--QMSP--VFYQRPLYRAVSAGRKT-----TIESWFNTAEAMCLP
39 FFL---LGVDSHTE-----YSCVDHITRWSLANCAN--GYASFIFTKTFKKCLRKHL
40 FSGE-----PIDTRHGSSRTCSPYL---KWAKVVLGSIPEQQTA-LRAQAV
41 ALLTQ-TRASGLADHRMMMQSITFKSVVTPYEH--IQM-----DE
42 PVLSRMIHPRMRCVLN-----NAKLGTGPSAIRGFPRSSGGQNLAALK----
43 TCTMTIDYDYLRSLSVTSTENRR--ISNSRD---VLDYC-----
44
45 -----ISQALNN-----
46 -----RQKVTEVTLSVVNEPS-KARTITVGDYALVQLLN
47 VAAHIFKDVCCCTQP-VRSGMRADRHLFNVWKDLHPQNTLWDDMGWSYETK-GM---PI
48 HALSSDLETATDYANPS-VGRQIWDCCLISGLEIQYPESS-PRALLELCRDLH---VGPR
49 TV--YYQK-I----FFCTKLRGWLMDPMTKVILTLAQ-----EYVLFRSNAGR
50 -----G-PTGR-LVG--SIVG-----DDLVIL---SRL
51
52
53
54
55
56
57
58
59
60

1
2
3 RHHLGWYLDLRSLDFRVSDDDTFISSDFMFYCEETSR-VPQGPGQSVVA-RTKYSHGT-
4 ---SCGYIDTPRIRLLIP-TR---PDE--DRFSNTNLGRFSLLKEYQWCLGNN---
5 -SDLAPLFRRRAIGYQNCLVPQDADT---QCPFMPVEAGGNGSYTT---DTAFWSKLVLR
6 RSKDPRTTHFRV--NQLLSNEYAY-RWIRSDRPLRGLAHRYSVVRQVFEDETRPEGSV
7 VKAEHEEERMILRSF--KQFLSPETAFMKLEQARWFNGLFDGRGVIPP-----NMQY
8 DREILLPSVGQSGPET---LDPPVIRRFL--ER-----WRNPGFTYRDDYRFLVRRD
9 FLPNEDPLFIEVDNEPDKLS-----ASVVPDDL
10 IGKTFEEAVNKILTADLYLDDNLWSR-LPLLIESDSMVMARWR--ELLRRDTYVETLVLI
11 STDRKLLR---MIDLFTVMRPDGYIVE-----AILMHPALFLTGTGYM-DYD
12 QYNGIPG-
13 RVVIQDPGSIFHQDLYLPQRRSWWEVTVMLEPFRTQEREGVTTIYHGYMEVR
14 PELDRK-----PLPRSNASRIWRHRE-----
15 >GCRV01019573.1
16 -----RHVGCAIMQSADR_SWAPPLPSLGGKVGPIMVNPNLS
17 LSDKLKA---YYLELSSSPSMWIRRVPYATLRAVGYS-----GAKHLVDPWL
18 YQMKVLLKTLF--KASR--HFVAKLESACRAGRIN-----HIKQWWITADA VVLP
19 YLL---ELKENDAD---YQVVDNLTRFALESCAN--NYSGFIGRLKKVKKLIRKSL
20 ALDQ-----EIPFMRDMDYPIKQYKASSPSEGGIHAMA VHC
21 NLWCQ-TRGAGLADSKMVKAETKFLDCIQRPK---VSI-----KL
22 DRQILKDTTLCCRAANG-----VHAHLSAGPKACFESTREEGGQTAFLSKL---
23 ARTRVVTHSYNYETLERVPLR-RR--VGSPKD---LLDYC-----
24 -----
25 -----IDWSL-----N
26 E-----PVLRKVVRLLHTVAEPA-KARVITVAPFAYNRLMG
27 VCAHMIAPTLRSTP-LRSGMTASRHLWGFTHDTLHPGDKLWEQGI-----GC-----
28 EALSTDWEVATDNGNVS-VARQVWSEIIRHSK-ID-GA-PLGFLNLCKTLY---CEDR
29 YV--LHKG-----KIHRKTRGWFMDMMTKVIITLVH-----DYCARV
30 ALL-----KCY--SLVG-----DDFVAM----GY
31 NENLQAYLQEVARLGIVISDDDTYISPRFMYYCEEMCL-VPHTT-----SQLPVVQIA
32 MGKSKISYIDTARLRLMLP-----CKQENQIFSATQVGRYSLLGKETRWALSTH---
33 -REVAPIFERGVVLQHLLLPRDQDT---QCPFTPIELGGDGSYTS---DSNFFKNVIEK
34 KSKNPAETHYRI---ASLMRGA-FGYRYICTERVDQVSHKYRMYVPLFQELVKYLP-EEA
35 VIPIHGNRTLTLR-ARLLETPQRTILRIVKEAYYREFK--GVPVAKLP----TLRL
36 EA PPPMPSTGGATPFY---SASFLEH-----WRNPGFKFRDQDPYLVRTD
37 IALTLDHMSLGWDFSIRGPPK--RALEDWLSQNLSLFDGNANEVLAKL---LSNEDLPE
38 GAKSRLNLFFESDGWVLNDL--SR-AEIPPGPIYLISNDIR--LGADVSRVMEGKQDP
39 R-----EVYCVRPILSIIGRMD-----EVEDLVPAPALSL-EDP
40 GSFSFET--LVT-----LGEDFPEWAFNPLL
41 KTKFSG-----VFIRVPEIYAD-----
42 >NC_032457.1
43 -----
44 -----MSG--HFQRALSKLCRSYDYL-----TIQRWWYTVDGLVLP
45 LLL--THPDPSQ-----METIDRLIKFGFENCAN--NYSHFISRMKRLKKLLRKEF
46 AEGTD-----LTSGQLFVDLSTYRKVFY----RYLDLTLSER-----SATECV
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 MAWTQ-
4 ARAAGLADRRMIEKSLEKAETLSTPRRVYLRGRWREIHAHYQGFIPVLEPEVP
5 RGPLLDVVRRVANKVRG-----ETAHISTGPHACLEIKQAFFGQAAGIAY---
6 LCKRDVLTVSYDPVTLERTFH-PARPCASAQD----LLSWA-----
7 -----ISRALDAP-----
8 -----
9 -----
10 -----VHTRMVRIVAVSEPS-KARTVTVGALAYQIILG
11 VVSKIFQPALASGI-SRTGLEGTRNLYEALNNDFDPSNGLWGPLRMDRT--GK---FPV
12 FALTSDLEEATDYGDLA-VARQILQALLMRCRD-IP-GF-PLGLAVLAKSLF---LSSR
13 II--IRPG-YGHGLVHWFRKRNGWLMGDRMTKVVLTLAH-----EIGIL-----
14 -----SAGI-QFA--RICG-----DDVFAL---SKA
15 PAQLQRYHKVMTDLGFKISEDDHFISRRILFYCEEVSL-VPQDAK-DLPS-VCNRRSE--
16 ---PSCYVDYPRIRLLP-----IKVETNALSYTDTGRFHLLGKEMRWVFQNS---S
17 QQADPFIRASLLQHIAIAMPRDILS-----PFLPQELAGDGAFPH---SAEFLLSVIER
18 KSINYDECLYRI---HSLTHGKWGF-RYLRADNINEVVHKYHQMPKLVLESVLP---
19 PDAVVRASEVLISSLKVKGLETPEKTFRLYRSYYWYKVLHGMKAPV-----
20 -----LNFDEDRT-----RVHGGREFKLDIP-----
21 -----
22 -----
23 -----
24 -----PDVLVERFYT-TWRD-----
25 -----SGFTF---
26 -----
27 -----
28 ----->GFUS01019776.1
29 -----NHVRWSPTLPSYGGRLCPRVLI-PANL
30 PDFASKL-EAYYLHREGSSSIFVRRVDVGWLRKHGHERLI-----AQGPFLADLPGT
31 VLYKYLIASL GSL-TMSA--HFQRSQLCLAGKM-----AIKRFQTTDGVVLP
32 LLL-SVEDPDQH-----YADIDRLTKFSLENCAN--NYSHFIGRLKRVRRLLRKDY
33 AEMG-----TSAKSLTDMMTYRGM-----YLKHAPKNRDQYRDVQYL
34 MCWSQ-TRATGLADERMIRKSLDKAEITLSTPRRPRRSIPREDWVQQSTK-LPLGEPVID
35 RALLDSVCERIYRRCDG-----RTGHL SAGPNACIESPRNKGGQTGAISRL---
36 TRMKVLHSTYDFKTLEVTRHPDRR--VGSVQD---LLSWC-----
37 -----
38 -----
39 -----
40 -----IQESLEHS-----
41 -----TYVKCVKLHAVAEPS-KARTITVCTLPYLLIVG
42 SLAKLLQPSIASDI-TRGGLESSRNFWNFLYQDLDPTADVWHRLRDEKAED-GD---GKT
43 YALSSDLEEATDYGDIG-VARQILDGILRSASR-IP-GF-PVGLALLAKTLF---LGKR
44 FC--FRSE-NGK-W-RYFVKRSGWLMGDRLTKVILTLAH-----ELAIR-----
45 -----ASGL-HIA--RICG-----DDVTAF---SKS
46 RSRLARYLVVIREYGFVSEDDSYISSRLAFYCEEAA--IPPVRFLDLPS-VCTRRT--
47 ---VPCYVDYPRIRLLP-----TSSEVKAYTYTDVGRFAQLGKEMRWILQGE---
48 IHPARDVYVRASMLQHLLVPQPADT---LCPFTPVEVGGDGAFPH---SENFFLEVLA
49 KARNPTEIRFRL---HQLHRGVWGY-RLVRHQYLDNVIHKHHMLVPKIEELRSQIR-ECV
50 PDALIEGDELMLNSISVSGLETPERTFFRMWRAWYWSEIFRGHDPEI-----TLEI
51 SRECNLPR---SEVTL---DALGFYRV-----WRGKGFTFQNRVDYLVITK
52 MVQPLDYMNLGWNFGVRDLPT---ENPDADQGRFLLNEENSSSFIEYI---RTGRPMDQ
53 VVRDSLHRMVESDSFLKWEFS---RRCHADIPQLILLVSTDVN--LGIDFQRLAATRRPD
54
55
56
57
58
59
60

1
2
3 -----YEDAPLVFLIKPTDYLGYGFMSHYESWAADLTQHWDVIVDFGCVV-YNQ
4 FTLDSPE-----FEWFEDPDVVHNPVRDASLRVYYLSLKEGE
5 GGSERC-----SLTTQPASQWRKGFFPSQYVREGLLGN-----
6 >GFOW01018115.1
7 -----SIMSQKRWSPTLPSFGGRLCPRVYL-PEGY
8 PSDYCSRLEQFYLERSSSSIFVKKVDVTWLAKGHQDQLFLDGR---FRAPSLGDLEWT
9 VLYKYLMSGISQ--TMSL--NFQRKLS DLCRYGKIG-----MIKRWYQTV DGLVLP
10 LLL-SVEDPDLH-----YAEIDKLICKFENCAN---NYSHFISRMKSLKRLRKDY
11 ALDT-----ESSKVLSDVSTYR-----KIYLRFFPAGRGVIPA-DAIRYV
12 LIWSQ-TRASGLADD RMCHQSLVKCEKTLSTPRRARRIAPPEERRYSYRG-LPLGEPVVP
13 KIHL SRICERVARRLRG-----ETAHLSAGPNACLESTRQEGGQTGYISRV---
14 CRSKSLHAEYDFRTLECKQCTPRP--VRSQDD---LLSWA-----
15 -----
16 -----IQEVL EH-----
17 -----PTYVRCVRLHSVAEPS-KARTITVCTL PYLLIVG
18 VLSKLLQPAIASDI-TRAGLHASRNLWKFLYEDLDPGADLWFTLRDQFKVR-GQISNLPI
19 LALSSDLSEATDYGDLG-VARQILAGLLHSCSE-VP-GF-PVGLGQLAKTLF---IGKR
20 FC--LRCE-NGR-Y-RFFVKRNGWLMGDRLTKVILTLSH-----EIAL-----
21 -----R-ESGL-YIA--RICG-----DDVIAF---SRS
22 RDDLERYHQVVTDLGFKVSEEDHYISSKLA FYCEEACI-PP TYR-ELPQ-VAIRRKTG-
23 ----SIYIDYPRIRLLIP-----TCVETHAHSYTDIGRF SLLGRETRWVYQNG----
24 APHVAALFKRASLLQHIIHV PRTADT---MCPFTPLQIGGDDSF PF--SGRFLEAVVSK
25 KARNESETRFRM---SSLMTGKWGY-RLVR NANLDNVVHKYHTLVPKLKE RNLLP---
26 EHAILEGNQFLLNSIKIDGVESP RTFFRIWRAWYWHQILKG SNPPEL-----VLGV
27 RKEANLPQ---TKVN----IDFEYFFHI-----WKDLGFTFENVPDYL VSTR
28 DVHVLDYMNMGWWFGTPVGK-TTPDIFDIGGRFLLNEENLESFINHV---KDNRELPQ
29 VVRSTIYKYVESDSYIKWEFS-KRDD-LDSLTRIIFLVSSDLK--LGAE LARMARSRYSS
30 N-----LADAPVVVVVPCWVYS-----YGMMDIPIDWST
31 NSMSCPD--PIVIVDQ GSLTFDQM VNSPEF SFFDKP RVTRWYRSNPWL GTQCL DIVE DG
32 TTPSEV-----WSLTTQHPSQWRPGHFPYVPLRT-----
33 >GFRX01320111.1
34 -----FFKKKTSHMMTQNRWAPTLPSFGGRWAPRVAI-EAEY
35 SASYLSRLERYFLRRQSCLSMF SHRVEAHALREC FGKDILTGYMVT LGVTPR LEQ LLD
36 D
37 ELYRYLVKSFG---KISFSKNFLRKLKSLCYSKDLA-----TIKRW WQTV DGLVYP
38 LLL-SVEDPDLY-----YAEIDHLIKF GLEN CAN---NYAHFISRMKSLKRLRKDY
39 AEMK-----DSAKALTDVSTYR-----KIMMKFFPKD--EANPHLT VQYL
40 LCWTQ-TRASGLADRHMIRKSLDKCEATLSEVRRIARIST-PETREYSYRGLPEGEPELP
41 QDILNDIAWRVSDKVKG-----ETAHLSAGPNACIESTRQGGQTGYIKHLA--
42 RMKDSLHTRYNFCTLEPEPITPSP--VRSQND---LLSWA-----
43 -----
44 -----I-QEALEH-----
45 -----PTYVRCVRLHAVPEPS-KARTITVCSLPYLLIVG
46 VMAKLLQPAIASDI-SRGGLHASRNLWNFLNDLPSLDLWG YL RDDGDPKVPI-----
47 MGLSSDLSEATDYGDIG- VARQIFSRVLEKCSS-VP-CF-PTALGNLAKTLF---IGRR
48 YC--LRYH-D-GKGYRFFVKRNGWLMGDRITKVILTLAH-----EIVVR-----
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 -----SSGL-FCA--RICG-----DDVIAF----SRR
4 REDLERYLSEVTRLGFKVSWEDSYISSRLAFYCEEACI-PLQRFR-EIPS-VCTRSGD--
5 ----QSMYVDYPRIRLLIP-----TRVETQAYSFTDTGRFSLLKESRWVHQNK---
6 LNSQTKLFRASLLQHLHVSQSPDT---LCPYLPIELGGDGAFFP--SHTFLMEVVRA
7 KARDLRETHFRM---LQLRSKWGF-RFVRSMNLNDVVHKYHLLVPKIKELRELIP---
8 PESIIEGNELLGSIKIAGVESPEKTFFRIWRAWYWHISILQTGREPP-----KLEL
9 GVSRRHFPREETEVE-----MTEEHFHQFFFH-----WKDSGFTFRDTPDYLVDTR
10 YIHVKDYMNMNWFG GIRPAWVS YRPTMEQVGGRMLLNETNMESFIDHI---
11 KDGRPLNP
12 IVGRTLHKYVEADSYILSECARNWSR-
13 LNPPDMCVLVSGDKRLACELARIQRAYYVPERD
14 I-----YDKVPWVMMVDPNHYLFGMMDPYEDYAEAYSSRFEVVVDFGAVA-YSA
15 YLTGGVD-----MFYENPRI DT KYLWNNPQLGVFEASLDEKS
16 ATSKEVLVSLTTQPRDQVLKRRIPWNFGSGLRGATMSP-----
17 >GFJG01122116.1
18 -----DRPNNKLGEAPGPNDHTGKRPRS DT QEMSDAPPVKRPKLA VR RLLYNVVE
19 SVIKTAK-RTKIS PPEPQLPDVAQHTEESQYKCDISV VENNVLREIC TR LAPYEYKGIVK
20 REFRNHPA HRRWYFMIGR--RLTDRTNQFVQSLTFR-----DYSTLEQIWRAVMHT
21 VLLNPDFDPSRFDPDL---IESVRSFKLWCITYASSHKNGIVKLGKALKTLASWMQWFM
22 LDNQANTCPNRVEIVPGWNRSTKRPTFPWFSGHLSHCRDFNRVKSD-----RDLHLL
23 AQMRTFGRALPCPPKILVDEGLKATLDV LATER-----KT
24 PDHVLEC ISSAVSYLN T--SWKISNMVRESHLSISVSASYDHTQEKGGSATQRE---
25 ILSPFLDH KVDQLLDGNVPIKGNRLFIDCFGNIMLHGFSMGRSIRT CRLVDGLYATHATV
26 YGREEISKNLGKPMFGRR T-----GNVLLASI-MKVQNYG-----
27 -----HFEPSPKGFLRTDDNQINIPYWDE-----SIIDEGLPSF
28 I-----EEKPFP CRFIALAE PGFKVRPLTIGYTPMMNVLK
29 NMRFMMQS VLLHD PRLRVGLRSTNKLWT VLK-----LLKQKKLEYGE--GL-----
30 VLQSSDYKSATDYIPLD-VIRTIWQTC DGLTK-----AHPF-RVFRSLI---WHPR
31 QIWADIPDDQGKYTRQQHRCGS FMGE PMSYMTLCILN-----L AVELLSQYRLRTLS
32 LTLP IEYDGLNV LH-GRGI-NLDVII VVG-----DDVLA F----RRD
33 MGIIHIFHEVATQCGFIFSEGKDGYSSRIGIICEDHLL-VDRREN---P-VTIR-----
34 -----FLDIVKSRLLT NVAR---TH---GDNRASVFGKGRMLNMVLDYLPNTM--FK
35 EHWL AYRSV LERTIRLKLSDDL-----PWFLPPSLGGLGLQLS-----GDIPDS
36 GWKFICFLEAR M---SSAASWQDQL-ALVWSFRHL-SNRQPHGWEPEQGT VEMA VE---
37 -----CLKEL--KHAPKEQSLDPNLEKVIWSTDHRRFLRISGAAAPFVGSAVQL
38 MHDHEY YRTEAEKLGL---VPLADVID-----QFERLTTFQS ALE-----
39 -----
40 --GTAEHRFKTIQQWIRRSKK-YWPKNMMNVRKPLWFTNLSDF--ERRMIRRF DGF NIID
41 -----LCSRSV NYGP NLRIVY-----DVWSQRR RGM-PRI
42 GLGGNPL-SLKT-----
43 -----
44 -----
45 -----
46 -----
47 -----
48 -----
49 -----
50 -----
51 >GEWO01002579.1
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3
4
5 -----YSRSKGGKFSEFRD-----
6 RVVDEFTFTPINEIVGHMPKI-D---LIDVTGNVFLRAEDWVYGTIRDCAYPDLLRDEG
7 ASAD-----TRIGLFGLL-WAVLEVL-----
8 ---KDPDIETCTNDMAWPGLGSGISPQIF-----
9 L-----DEGTVRAKVVSLPEPGYKARVITIVPLAISIIGS
10 VARHLLDDMVQSDEMTKLGLSKDKLYSFLRKMNNGKISGTADVSDSCKFA-----
11 YAESVDLTTATDTPPLS-SVRSLLQGTIATYLY---DKHAGADFCQFAASLS---TLPR
12 TF--VLPG-GKVT--NPNDHRSGIMMGEGLSGTFLNVMSAVRFLLRPTIDVFSDYDGTT
13 AVDA----DDFIGC-NRDR-IQE--FLDGRERPMFEPDPTSTQSGDDVLF---DNLAP
14 GMARRYLVLLYRIFGLLPSGTTFFSSEFYGTFCEEVIV-RTHDSR-GWT-----
15 -----YVDTTKSKLF-----QTTS-VEGPEKLLGRIGAIASSLKYSGDEA---
16 LLRRAERVDAMISGHPTLSRILR--YDPLPLGLPECLGGIPHNG----SLSDFVAC
17 LDLVSLEILTGA---STVPDAVLEIEFSRFFDDLPSDQ---IHLVEDTVRNVLE---
18 -----VIGGW-----ESPSFVDSTIPFQVFRENFLPQG-----DLKY
19 EVWVEVAEAVSRRHGL---ISIRSAIE-----KISSMIRASIFLS-----
20 -----
21 -----GGVPPVNPLIRRNR--RQK-LL-----TRLR--QHSQSIGAVDNSTPR
22 T-----VVKRLRTCENRYFR-----FQDIS-DIF
23 ELHGYPs--FVV-----
24 --PFSL-----S-----
25 >GEWL01001678.1
26 -----LSDEFVPLGVWQVRP-SRQC
27 GSDEL-----LPSCMAFEGDVLQSWRAVEAALLL-----AGASAPHCSEVH
28 ELLLR-----TFVR--GGIRAFAL-----TLKGWGKAVRGLLS-
29 -----LEP-----LDTHLVGYTRLPKNCVCALRREGMGESDIRDF-----
30 -----LDFNGKLGPIPFLF-----SEFNKVLLPG---RRLSDIQRI
31 ASLTTLGRSAPPATQEVLRGAVDAFKNVTTTH-----AR
32 TFDLFTIGWEFGQYAKG-----KSPPVRSRIGLNTHSSYDYSRSLGGKFAEFRD---
33 RVIDDFLFWPIEEVVERKPNH-D---LIDITGNVFIRREDWSAGTTLRDCAYPLLERDEG
34 SSAD-----SRIGLLGLL-WAVLEVL-----
35 ---KDPNIVAETVGMVQPGLGSRDPPQIK-----
36 -----LGGFVPAKVVSLPEPGFKARVITVVPLAVSIIGS
37 AARHLLDNMVWADDMSRIGLSSKNLYDFMR-----VLFNGRIEGNATD--GVSCNHLP
38 SVESVDLTTATDTPSLQ-AVGSSLGATNGYLE---DSIAFEFSQMAVSLA---TSSR
39 EF--TLPG-GRKL--SYLLHESGIMMGEGLSGTFLNVMSLVRFLLPEAVIAFWDYKGNS
40 VAEA----DVFIDL-NSGW-IQL--FLDGQGRARFEPDPHSTQSGDDVILF---NDLRP
41 GEARRYLVLLYRIFGLKPSETTFYSSGTGTFCEEVCV-KTKSSN-GWT-----
42 -----FVDTVKARLF-----QVTS-TEGPEKILSRISAISDALKYNRDDT---
43 LILRAAEAVDAMISTHPTLSRIFR--YGLPLGLPECLGGIPHIG----LASDFVAC
44 LDEISLEILTGA---STVPDAVLEIEFSQYFDDLPDAQ---VELIRKLIQNLLG---
45 -----TIDGW-----ERPSGLVTAEPFQVYTRERILPQG-----DLQY
46 STWVGIATAISERHGL---SSVRTVVE-----KISAQVRASVFLS-----
47 -----
48 -----GGSVPPVNPLIRRNR--RDR-LL-----KKLR--EHTTSIGGVDLSTPK

1
2
3 T-----VVERLRSCEDRYIR-----FSDVQ-DVF
4 ELYGLPS--FVI-----
5 --PFLS-----S-----
6 >GADD01004427.1
7 -----SPKGPRVYPGRLGRAQMGAAPAASAPGGSPEALI-DFDL
8 GPISWDL-----KRWLPDNPVVAL-----WSVRGKANADDQ
9 RRSFN-----MTNM--AIRRSSLEKFRKLRA-----ERKTLSLTSGLLDN
10 LIA---SNPERIRTART---DPDVRKILTWAFLSLGLY---NVQHAISEWKKFCTLLKHRA
11 LQSET---PWPEPLSDFPFGFRDACPDCPALWARLCPWLIWSQG--ASSK-SDMTRI
12 CHLAT-SRNLPTASAEREKTLHASVLTSTPI---PDI-----ER
13 SSMLRMLSRRIGEMVK----KRKGRRSSGHCSLAQSACIEQPAGKGGRAFYIQS----
14 EFQTWASRKQTETISETTWFG-QP--FRLKEG---LPMWV--TMCRTEDASN-----
15 -----PTMAIGES---RQDFN-----
16 -----MNFDDFRHEDFLAGLDKATGYQLL--QWSLEEGIRRGCLLGDPF--RSSSGLAI
17 S-----KEKLPEIKADIVGEPPGKSRVMTIGEGWLTIFLQ
18 PASHHLAGLIEHHEAAWAGFKRGWQASEWAK-----RYRRKQHWTGED-----
19 YVLSSDLKTATDYCRHD-FSLSLLGFLESGQ----G-DNPYLRAAAELL---CSPR
20 YY--TGPVEG----SIGPTSRGILMGDPGSKIVLTLFN-----LAAEEEAFVRWATPR
21 MIPTISVGGMVTIADS-ATSW-KIRNLAVSG-----DDHTAR-----GP
22 KSYLKGITQAHIANGMEVSESSNFLSPIGGFYCEEALL-FRGLDLQKLRA-YSPFWDIP
23 YL--EHMHVDSLKLRCSPVMK----GDAT-RDEKNPAIGKASTFTGMLRWFAAGGW----
24 ACTRPPIFSARFEQRFKTFLPEGILR-----YLPRELGGVAAPCF---HLTAGMLESS
25 LTKLDPSVRSAl---NLVLAGERG--ELAKLLAKIPANTSMRGIPE--NLARDQVT----
26 -----QILETF---PGVQGRIDESQLKAMADDQMGD-QWWG-----DATA
27 RDRARFAR---RNN-L---ISVVDALN-----DIERPYLFREFLA-----
28 -----PGQHAVHE
29 GPSQFKTRIYDTTPWPVRVNR--FRS-LA----LSAKPDLL--HHHSAEGVSDGKLVG
30 M-----IQKGKGLPPPPRMYIP-----ERVIFGDDL
31 VTLQTPL-----
32 -----
33 >GFOL01002916.1
34 -----
35 -----
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----AISEWKKFCTLLKHRA
43 LQSET---PWPEPLSDFPFGFRDACPDCPALWARLCPWLIWSQG--ASSK-SDMTRI
44 CHLAT-SRNLPTASAEREKTLHASVLTSTPI---PDI-----ER
45 SSMLRMLSRRIGEMVK----KRKGRRSSGHCSLAQSACIEQPAGKGGRAFYIQS----
46 EFQTWASRKQTETISETTWFG-QP--FRLKEG---LPMWV--TMCRTEDASN-----
47 -----PTMAIGES---RQDFN-----
48 -----MNFDDFRHEDFLAGLDKATGYQLL--QWSLEEGIRRGCLLGDPF--RSSSGLAI
49 S-----KEKLPEIKADIVGEPPGKSRVMTIGEGWLTIFLQ
50 PASHHLAGLIEHHEAAWAGFKRGWQASEWAK-----RYRRKQHWTGED-----
51 YVLSSDLKTATDYCRHD-FSLSLLGFLESGQ----G-DNPYLRAAAELL---CSPR
52 YY--TGPVEG----SIGPTSRGILMGDPGSKIVLTLFN-----LAAEEEAFVRWATPR
53 MIPTISVGGMVTIADP-ATSW-KIRNLAVSG-----DDHTAR-----GP
54
55
56
57
58
59
60

1
2
3 KSYLKGITQAHIANGMEVSESSNFLSPIGGFYCEEALL-FRGLDLQKLRA-YSVPFWDIP
4 YL--EHMHVDSLKLRCLSPVMK----GDAT-RDEKNPAIGKASTFTGMLRWFAAGGW----
5 ACTRPIFSARFEQRFKTFLPEGILR-----YLPRELGGVAAPCF---HTAGMLESS
6 LTKLDPSVRSAI--NLVLAGERG--ELAKLLAKIPANTSMRGIPE--NLARDQVT---
7 -----QILETF---PGVQGRIDESQLKAMADDQMGD-QWWG-----DATA
8 RDRARFAR--RNN-L---ISVVDALN-----DIERPYLFREFLA-----
9 -----PGQHAVHE
10 GPSQFKTRIYDTTPWPVRVNR--FRS-LA----LSAKPDLLL--HHHSAEGVSDGKLVG
11 M-----IQKGKGLPPPRMYIP-----ERVIFGDDL
12 VTLQTPL-----
13 -----
14 -----
15 -----
16 >GFOL01000185.1
17 -----
18 -----
19 -----
20 --SFNL-----MTNM--AIRRSSLEKFRKLRHA-----ERKTLSDLTSGLLDN
21 LIA--SNPERIRTART---DPDVRKILTWAFLSLGLY---NVQHAISEWKKFCTLLKHRA
22 LQSET---PWPEPLSDFPFGFRDACPDCPALWARLCWLKIIWSQG--ASSK-SDMTRI
23 CHLAT-SRNLPNASAEREKTLLSHASVLTSTPI---PDI-----ER
24 SSMLRMLSRRIGEMVKG-----KRKGRSSGHCSLAQSACIEQPAGKGGRAFYIQS---
25 EFQTWASRKQTETISSETTWFG-QP--FRLKEG----LPMWV--TMCRTEDASN-----
26 -----PTMAIGES---RQDFN-----
27 -----MNFFDDFRHEDFLAGLDKATGYQLL--QWSLEEGIRRGCLLGDPF--RSSSGLAI
28 S-----KEKLPEIKADIVGEPPGKSRVMTIGEGWLTIFLQ
29 PASHHLAGLIEHHEAAWAGFKRGWQASEWAK-----RYRRKQHWTGED-----
30 YVLSSDLKTATDYCRHD-FSLSLLGFLEGSQ-----G-DNPYLRAAAELL---CSPR
31 YY--TGPVEG-----SIGPTSRGILMGDPGSKIVLTLFN-----LAAEEEAFVRWATPR
32 MIPTISVGMVTIADP-ATSW-KIRNLAVSG-----DDHTAR-----GP
33 KSYLKGITQAHIANGMEVSESSNFLSPIGGFYCEEALL-FRGLDLQKLRA-YSVPFWDIP
34 YL--EHMHVDSLKLRCLSPVMK----GDAT-RDEKNPAIGKASTFTGMLRWFAAGGW----
35 ACTRPIFSARFEQRFKTFLPEGILR-----YLPRELGGVAAPCF---HTAGMLESS
36 LTKLDPSVRSAI--NLVLAGERG--ELAKLLAKIPANTSMRGIPE--NLARDQVT---
37 -----QILETF---PGVQGRIDESQLKAMADDQMGD-QWWG-----DATA
38 RDRARFAR--RNN-L---ISVVDALN-----DIERPYLFREFLA-----
39 -----PGQHAVHE
40 GPSQFKTRIYDTTPWPVRVNR--FRS-LA----LSAKPDLLL--HHHSAEGVSDGKLVG
41 M-----IQKGKGLPPPRMYIP-----ERVIFGDDL
42 VTLQTPL-----
43 -----
44 -----
45 -----
46 -----
47 >GFOL01000028.1
48 -----SPKGPRVYPGRLGRAQMGAPAASAPGGSPEALI-DFDL
49 GPISWDL-----KRWLDPNPVVAL-----WSVRGKANADDQ
50 RRSFNL-----MTNM--AIRRSSLEKFRKLRHA-----ERKTLSDLTSGLLDN
51 LIA--SNPERIRTART---DPDVRKILTWAFLSLGLY---NVQHAISEWKKFCTLLKHRA
52 LQSET---PWPEPLSDFPFGFRDACPDCPALWARLCWLKIIWSQG--ASSK-SDMTRI
53 CHLAT-SRNLPNASAEREKTLLSHASVLTSTPI---PDI-----ER
54 SSMLRMLSRRIGEMVKG-----KRKGRSSGHCSLAQSACIEQPAGKGGRAFYIQS---
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 EFQTWASRKQTETISETTWFG-QP--FRLKEG----LPMWV--TMCRTEDASN-----
4 -----PTMAIGES---RQDFN-----
5 -----MNFDDFRHEDFLAGLDKATGYQLL--QWSLEEGIRRGCLLGDPF--RSSSGLAI
6 S-----KEKLPEIKADIVGEPPGKSRVMTIGEGWLTIFLQ
7 PASHHLAGLIEHHEAAWAGFKRGWQASEWAK-----RYRRKQHWTGED-----
8 YVLSSDLKTATDYCRHD-FSLSLLGFLEGSGQ-----GDNPYLRAAAELL---CSPR
9 YY--TGPVEG----SIGPTSRGILMDPGSKIVLTLFN----LAAEEEAFVRWATPR
10 MIPTISVGVMVTIADP-ATSW-KIRNLAVSG-----DDHTAR-----GP
11 KSYLKGITQAHIANGMEVSESSNFLSPIGGFYCEEALL-FRGLDLQKLRA-YSVFWDIP
12 YL--EHMHVDSLKRCLSPVMK---GDAT-RDEKNPAIGKASTFTGMLRWFAAGGW----
13 ACTRPIFSARFEQRFKTFLPEGILR-----YLPRELGGVAAPCF---HLTAGMLESS
14 LTKLDPSVRSAl---NLVLAGERG--ELAKLLAKIPANTS MRGIPE--NLARDQVT----
15 -----QILETF---PGVQGRIDESQLKAMADDQMGD-QWWG-----DATA
16 RDRARFAR---RNN-L---ISVVDALN-----DIERPYLFREFLA-----
17 -----PGQHAVHE
18 GPSQFKTRIYDTTPWPVRVNR--FRS-LA----LSAKPDLLD--HHHSAEGVSDGKLVG
19 M-----IQKGKGLPPPPRMYIP-----ERVIFGDDL
20 VTLQTPL-----
21 -----
22 -----
23 -----
24 -----
25 >GFOL01002977.1
26 -----PAASAPGGSPEALI-DFDL
27 GPISWDL-----KRWLPDPNVVAL-----WSVRGKANADDQ
28 RRSFNL-----MTNM--AIRRSSLEKFRKLRA-----ERKTLSDLTSGLLDN
29 LIA---SNPERIRTART---DPDVRKILTWAFLSLGLY--NVQHAISEWKKFCTLLKHRA
30 LQSET---PWPEPLSDFPFGFRDACPDCPALWARLCPWLKIIWSQG--ASSK-SDMTRI
31 CHLAT-SRNLPTASAAREKTLSSHASVLTSTPI---PDI-----ER
32 SSMLRMLSRRIGEMVKG----KRKGRSSGHCSLAQSACIEQPAGKGGRAFYIQS---
33 EFQTWASRKQTETISETTWFG-QP--FRLKEG----LPMWV--TMCRTEDASN-----
34 -----PTMAIGES---RQDFN-----
35 -----MNFDDFRHEDFLAGLDKATGYQLL--QWSLEEGIRRGCLLGDPF--RSSSGLAI
36 S-----KEKLPEIKADIVGEPPGKSRVMTIGEGWLTIFLQ
37 PASHHLAGLIEHHEAAWAGFKRGWQASEWAK-----RYRRKQHWTGED-----
38 YVLSSDLKTATDYCRHD-FSLSLLGFLEGSGQ-----GDNPYLRAAAELL---CSPR
39 YY--TGPVEG----SIGPTSRGILMDPGSKIVLTLFN----LAAEEEAFVRWATPR
40 MIPTISVGVMVTIADP-ATSW-KIRNLAVSG-----DDHTAR-----GP
41 KSYLKGITQAHIANGMEVSESSNFLSPIGGFYCEEALL-FRGLDLQKLRA-YSVFWDIP
42 YL--EHMHVDSLKRCLSPVMK---GDAT-RDEKNPAIGKASTFTGMLRWFAAGGW----
43 ACTRPIFSARFEQRFKTFLPEGILR-----YLPRELGGVAAPCF---HLTAGMLESS
44 LTKLDPSVRSAl---NLVLAGERG--ELAKLLAKIPANTS MRGIPE--NLARDQVT----
45 -----QILETF---PGVQGRIDESQLKAMADDQMGD-QWWG-----DATA
46 RDRARFAR---RNN-L---ISVVDALN-----DIERPYLFREFLA-----
47 -----PGQHAVHE
48 GPSQFKTRIYDTTPWPVRVNR--FRS-LA----LSAKPDLLD--HHHSAEGVSDGKLVG
49 M-----IQKGKGLPPPPRMYIP-----ERVIFGDDL
50 VTLQTPL-----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 >IABI01004424.1
4 -----LDPLPPCPPVAAVISTEGNGFVSQQSEPLARS-PKDV
5 GPPKGDL-----SSFVPTYGGFNAGIVDRPGGLATHLAWRNPLSEKILRRSRNGKPY
6 GVVAISNGVKP---SLRE--AILRRHPEICTYPDVVIRRLSEDGHRFNIRGAYLSIEDA
7 IVC---AVPEILSD----DPVFPLLRRWVLNTGAV-HEGVQVIKMWKSMIVSWRWS
8 LRSRT---EQPLAVPGSIFFNEKGHFVGFPEESILSSVMEDIFITG--ITSK-HDGTRV
9 CHLCS-ARGMPPAGQKKSEDSLREHHETLTGQRDD--ISQ-----ER
10 LDLVKKLARRIGRKVATVL-SKEKGDFRSDGHVSATTSAVEDTVKSGGRANTLAT---
11 QFIPWAKKVATEPLEGNDILG-NA--VRERPG----HARWR--TTCREVPVPFERDCDC
12 AGEDGVRPNSSNGLPAVHSLEANNREECPNLGLFSSI-WRMPEFD-----
13 -----MPTNLDNPLTGVDENIKQML--QWAHHQGCSAGYIEGPLVGSREEFNTSK
14 V-----VKPIPVR-SVVERPGGKARVSKNEAWFTFLQ
15 PFGHHSVGLSRLPSTRAGLSAAAQGYEWAK-----SASKSTGIDGAS-GM-----
16 LFLSSDLKTATDFCDHR-VSKSLLEGFYEGVGT----NPDPYIRLCIEIL---CSPR
17 IL-----E---DGSITSRGIMMGEPGCKAVLTLFN----LVAEEEAFTAHCQNY
18 PEIL----DLAVVPLRKSW-RVP--ACAG-----DDHIAY-----GP
19 ETYLRRITRNHSNGMQVSWPSNFVPIAAIYCEEMILKPHLRF----EVGYLPKAD
20 YR--ITVHIDAMKVRLSPYRKETSGGPND-SDKTNPAVGKSSFTKLDWQPKGW---
21 ELARKPFLERFRSRMMQYLPHIEAG----YQYLPRYLGGLGVPDV----IPSSLRTL
22 SADLSPRHVAII---EAIMDGRGTP-AIRRCLSRFAKSDEARGIDLSNTAAEDENI---
23 -----ALMKRA----IQGFDTPEEGVLTFEAVREKLGLSFDQAE----NMRR
24 RDFNRKAG--RELQL---ISITDALR-----EIERPYITVELLK-----
25 -----
26 GKASEPKSRFDVRPWVKRYAK-LQDD-LETG--VPKNPNNVL--VASTLDEVFNA---
27 -----VITGNDEYIPKERISR-----VISAFIGVEV-VR-
28 -----
29 -----
30 -----
31 -----
32 -----
33 -----
34 >GEWH01000774.1
35 -----KFNLDSTPCARVAAALDDARYGEAQATRYSIPDPDAT
36 GVELVTR---IPWRNPLTEKSEVSFGGKRRQRRKPL-----KRPSYHSMIKPS
37 QRFRILS-----VTPIIKECSRQLNWLSSEDGD-----RFNRVRGCFLGIEDA
38 IYS--SCPELRDD----DPIWSKLLKWTFATGSSRHSGVDTVIKSWKSFATWVKWAA
39 LRSET---EEPAVIRGFPFF-DENGFTGFPADSQFSELIEPILRTG--ITSK-GEGTRI
40 CHLIT-SRGLPPAGQSKADASLDEHKVLTSDSG--ISD-----ER
41 VEEVRVIAKRIGKLVRGCIVKEGEDRFRSDGHVSATSSACVESTVKQGGRAASMSA---
42 NLLTFLNEIPKTETKGVDILN-NP--FIETGD----LPRWR--TTCRDEVNELSRDCDC
43 FET-----CLNKWECKNNNTWSSSY-WSIPDFE-----
44 -----SDLENMYDNPLMGVDHAIGPQMV--QWAYLEGRYRGVFTGPRVGDKQQDFLTC
45 T-----VSRPPQIIRSVVAEPGGKARVVTKGEDWLTFLQ
46 PFSHHTIGLLRRLPSARAGLSAAAMGYEWTK-----ALSRKEGLDKED-NL-----
47 YFLSSDLKTATDYCDHR-ISKALLEGFYSGIGW----YKDHSYIKLAIDLL---CSSR
48 EC--E---G----EFT--KRGILMGDPGTCAVLTFLN----LVAEEESFQDWK---
49 -----RKNPER-EVKW-RCP--AFAG-----DDHVAL-----GP
50 KEYVSQITVNHARNGMKVGWESNFVSKIGAFYCEELLAHVNGTKF----EVGYLPKAE
51 YE--STIHVDAMKIRLMSPYRKETSAAPND-SDKANPAIGKASSFYKKLSWFQDGW---
52 KSACAPLLQRFRRLRMMMDYLPAENQV----YQYLPYYLGGLDMPLL-EGNPESVSDWTEL
53
54
55
56
57
58
59
60

1
2
3 IPQLDLAHQSAI---QMVIDGKADS-WLLRCLSRFAASNDYARGIDLSNRYAEDSIR----
4 -----GLLQIKKISTQETESYTYEQVVEELVKTGISK-TSIE-----DMRK
5 RDFGRKAQ---DQLKL---ISMSSALN-----LIEKPYIVTNLLR-----
6 -----G
7 EGVDSQKQKYDLVPWKTRYNR-FIKD-LESVG--VSAEGYEYK--TGMTKDDIFNR----
8 -----FVSGKDIYIPERRITS-----ILTAYKVGW-WST
9 PSLLTET--KLS-----
10 -----
11 -----
12 >GEWH01001802.1
13 -----
14 -----
15 -----
16 -----
17 -----
18 -----PILRTG--ITSK-GEGTRI
19 CHLIT-SRGLPPAGQSKADASLDEHFKVLTSDSG---ISD-----ER
20 VEEVRVIAKRIGKLVRCIVKEGEDRFRSDGHVSATSSACVESTVKQGGRAASMSA----
21 NLLTFLNEIPKTETKGVDILN-NP--FIETGD----LPRWR--TTCRDEVNELSRDCDC
22 FET-----CLNKWECKNNNTWSSSY-WSIPDFE-----
23 -----SDLENMYDNPLMGVDHAIGPQMV--QWAYLEGRYRGVFTGPRVGDKQQDFLTC
24 T-----VSRPPQIIRSVVAEPGGKARVVTKGEDWLTFLQ
25 PFSHHTIGLLRRLPSARAGLSAAAMGYEWTK-----ALSRKEGLDKED-NL-----
26 YFLSSDLKTATDYCDHR-ISKALLEGFYSGIGW----YKDHSYIKLAIDL---CSR
27 EC-----EGEFTKRGILMGDPGTCAVLTFLN-----LVAEEESFQDWK---
28 -----RKNPER-EVKW-RCP--AFAG-----DDHVAL-----GP
29 KEYVSQITVNHARNGMKVGWESNFVSKIGAFYCEELLA-HVNGT--KFEV-GYLPKA--E
30 YE--STIHVDAMKIRLMSPYRKETSAAPND-SDKANPAIGKASSFYKKLSWFQDGW----
31 KSACAPLQRFRRLRMMDYLPENQV----YQYLPYYLGGLDMPLL-EGNPESVSDWTEL
32 IPQLDLAHQSAI---QMVIDGKADS-WLLRCLSRFAASNDYARGIDLSNRYAEDSIR----
33 -----GLLQIKKISTQETESYTYEQVVEELVKTGISK-TSIE-----DMRK
34 RDFGRKAQ---DQLKL---ISMSSALN-----LIEKPYIVTNLLR-----
35 -----G
36 EGVDSQKQKYDLVPWKTRYNR-FIKD-LESVG--VSAEGYEYK--TGMTKDDIFNR----
37 -----FVSGKDIYIPERRITS-----ILTAYKVGW-WST
38 PSLLTET--KLS-----
39 -----
40 -----
41 >KX373304.1
42 -----MPTIRNELNPRVSF-----
43 -----FAEKHGKGTFKCPLSLR-----HYLGDSRHL
44 RKKCLLHP-----RAGLLKEVLGKSWKTFKCFKTS-----SIVRAHELYLCYLDI
45 CVC---LDPEILS-----IESNRSIGVKVFNTACF--STSTATSLYKEFVENVLILL
46 KCKE-----GEERWTRQNPFKRMMTW---KIFQDAIASV--QLSKGKSLERV
47 LSVLT-TRNFPEDPGKSIQKKVDFIGIISQQPPKGWKSE-----EY
48 RSHFLESIEEIANECES-----TTVSDCHISVTAAGSLKKTVREGGKFAEMIE---
49 EVKGFLAETPSDEELYEFANL-K---WTCNTN---EPRWK--TFIGIIGEINPMS---
50 -----AMTLFTEGV-DFLNEVPNQYLGPIDFS GGSLP
51 PGIPEEDELFITDLRPFRMGLGAQFGNQLL--LYCC-----VFY
52
53
54
55
56
57
58
59
60

1
2
3 K-----KDELPMIRASPVLEGGDKVRWITMASWRDLVIQQ
4 AAATIFRSLMESHKEMKPIFSRANLAWVYLN-----KAREIGPSD-----
5 ICYVSDYSSATDTVDRE-FAEFILTNFIKRFRN--RLSDPLLNFLELGIRNA---VSPK
6 VV--IFPT-G----ERITSSRGVMGEPMSKVILTLIM----FTIGKAAKSIYKLRF
7 -----PKSMEK-LTFW--AP----G-----DDL VAT----GP
8 TEYIDIYSELAKILQILNHSKVFKSRTVFKLCEQWFW-VPGLKS-SVGT-WAITDPGK
9 YR--ESAWVDTVVKLKLGMAMSL---SNHSHFEERNEVIGKAKALSKILRWLPSDT--YP
10 LEYKKLLRTWFLCRFEPKLPRTDSK--TFAFLMLPGHLGGFDLLL-----DQEIREA
11 YQKVSSFSRILF--NNYYDDFIVG---SYLRNMLRNRSYRGFAL---SEDYNL---
12 -----TVEKLV-----KYYIDLLEEAPYEEYKD-SKVE-----GLSF
13 RETDRALF---KAG-F---LTENRIRD-----IVSRPLAFGSIWG-----
14 -----
15 -----
16 --RTLKHSPYNTTPISVRLAK-LWKN-LT-----GDTERWK--NIS-LKGEISEDDFV
17 V-----LCKENKGPVLK VYR-----RNDSL-EDE
18 VLIGLPS--MKV-----SLGIL---
19 -----
20 -----
21 -----
22 >KU935604.1
23 -----MPTIRNELNPRVSF-----
24 -----FAEKHGKGTFKCPLSLR-----HYLGDSRHL
25 RKKCLLHP----RAGLLKEVLGKSWKTFKCFKTS-----SIVRAHELYLCYLDI
26 CVC---LDPEILS-----IESNRSIGVKVFNNTACF--STSTATSLYKEFVENVLILL
27 KCKE-----GEERWTRQNPFKRMMTW---KIFQDAIASV--QLSKGKSLERV
28 LSVLT-TRNFPEPDGKSIQKKKVDFIGIISQQPPKGWKSE-----EY
29 RSHFLESIEEIANECES-----TTVSDCHISVTAAGSLKKTVREGGKFAEMIE---
30 EVKGFLAETPSNEELYEFANL-K---WTCNTN---EPRWK--TFIGIIGEINPMS---
31 -----AMTLFTEGV-DFLNEVPNQYLGPIDFSGGSLP
32 PGIIPEDLFITDLRPFMRMGLGAQFGNQLL--LYCC-----VFY
33 K-----KDELPMIRASPVLEGGDKVRWITMASWRDLVIQQ
34 AAATIFRSLMESHKEMKPIFSRANLAWVYLN-----KAREIGPSD-----
35 ICYVSDYSSATDTVDRE-FAEFILTNFIKRFRN--RLSDPLLNFLELGIRNA---VSPK
36 VV--IFPT-G----ERITSSRGVMGEPMSKVILTLIM----FTIGKAAKSIYKLRF
37 -----PKSMEK-LTFW--AP----G-----DDL VAT----GP
38 TEYIDIYSELAKILQILNHSKVFKSRTVFKLCEQWFW-VPGLKS-SVGT-WAITDPGK
39 YR--ESAWVDTVVKLKLGMAMSL---SNHSHFEERNEVIGKAKALSKILRWLPSDT--YP
40 LEYKKLLRTWFLCRFEPRLPRTDSK--TFAFLMLPGHLGGFDLLL-----DQEIREA
41 YQKVSSFSRILF--NNYYDDFIVG---SYLRNMLRNRSYRGFAL---SEDYNL---
42 -----TVEKLV-----KYYIDLLEEAPYEEYKD-SKVE-----GLSF
43 RETDRALF---KAG-F---LTENRIRD-----IVSRPLAFGSIWG-----
44 -----
45 -----
46 -----
47 -----
48 --RTLKHSPYNTTPISVRLAK-LWKN-LT-----GDTERWK--NIS-LKGEISEDDFV
49 V-----LCKENKGPVLK VYR-----RNDSL-EDE
50 VLIGLPS--MKV-----SLGIL---
51 -----
52 -----
53 >KY628364.1
54 -----MPTIRNELNPRVSF-----
55 -----FAEKHGKGTFKCPLSLR-----HYLGDSRHL
56
57
58
59
60

1
2
3 RKKCLLHP-----RAGLLKEVLGKSWKTFKCFKTS-----SIVRAHELYLCYLDI
4 CVC---LDPEILS-----IESNRSIGVKVFNTACF---STSTATSLYKEFVENVLILL
5 KCKE-----GEERWTRQNPFKRMMTW---KIFQDAIASV--QLSKGKSLERV
6 LSVLT-TRNFPEPDGKSIQKKVDFIGIISQQPPKGWKSE-----EY
7 RSHFLESIEEIANECES-----TTVSDCHISVTAAGSLKKTVREGGKFAEMIE---
8 EVKGFLAETPSNEELYEFANL-K---WTCNTN---EPRWK--TFIGIIGEINPMS---
9 -----AMTLFTEGV-DFLNEVPNQYLGPIDFSGGS LP
10 PGIPIEEDLFITDLRPFRMGLGAQFGNQLL--LYCC-----VFY
11 K-----KDELPMIRASPVLEGGDKVRWITMASWRDLVIQQ
12 AAATIFRSLMESHKEMKPIFSRANLAWVYLN-----KAREIGP-SD-----
13 ICYVSDYSSATDTVDRE-FAEFILTNFIKRFRN--RLSDPLLNFLELGIRNA---VSPK
14 VV--IFPT-G-----ERITSSRGVFMGEPMMSKVILTLIM-----FTIGKAAKSIYKLRF
15 -----PKSMEK-LTFW--AP----G-----DDL VAT-----GP
16 TEYIDIYSELAKILGQILNHSKVFKSRTVFKLCEQWFW-VPGLKS-SVGT-WAITDPGK
17 YR--ESA WVDTVKLKLLGAMSL---SNHSHFEERNEVIGKAKALSKILRWLPSDT--YP
18 LEYKKLLRTWFLCRFEPKLPRTDSK--TFAFLMLPGHLGGFDLLL D-----DQEIREA
19 YQKVSSFSRILF--NNYYDDFIVG---SYLRNMLRNRSYRGFAL---SEDYNL---
20 -----TVEKLV-----KYYTDLLEEAPYEEYKD-SKVE-----GLSF
21 RETDRALF---KAG-F---LTENRIRD-----IVSRPLAFGSIWG-----
22 -----
23 --RTLKHSPYNTTPISVRLAK-LWKN-----LTGDTERWK--NIS-LKGEISEDDFV
24 V-----LCKENKGPVLK VYR-----RNDSL-EDE
25 VLIGLPS--MKV-----SLGIL---
26 --PHDG-----YISW-----
27 >JN400242.1
28 -----
29 -----RSTLPQSWE-----RGDRYHSDL LKR
30 DLMIRL-----TRDL--GLSGRHADRLVNRP AI-----HLKRIEEFISGLVDS
31 FWL--GDPRTFALGSPD--FKTFRTIVRKIFAVGTS--NLGMLMANWKEWTNWFTHSV
32 CETG-----LREEAKLSNHNMFHRLARL---SVVQKALSGG--VPLM-ELGSIV
33 AHLTS-TRQMPYMLPTEVKAKADFIEIISTPFH--VEP-----SH
34 HRQM VQCAARMGR LCMS-L-RNGVPVSDRSSHF SATSSGELDHSLTRGGQAQALKD---
35 AIDRWLIPASGSTYMEDTPFG-V---AEHREG---VPLWK--TLFVDQETQLEL---
36 -----IFSDFGD S-LDWIKDVP-----
37 -----GRVYGLDDYTGRQILYVAWKE-----
38 -----MEDIPHIRASTVPELGNKARI VTLSAFWLNV LQA
39 PLSHIMKEVLKYHPSCFASFTRGEQAWHAAS-----GLGRLNPR SVA--GY-----
40 SVLSSDLKNATNAQHIA-LTRDMLRAFIYSSGL---VV-SKAYVNLVLD T-----ICPR
41 LV--EL-D-G---EFTLSCRGIMMGEAIAKPSMTLLN-----LVVEGLAFLKYENKL
42 VLLN----TDKAAP-SRRW-RCY--HVGG-----DDHLAV-----GP
43 DQYLDLITHNHELSGSIISPDKHA KS RKMVKYCERVLL-VQNLQ-----YNT PDRK-P
44 FE--HGLIVDSIKVRLLEKGQS---TLIA-KDNKNVAVGKSQQLVKSLDWLPKQL--YS
45 KGFIRSIQHLFIKRMGSLLPNRNRDEEA FHSVCLPKILGGYGLGLA-----EDLLTH
46 LSKSKPEIQQTV--SAVLYGI TR-RARNILSRLNTTISDRSIQG---TKWYED---
47 -----EMVQQF-----REYPDMVGAIDGREMRR--RFP-----AST-
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3
4
5
6
7
8
9 >JN400243.1
10
11 -----RSTLPQSWE-----RGDRYHSDLKR
12 DLMIRL-----TRDL--GLSGRHADRLVNRAI-----HLKRIEFISGLVDS
13 FWL---GDPRTFALGSPD--FKTFRTIVRKIFAVGTS---NLGMLMANWKEWTNWFTHSV
14 CETE-----LREEAKLSNHNMFHRLARL---SVVQKALSGG--VPLM-ELGSIV
15 AHLTS-TRQMPYMLPTEVKAKADFIEIISTPFH---VEP-----SH
16 HRQMVQCAARMGRCLCMS--LRNGVPVSDRSSHFATSSGELDHSLTRGGQAQALKD---
17 AIDRWLIPASGSTYMEDTPFG-V---AEHREG----VPLWK--TLFVDQETQLEL---
18 -----IFSDFGDSDL-DWIKDVP-----
19 -----GRVYGLDDYTGRQILYVAWKE-----
20 -----MEDIPHIRASTVPELGNKARIVTLSAFWLNVLQA
21 PLSHIMKEVLKYHPSCFASFTRGDQA WHAAS-----GLGRLNPRSV A--GY-----
22 SVLSSDLKNATNAQHIA-LTRDMLRAFIYSSGL---VV-SKAYVNLVLD T-----ICPR
23 LV--EL-D-G----EFTLSCRGIMMGEAIAKPSMTLLN-----LVVEELAFLKYENKL
24 VLLN----TDKAAP-SRRW-RCY--HVGG-----DDHLAV-----GP
25 DQYLDLITHNHELSGIISPDKHAKSRKMVKYCERVLL-VQNLQ-----YNTPDRK-P
26 FE--HGLIVDSIKVRLLEKGQS---TLIA-KDNKNVAVGKSSQLVKSLDWLPKQL--YS
27 KGFIRSIQHLFIKRMGSLLPNRNRDEEA FHSVCLPKILGGYGLGLA-----EDLLTH
28 LSKSKPEIQQT V--SAVLYGIETR-RARNILSRLNTTISDRSIQG---TKWYED---
29 -----EMVQQF-----REYPDMVGAIDGREMRR--RFP-----AST-
30
31
32
33
34
35
36
37
38
39
40 >JN400241.1
41 -----MQEDLILVQPDTDWVGGVNRSWRSL P-VLN
42 GVEYI-----LPRGSLVGLKLHGRSTLPQSWE-----RGDRYHSDLKR
43 DLMIRL-----TRDL--GLSGRHADRLVNRAI-----HLKRIEFISGLVDS
44 FWL---GDPRTFALGSPD--FKTFRTIVRKIFAVGTS---NLGMLMTNWEWTNWFTHSV
45 CETE-----LREEAKLSNHNMFHRLARL---SVVQKALSGD--VPLM-ELGSIV
46 AHLTS-TRQMPYMLPTEVKAKADFIEIISTPFH---VEP-----SH
47 HRQMVQCAARMGRCLCMS-L-RNGVPVSDRSSHFATSSGELDHSLTRGGQAQALKD---
48 AIDKW LIPASGSTYMEDTPFG-V---AEHREG----VPLWK--TLFVDQETQLEL---
49 -----IFSDFGDSDL-DWIKDVP-----
50 -----GRVYGLDDYTGRQILYVAWKE-----
51 -----MEDIPHIRASTVPELGNKARIVTLSAFWLNVLQA
52 PLSHIMKEVLKYHPSCFASFTRGDQA WHAAS-----GLGRLNPRSV A--GY-----
53 SVLSSDLKNATNAQHIA-LTRDMLRAFIYSSGL---VV-SEAYVNLVLD T-----ICPR
54
55
56
57
58
59
60

1
2
3 LV--EL-D-G----EFTLSCRGIMMGEAIAKPSLTLLN----LVVEELAFLKYENKL
4 VLLN----TDKAAP-SRRW-RCY--HVGG-----DDHLAV----GP
5 DQYLDLITHNHELSGIISPDKGSRKMKYCERVLL-VQNLQ-----YNTPDRK-P
6 FE--HGLIVDSIKVRLLEKGQS---TLIA-KDNKNVAVGKSSQLVKSLDWLPQL--YS
7 KGFIRSIQHLFIKRMGSLLPNRNRDEEAHSVCLPKILGGYGLGLA-----EDLLTH
8 LSKSKPEIQQTV--SAVLYGIETR-RARNILSRLNTTISDRSIQG---TKWYED---
9 -----EMVQQF-----REYPDMVGAIDGREMRR--RFP-----ASTY
10 QESKAHAA--SAG-W---LPIDQLAR-----RVTRGHLFGEMLS-----
11 -----
12 -----
13 -GGAKRQNVFKTRNWIQEFT--LRKR-LEEEVGYLPLPEYADL--EVKTLRNALGAMTTD
14 L-----FIDSQVTSFDKGVD-----DSFDFFDGEL-LRT
15 YEAHQPN--LVV-----GLQFIGLD---
16 --STCN-----AFTRRVRRRY-----
17 >KU295726.1
18 -----MQEDLILVQPDTDWVGEVNRSWRSL--PVLN
19 GVEYI-----LPRGSLVGLKLHGRSTLPQSWE-----RGDRYHSDLKR
20 DLMIRL-----TRDL--GLSGRHADRLVNRPAl-----HLKRIEEFISGLVDS
21 FWL--GDPRTFVLGSPD--FKTFRTIVRKIFAVGTS---NLGMLMTNWKEWTNWFTHSV
22 CETE-----LREEAKLSNHNVFHRLARL---SVVQKALSGG--VPLM-ELGSIV
23 AHLTS-TRQMPYMGPLTEVKAKADFIEIISTPFH---VEP-----SH
24 HRQMVQCAARMGRCMS-L-RNGVPVSDRSSHFATSSGELDHSLTRGGQAQALKD---
25 AIDRWLIPASGSTYMEDTPFG-V---AEHREG---VPLWK--TLFVDQETQLEL---
26 -----IFSDFGDSL-DWIKDVP-----
27 -----GRVYGLDDYTGRQILYVAWKE-----
28 -----MEDIPIHIRASTVPELGNKARIVTLSAFWLNLQA
29 PLSHIMKEVLKYHPSCFASFTRGDQAWHAAS-----GLGRLNPRSVA--GY-----
30 SVLSSDLKNATNAQHIA-LTRDMLRAFIYSSGL---VV-SEAYVNLVLDT---ICPR
31 LV--EL-D-G----EFTLSCRGIMMGEAIAKPSLTLLN----LVVEELAFLKYENKL
32 VLLN----TDKAAP-SRRW-RCY--HVGG-----DDHLAV----GP
33 DQYLDLITHNHELSGIISPDKGSRKMKYCERVLL-VQNLQ-----YNTPDRK-P
34 FE--HGLIVDSIKVRLLEKGQS---TLIA-KDNKNVAVGKSSQLVKSLDWLPQL--YS
35 KGFIRSIQHLFIKRMGSLLPNRNRDEEAHSVCLPKILGGYGLGLA-----EDLLTH
36 LSKSKPEIQQTV--SAVLYGIETR-RARNILSRLNTTISDRSIQG---TKWYED---
37 -----EMVQQF-----REYPDMVGAIDGREMRR--RFP-----ASTY
38 QESKAHAA--SAG-W---LPIDQLAR-----RVTRGHLFGEMLS-----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -GGAKRQNVFKTRNWIQEFT--LRKR-LEEEVGYLPLPEYADL--EVKTLRNALGAMTTD
46 L-----FIDSQVTSFDKGVD-----DSFDFFDGEL-LRT
47 YEAHQPN--LVV-----GLQFIGLD---
48 --STSN-----AFTRRVGRY-----
49 >KU295722.1
50 -----MQEDLILVQPDTDWVGGVNRSWRSL--PVLN
51 GVEYI-----LPRGSLVGLKLHGRSTLPQSWE-----RGDRYHSDLKR
52 DLMIRL-----TRDL--GLSGRHADRLVNRPAl-----HLKRIEEFISGLVDS
53 FWL--GDPRTFVLGSPD--FKTFRTIVRKIFAVGTS---NLGMLMTNWKEWTNWFTHSV
54 CETE-----LREEAKLSNHNMFHRLARL---SVVQKALSGG--VPLM-ELGSIV
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 AHLTS-TRQMPYMLPTEVKAKADFIEIISTPFH---VEP-----SH
4 HRQMVQCAARMGRLCMS-L-RNGVPVSDRSSHFATSSGELDHSLTRGGQAQALKD----
5 AIDRWLIPASGSTYMEDTPFG-V---AEHREG---VPLWK--TLFVDQETQLEL----
6 -----IFSDFGDSDL-DWIKDVP-----
7 -----GRVYGLDDYTGRQILYVAWKE-----
8 -----MEDIPIHIRASTVPELGNKARIVTLSAFWLNVLQA
9 PLSHIMKEVLKYHPSCFASFTRGDQAWHAAS-----GLGRLNPRSVAA-GY-----
10 SVLSSDLKNATNAQHIA-LTRDMLRAFIYSSGL---VV-SEAYVNLVLDT----ICPR
11 LV--EL-D-G---EFTLSCRGIMMGEAIAKPSLTLLN---LVVEELAFLKYENKL
12 VLLN----TDKAAP-SRRW-RCY--HVGG-----DDHLAV----GP
13 DQYLDLITHNHELSGSIISPDKGSRKMVKYCERVLL-VQNLQ-----YNTPDRK-P
14 FE--HGLIVDSIKVRLLEKGQS---TLIA-KDNKNVAVGKSSQLVKSLDWLPKQL--YS
15 KGFIRSIQHLFIKRMGSLPNRNRDEEAHSVCLPKILGGYGLGLA-----EDLLTH
16 LSKSKPEIQQTV--SAVLYGIETR-RARNILSRLNTTISDRSIQG---TKWYED---
17 -----EMVQQF-----REYPDMVGAIDGREMRR--RFP-----ASTY
18 QESKAHAA--SAG-W---LPIDQLAR-----RVTRGHLFGEMLS-----
19 -----
20 -----GGAKRQNFKTRNWIQEFTLRKR-LEEEVGYLPLPEYADL--EVKTLRNALGAMTTD
21 L-----FIDSRQVTSFDKGVD-----DSFDFFDGE-LRT
22 YEAHQPN--LVV-----GLQFIGLD---
23 --STCN-----AFTRRVGRGY-----
24 >KU295727.1
25 -----MQEDLILVQPDTDWVGGVNRSWRSLL-PVLN
26 GVEYI-----LPRGSLVGLKLHGRSTLPQSWE-----RGDRYHSDLKR
27 DLMIRL-----TRDL--GLSGRHADRLVNRPAl-----HLKRIEFISGLVDS
28 FWL---GDPRTFVLGSPD--FKTFRTIVRKIFAVGTS---NLGMLMANWKEWTNWFTHSV
29 CETE-----LREEAKLSNHNMFHRLARL---SVVQKALSGG--VPLM-ELGSIV
30 AHLTS-TRQMPYMLPTEVKAKADFIEIISTPFH---VEP-----SH
31 HRQMVQCAARMGRLCMS-L-RNGVPVSDRSSHFATSSGELDHSLTRGGQAQALKD----
32 AIDRWLIPASGSTYMEDTPFG-V---AEHREG---VPLWK--TLFVDQETQLEL----
33 -----IFSDFGDSDL-DWIKDVP-----
34 -----GRVYGLDDYTGRQILYVAWKE-----
35 -----MEDIPIHIRASTVPELGNKARIVTLSAFWLNVLQA
36 PLSHIMKEVLKYHPSCFASFTRGDQAWHAAS-----GLGRLNPRSVAA-GY-----
37 SVLSSDLKNATNAQHIA-LTRDMLRAFIYSSGL---VV-SEAYVNLVLDT----ICPR
38 LV--EL-D-G---EFTLSCRGIMMGEAIAKPSLTLLN---LVVEELAFLKYENKL
39 VLLN----TDKAAP-SRRW-RCY--HVGG-----DDHLAV----GP
40 DQYLDLITHNHELSGSIISPDKGSRKMVKYCERVLL-VQNLQ-----YNTPDRK-P
41 FE--HGLIVDSIKVRLLEKGQS---TLIA-KDNKNVAVGKSSQLVKSLDWLPKQL--YS
42 KGFIRSIQHLFIKRMGSLPNRNRDEEAHSVCLPKILGGYGLGLA-----EDLLTH
43 LSKSKPEIQQTV--SAVLYGIETR-RARNILSRLNTTISDRSIQG---TKWYED---
44 -----EMVQQF-----REYPDMVGAIDGREMRR--RFP-----ASTY
45 QESKAHAA--SAG-W---LPIDQLAR-----RVTRGHLFGEMLS-----
46 -----
47 -----GGAKRQNFKTRNWIQEFTLRKR-LEEEVGYLPLPEYADL--EVKTLRNALGAMTTD
48 L-----FIDSRQVTSFDKGVD-----DSFDFFDGE-LRT

1
 2
 3 YEAHQPN--LVV-----GLQFIGLD---
 4 --STCN-----AFTRRVGRGY-----
 5 >HAFC01092190.1
 6 -----MRELSYKLIRVTIPRGATGKSVSEAYRAIL-PCLL
 7 PTDQEDV----YNYTLPLGIFKLKFQGSPLPQLWEK-----ADSVYLSEREKK
 8 IFRFHL-----KTGL--GLGKRVTEVFLSRPRS-----HYLRTFGFIDGLVDA
 9 FWA---ADPKFFLVGSKD--YKIFKSIVRKIIKGCT---DLGRLVDDWKDWGTRIFHQV
 10 ARTKT-----IGSLPEAKPENIFRILSEL---SWIKSILA KVKDLDs-LDGNNL
 11 ASIRS-SRQFPYMGAKTEKSLERFKEVVSTPFI---VDP-----ET
 12 SWQMVA TARRIGGICRA---IRTRPIPDMAAHFSVSSSGEYTYTINKGAFSQAVFD---
 13 GIVPYLMRI-IDTSSEETPFG-L---VHYQSG---KEVWR--YLFRDEEISQ---
 14 -----GEFL-DNRKGFP-----
 15 -----KCQLPRLLGLDETFGKQVLWAAWKA-----SLE
 16 F-----QGKPIPVRAEIVPEMGNKARVVTCPWWLQILQA
 17 PLAHCACIAMMKWHPSVYSSFHKMDQAWCASV-----GLHKVKDPDLT--KY-----
 18 WLLSSDLKDATNAQSID-LTKAMLKAFLLGSNL---LSGREAYVNVCLELLGSTRLGSR
 19 EV--HFKD-K---TVVLSKSGIMMGEP LAKVGLTLLN-----LSVEELALLTHIGRL
 20 DLLL-----KPDPTP-KLK W-RFC--HIGG-----DDHLAW-----GP
 21 EPYLEKITELHYKVGSHLSPGEHHSRIFVRYTEKVLN-VMNMGKSPMDA-QVN KIDAF-
 22 ---KRIMVDGIKIRLLESQGS---TMIS-KDNKNVAIGKAGQLVRSIQWLEPKI-GFT
 23 ESRIDSIRDLFCLRLNRFLPSQHRNWRCYQATKLPRVLLGGFGLGIT-----PSEQYQA
 24 YVDS P SPIRWVV--SKWKAGEPIV-DELQILSFLNSNVSTRGIAS---VEAYRK---
 25 -----EILDQL-----QTYPNLVNAKSWHDICE--LFPVEQ-----DAPS
 26 SMIVDKAA--EAG-Y---LTFEKFAE-----MSTRGAIFQNLLL-----
 27 -----
 28 --ATEEVKIFQTRPYIQT MRR-VTHQ-LQLRG--VDQFGDTPI--EQADFEEM LKEIGRP
 29 I-----FFNIDQVTTCDIGEYD-----PLKPEEETWNFQDTTF-REL
 30 YTKGTPS--LIV-----GLSFLGLRY--
 31 -----
 32 >NC_035126.1
 33 -----
 34 -----
 35 -----
 36 -----
 37 -----
 38 -----
 39 -----
 40 -----MVSGLVDA
 41 LWC---ADEFVFVLGSID--LKTLKALVRKJFCVGT F---NLGLLVDDYKA WANRLFHK A
 42 ASTQT-----IGQMEPLRCNNIFKMLDSI---EYIKAIQDSG-SDCSF-INLQKL
 43 SHLVS-SRQLPYMGLKTEIKSLKGFKDV VTS D YR---PTL-----ET
 44 LDLLTKAARRIGGICRS-I-RKDRPIPDAE AHISMTCSGEATTSIIDGGQACAVRE---
 45 GFERIMGRIAESDLLEDTPFG-V---ASHREG---YALWR--TIFREEPVE-----
 46 -----GGSLYDPLYQGYPKDQP-----
 47 -----GRYLGLDRVLGKQLLYVAWKE-----S
 48 I-----LNPTIVL RASIVPEMGNKARVVTMSPYWVQLLQA
 49 PLAHLTIAGMRLHPSVFSSFARQDQAWEAAK-----GISRLNIDNWE--DL-----
 50 HVLSSDLKDATNAQQQL-LTRTMLRAFMNGYGM---IH-RNTYSELALST---IVER
 51 LI--TFED-Q---TSVWATTGIMMG EPIAKPSLTLLN-----LAIEELAFLTYTNRL
 52 DLLD---SDLPSP-YRAW-RFI--HIGG-----DDHLAI-----GP
 53 KKYLELITNYHLLAGSHIDPGKHGYSQVAVKYTERILS-VKNFR-----FKAPFDYDN
 54
 55
 56
 57
 58
 59
 60

1
2
3 YP--KSMIVDSVKVRLLERGQS---TMQK-KDNKNVAIGKSRRMSGCLDWLPSDPAYWP
4 SDKKISIRNLFINRMGPLLPSKSLHPKYHSVQLPSILGGFNLGLK-----SEIYQA
5 YTLAPPPIRWVL---NKAAGVDVT-AELKILSLLRNISDRGTDP---LQEYRA---
6 -----SIIDQL-----YSYPQLIGAITWKEILS--KFPP-----QNEN
7 VRWSIYAA---REAHY---LSIEEFAE-----WSTRGNTFVQLIM-----
8 -----
9 -----
10 --GRGKNNVFNTRPYVRTFAR-VWTE-LEALQ--TDLYGDSPL--GEGEFNKTLRGLSRC
11 L-----YFDTSQSTTIDIGYYK-----PDDPDREEFEVDAKY-LEA
12 FDRMLPT--LVV-----GKKFIGFRN--
13 -----
14 >HACV01003050.1
15 -----AECMNFCGKLIQAQHLRDCVGKRRNGSPWDKLL-PVLG
16 NPNGQQS-NTLLYYHMPGSLLSQTFKGRSLRKQKFE-----KESKYFSRPTY
17 GLFVLL-----TRDI--GVTKQIAKILVDRPV-----HFKRIQEFinGLVDS
18 LWL---ANEQVFLECSQE--LLIVKKTIRKLFCVGT-----NLKDLVDQWKEWGNYLHTV
19 ARTET-----IGPLKAVAKNNIFRTLNGI----PYISRMYHTN--EPDM-LLLQHL
20 SHLVS-SRQMPYMGFKTELDSRKKFKEVLLSSYE---PPK-----DF
21 LQKISMAARRIGGICRT---LRPKLNSGETHISVTSSGTITHSIEQGGQAAAVME---
22 AMKRILLVVPEESFEEDTPFG-T--AVHKKG----IPLWK--TLYRSSNDDM-----
23 -----SSYFLEPY-NLIKEMS-----
24 -----GRFRGLDRVTVGAQIMYVAWRELEATPILRAEVVPEMGNKARHVTL
25 SAYWLNLQSPLCHELLIDSMDKYHWRELEATPILRAEVVPEMGNKARHVTLSAYWLNL
26 QS
27 PLCHLLIDSMDKYHPSVFSSNRQDQAFEAVK-----GMCLKKELQLLP-GH-----
28 AVLSSDLKDATNAQQWS-VTKAILKGYISGANL---SF-SDRYINLVLDL----IGPR
29 LV--KFCD-G----TTVLSQVGIMMGEAIAKPSLTLLN-----LSIEELAFLEYNNAT
30 ELLT----SNDPAP-YKQW-RFL--HIGG-----DDHLAI-----GP
31 VCYLNQITHNHMLAGSHIDPGKHGFSQICVRYTERLIN-IRNFQ-----YKQPFHRDD
32 YS--KSIIVDSVKVRLLERGLS---TEMK-RDNKNAAIGKSQTLGGTLEWLPIDDRFWT
33 ETKKESIRALFVERMGELLPRKAVNPRAYAAIHLQPQYFGGYGLGFK-----REWGKY
34 LRDSPAPHRGGLV---MLAHLGVDVR-HHLKVFRKLNSNLS-----
35 -----
36 -----
37 -----
38 -----
39 -----
40 -----
41 -----
42 -----
43 -----
44 -----
45 -----
46 -----
47 >GACM01002912.1
48 -----NFTGNLIQAHLSECAGKDRKRSPWDRL--PVLG
49 NPNKVDV--NRQFYLMPSGSLLSETYKGSSLMRQKFH-----KRERFFNPRQKA
50 GLRIIL-----TRDI--GLSQRVAVQLVNRPIV-----HYKRIEEFISGFVDS
51 LWL---AEEQIFLDGKW--IKLIRKLIRKIFCVGTY---NL GSLVDQWKEWGNNLHFTL
52 AESTT-----IGVLRKPAENNIFRLLDTI---DYISAVYRGD---KSMKLMQHV
53 SHLIS-SRQMPYMGSSTEEKAWDKFQSVLTSDFV---PSE-----KT
54 IVQLQMAARRVGGICRS---IRSRRIPDGVAHISVTSSGEYSYPIAKGAQAAVKA---
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----

1
2
3 AMERILTVVPLEDQEEDTPFG-L---VRHHKG----IPIWK--TLFRTEPLD-----
4 -----TELPFLAPY-ALIKEQE-----
5 -----GRFAGLDRVTGKQIMYVAWKE-----
6 -----YRPLPVLRAEVVPEMGNKARLVTISDYWLNLQS
7 PLSHVLIDAMKFHPSVFSSFHRQDQAFEATK-----GLCSLKRKSLLK-GE-----
8 AVLSSDLQDATNAQQWS-VTIAMLRGFIQGFGL---SF-RPEYVELVLST---IGPR
9 LV--CFRD-E---ISILSKVGIMMGAEAIKPSLTLLN-----LSIEELAFLRHCR
10 AE ELLY----SVEPAP-YRDW-RYI--HIGG-----DDHLVK-----GP
11 IPYLNLTQIHLDAGSHIDPGKHGFSRICVKYTERLLN-LSNLE-----HGCPFDPSD
12 YS--RSTIVDSVKVRLLERGQS---TLLK-KDNKNVAIGKSTQLGGCLEWLPIDDRFY
13 ETKKASIRALFIERMGSLLPRKAVNPRAFAAIHLPTIVGGYGLGMS-----SELQKF
14 LEASPEPHKGLL---MKAFCGVNVK-EDLKIFRVLNTNTSDRGVEN---IQQFQQ---
15 -----KIIDQL-----SEYPQMVNAMDWWELKR--KFPDP-----SGNS
16 KRTIALAA---DAG-I---LSFEEFAK-----RATRGNLFQALLM-----
17 -----
18 -----
19 -----
20 --GRKDLKVFNTPFVRTYKNIVWDE-AESRGLLTWSEGFNLN--NDE-IAIAIKNIAPQ
21 W-----YFDINQETAMDTGHWD-----PENPETETWDFKDSTY-IDK
22 YTEGFPS--FDV-----GFKVLGLRH--
23 -----
24 -----
25 >GFCJ01029165.1
26 -----
27 -----
28 -----
29 -----
30 -----
31 -----
32 SHLIS-TRHFPTGGKSVERQAVSDFVEVTTLKN--VDS-----DF
33 LKKVRSASTIIGARLKE----NSRPNLQTHFSLSASGSLNKTISEGGRGQE
34 VKE---AIENLKFIPIEDKEIKILGWT---VNDIAG---RPRWQ--TWFRDPDWENPE---
35 -----PDVPFGELLKDTLLGYP-----
36 -----LYRQGFDEVLGKQIMALAANT-----FLEEV
37 S-----AGRRPSVRVITVPEPGGKARI
38 VTTGPWWLYILEQ
39 PFSHAMRSLLVHPSAIAGLCRADQAWQYLR-----LLSNAKKYL
40 GKG--SN-----YCLSSDLKSATDAIPFE-ICEAVFEGFFDGIGQ-----SPTMSPVIA
41 KLL---TCGR EC--ILKS-G---FSYTTVRGIFMGEPM
42 AKA
43 PI
44 LTLV
45 NDKH
46 DDP
47 FPE
48 Q
49 C
50 ID
51 A
52 T
53 R
54 I
55 Q
56 T
57 R
58 F
59 Q
60 P
61 L
62 R
63 F
64 R
65 F
66 R
67 F
68 R
69 F
70 R
71 F
72 R
73 F
74 R
75 F
76 R
77 F
78 R
79 F
80 R
81 F
82 R
83 F
84 R
85 F
86 R
87 F
88 R
89 F
90 R
91 F
92 R
93 F
94 R
95 F
96 R
97 F
98 R
99 F
100 R
101 F
102 R
103 F
104 R
105 F
106 R
107 F
108 R
109 F
110 R
111 F
112 R
113 F
114 R
115 F
116 R
117 F
118 R
119 F
120 R
121 F
122 R
123 F
124 R
125 F
126 R
127 F
128 R
129 F
130 R
131 F
132 R
133 F
134 R
135 F
136 R
137 F
138 R
139 F
140 R
141 F
142 R
143 F
144 R
145 F
146 R
147 F
148 R
149 F
150 R
151 F
152 R
153 F
154 R
155 F
156 R
157 F
158 R
159 F
160 R
161 F
162 R
163 F
164 R
165 F
166 R
167 F
168 R
169 F
170 R
171 F
172 R
173 F
174 R
175 F
176 R
177 F
178 R
179 F
180 R
181 F
182 R
183 F
184 R
185 F
186 R
187 F
188 R
189 F
190 R
191 F
192 R
193 F
194 R
195 F
196 R
197 F
198 R
199 F
200 R
201 F
202 R
203 F
204 R
205 F
206 R
207 F
208 R
209 F
210 R
211 F
212 R
213 F
214 R
215 F
216 R
217 F
218 R
219 F
220 R
221 F
222 R
223 F
224 R
225 F
226 R
227 F
228 R
229 F
230 R
231 F
232 R
233 F
234 R
235 F
236 R
237 F
238 R
239 F
240 R
241 F
242 R
243 F
244 R
245 F
246 R
247 F
248 R
249 F
250 R
251 F
252 R
253 F
254 R
255 F
256 R
257 F
258 R
259 F
260 R
261 F
262 R
263 F
264 R
265 F
266 R
267 F
268 R
269 F
270 R
271 F
272 R
273 F
274 R
275 F
276 R
277 F
278 R
279 F
280 R
281 F
282 R
283 F
284 R
285 F
286 R
287 F
288 R
289 F
290 R
291 F
292 R
293 F
294 R
295 F
296 R
297 F
298 R
299 F
300 R
301 F
302 R
303 F
304 R
305 F
306 R
307 F
308 R
309 F
310 R
311 F
312 R
313 F
314 R
315 F
316 R
317 F
318 R
319 F
320 R
321 F
322 R
323 F
324 R
325 F
326 R
327 F
328 R
329 F
330 R
331 F
332 R
333 F
334 R
335 F
336 R
337 F
338 R
339 F
340 R
341 F
342 R
343 F
344 R
345 F
346 R
347 F
348 R
349 F
350 R
351 F
352 R
353 F
354 R
355 F
356 R
357 F
358 R
359 F
360 R
361 F
362 R
363 F
364 R
365 F
366 R
367 F
368 R
369 F
370 R
371 F
372 R
373 F
374 R
375 F
376 R
377 F
378 R
379 F
380 R
381 F
382 R
383 F
384 R
385 F
386 R
387 F
388 R
389 F
390 R
391 F
392 R
393 F
394 R
395 F
396 R
397 F
398 R
399 F
400 R
401 F
402 R
403 F
404 R
405 F
406 R
407 F
408 R
409 F
410 R
411 F
412 R
413 F
414 R
415 F
416 R
417 F
418 R
419 F
420 R
421 F
422 R
423 F
424 R
425 F
426 R
427 F
428 R
429 F
430 R
431 F
432 R
433 F
434 R
435 F
436 R
437 F
438 R
439 F
440 R
441 F
442 R
443 F
444 R
445 F
446 R
447 F
448 R
449 F
450 R
451 F
452 R
453 F
454 R
455 F
456 R
457 F
458 R
459 F
460 R
461 F
462 R
463 F
464 R
465 F
466 R
467 F
468 R
469 F
470 R
471 F
472 R
473 F
474 R
475 F
476 R
477 F
478 R
479 F
480 R
481 F
482 R
483 F
484 R
485 F
486 R
487 F
488 R
489 F
490 R
491 F
492 R
493 F
494 R
495 F
496 R
497 F
498 R
499 F
500 R
501 F
502 R
503 F
504 R
505 F
506 R
507 F
508 R
509 F
510 R
511 F
512 R
513 F
514 R
515 F
516 R
517 F
518 R
519 F
520 R
521 F
522 R
523 F
524 R
525 F
526 R
527 F
528 R
529 F
530 R
531 F
532 R
533 F
534 R
535 F
536 R
537 F
538 R
539 F
540 R
541 F
542 R
543 F
544 R
545 F
546 R
547 F
548 R
549 F
550 R
551 F
552 R
553 F
554 R
555 F
556 R
557 F
558 R
559 F
560 R
561 F
562 R
563 F
564 R
565 F
566 R
567 F
568 R
569 F
570 R
571 F
572 R
573 F
574 R
575 F
576 R
577 F
578 R
579 F
580 R
581 F
582 R
583 F
584 R
585 F
586 R
587 F
588 R
589 F
590 R
591 F
592 R
593 F
594 R
595 F
596 R
597 F
598 R
599 F
600 R
601 F
602 R
603 F
604 R
605 F
606 R
607 F
608 R
609 F
610 R
611 F
612 R
613 F
614 R
615 F
616 R
617 F
618 R
619 F
620 R
621 F
622 R
623 F
624 R
625 F
626 R
627 F
628 R
629 F
630 R
631 F
632 R
633 F
634 R
635 F
636 R
637 F
638 R
639 F
640 R
641 F
642 R
643 F
644 R
645 F
646 R
647 F
648 R
649 F
650 R
651 F
652 R
653 F
654 R
655 F
656 R
657 F
658 R
659 F
660 R
661 F
662 R
663 F
664 R
665 F
666 R
667 F
668 R
669 F
670 R
671 F
672 R
673 F
674 R
675 F
676 R
677 F
678 R
679 F
680 R
681 F
682 R
683 F
684 R
685 F
686 R
687 F
688 R
689 F
690 R
691 F
692 R
693 F
694 R
695 F
696 R
697 F
698 R
699 F
700 R
701 F
702 R
703 F
704 R
705 F
706 R
707 F
708 R
709 F
710 R
711 F
712 R
713 F
714 R
715 F
716 R
717 F
718 R
719 F
720 R
721 F
722 R
723 F
724 R
725 F
726 R
727 F
728 R
729 F
730 R
731 F
732 R
733 F
734 R
735 F
736 R
737 F
738 R
739 F
740 R
741 F
742 R
743 F
744 R
745 F
746 R
747 F
748 R
749 F
750 R
751 F
752 R
753 F
754 R
755 F
756 R
757 F
758 R
759 F
760 R
761 F
762 R
763 F
764 R
765 F
766 R
767 F
768 R
769 F
770 R
771 F
772 R
773 F
774 R
775 F
776 R
777 F
778 R
779 F
780 R
781 F
782 R
783 F
784 R
785 F
786 R
787 F
788 R
789 F
790 R
791 F
792 R
793 F
794 R
795 F
796 R
797 F
798 R
799 F
800 R
801 F
802 R
803 F
804 R
805 F
806 R
807 F
808 R
809 F
810 R
811 F
812 R
813 F
814 R
815 F
816 R
817 F
818 R
819 F
820 R
821 F
822 R
823 F
824 R
825 F
826 R
827 F
828 R
829 F
830 R
831 F
832 R
833 F
834 R
835 F
836 R
837 F
838 R
839 F
840 R
841 F
842 R
843 F
844 R
845 F
846 R
847 F
848 R
849 F
850 R
851 F
852 R
853 F
854 R
855 F
856 R
857 F
858 R
859 F
860 R
861 F
862 R
863 F
864 R
865 F
866 R
867 F
868 R
869 F
870 R
871 F
872 R
873 F
874 R
875 F
876 R
877 F
878 R
879 F
880 R
881 F
882 R
883 F
884 R
885 F
886 R
887 F
888 R
889 F
890 R
891 F
892 R
893 F
894 R
895 F
896 R
897 F
898 R
899 F
900 R
901 F
902 R
903 F
904 R
905 F
906 R
907 F
908 R
909 F
910 R
911 F
912 R
913 F
914 R
915 F
916 R
917 F
918 R
919 F
920 R
921 F
922 R
923 F
924 R
925 F
926 R
927 F
928 R
929 F
930 R
931 F
932 R
933 F
934 R
935 F
936 R
937 F
938 R
939 F
940 R
941 F
942 R
943 F
944 R
945 F
946 R
947 F
948 R
949 F
950 R
951 F
952 R
953 F
954 R
955 F
956 R
957 F
958 R
959 F
960 R
961 F
962 R
963 F
964 R
965 F
966 R
967 F
968 R
969 F
970 R
971 F
972 R
973 F
974 R
975 F
976 R
977 F
978 R
979 F
980 R
981 F
982 R
983 F
984 R
985 F
986 R
987 F
988 R
989 F
990 R
991 F
992 R
993 F
994 R
995 F
996 R
997 F
998 R
999 F
1000 R
1001 F
1002 R
1003 F
1004 R
1005 F
1006 R
1007 F
1008 R
1009 F
1010 R
1011 F
1012 R
1013 F
1014 R
1015 F
1016 R
1017 F
1018 R
1019 F
1020 R
1021 F
1022 R
1023 F
1024 R
1025 F
1026 R
1027 F
1028 R
1029 F
1030 R
1031 F
1032 R
1033 F
1034 R
1035 F
1036 R
1037 F
1038 R
1039 F
1040 R
1041 F
1042 R
1043 F
1044 R
1045 F
1046 R
1047 F
1048 R
1049 F
1050 R
1051 F
1052 R
1053 F
1054 R
1055 F
1056 R
1057 F
1058 R
1059 F
1060 R
1061 F
1062 R
1063 F
1064 R
1065 F
1066 R
1067 F
1068 R
1069 F
1070 R
1071 F
1072 R
1073 F
1074 R
1075 F
1076 R
1077 F
1078 R
1079 F
1080 R
1081 F
1082 R
1083 F
1084 R
1085 F
1086 R
1087 F
1088 R
1089 F
1090 R
1091 F
1092 R
1093 F
1094 R
1095 F
1096 R
1097 F
1098 R
1099 F
1100 R
1101 F
1102 R
1103 F
1104 R
1105 F
1106 R
1107 F
1108 R
1109 F
1110 R
1111 F
1112 R
1113 F
1114 R
1115 F
1116 R
1117 F
1118 R
1119 F
1120 R
1121 F
1122 R
1123 F
1124 R
1125 F
1126 R
1127 F
1128 R
1129 F
1130 R
1131 F
1132 R
1133 F
1134 R
1135 F
1136 R
1137 F
1138 R
1139 F
1140 R
1141 F
1142 R
1143 F
1144 R
1145 F
1146 R
1147 F
1148 R
1149 F
1150 R
1151 F
1152 R
1153 F
1154 R
1155 F
1156 R
1157 F
1158 R
1159 F
1160 R
1161 F
1162 R
1163 F
1164 R
1165 F
1166 R
1167 F
1168 R
1169 F
1170 R
1171 F
1172 R
1173 F
1174 R
1175 F
1176 R
1177 F
1178 R
1179 F
1180 R
1181 F
1182 R
1183 F
1184 R
1185 F
1186 R
1187 F
1188 R
1189 F
1190 R
1191 F
1192 R
1193 F
1194 R
1195 F
1196 R
1197 F
1198 R
1199 F
1200 R
1201 F
1202 R
1203 F
1204 R
1205 F
1206 R
1207 F
1208 R
1209 F
1210 R
1211 F
1212 R
1213 F
1214 R
1215 F
1216 R
1217 F
1218 R
1219 F
1220 R
1221 F
1222 R
1223 F
1224 R
1225 F
1226 R
1227 F
1228 R
1229 F
1230 R
1231 F
1232 R
1233 F
1234 R
1235 F
1236 R
1237 F
1238 R
1239 F
1240 R
1241 F
1242 R
1243 F
1244 R
1245 F
1246 R
1247 F
1248 R
1249 F
1250 R
1251 F
1252 R
1253 F
1254 R
1255 F
1256 R
1257 F
1258 R
1259 F
1260 R
1261 F
1262 R
1263 F
1264 R
1265 F
1266 R
1267 F
1268 R
1269 F
1270 R
1271 F
1272 R
1273 F
1274 R
1275 F
1276 R
1277 F
1278 R
1279 F
1280 R
1281 F
1282 R
1283 F
1284 R
1285 F
1286 R
1287 F
1288 R
1289 F
1290 R
1291 F
1292 R
1293 F
1294 R
1295 F
1296 R
1297 F
1298 R
1299 F
1300 R
1301 F
1302 R
1303 F
1304 R
1305 F
1306 R
1307 F
1308 R
1309 F
1310 R
1311 F
1312 R
1313 F
1314 R
1315 F
1316 R
1317 F
1318 R
1319 F
1320 R
1321 F
1322 R
1323 F
1324 R
1325 F
1326 R
1327 F
1328 R
1329 F
1330 R
1331 F
1332 R
1333 F
1334 R
1335 F
1336 R
1337 F
1338 R
1339 F
1340 R
1341 F
1342 R
1343 F
1344 R
1345 F
1346 R
1347 F
1348 R
1349 F
1350 R
1351 F
1352 R
1353 F
1354 R
1355 F
1356 R
1357 F
1358 R
1359 F
1360 R
1361 F
1362 R
1363 F
1364 R
1365 F
1366 R
1367 F
1368 R
1369 F
1370 R
1371 F
1372 R
1373 F
1374 R
1375 F
1376 R
1377 F
1378 R
1379 F
1380 R
1381 F
1382 R
1383 F
1384 R
1385 F
1386 R
1387 F
1388 R
1389 F
1390 R
1391 F
1392 R
1393 F
1394 R
1395 F
1396 R
1397 F
1398 R
1399 F
1400 R
1401 F
1402 R
1403 F
1404 R
1405 F
1406 R
1407 F
1408 R
1409 F
1410 R
1411 F
1412 R
1413 F
1414 R
1415 F
1416 R
1417 F
1418 R
1419 F
1420 R
1421 F
1422 R
1423 F
1424 R
1425 F
1426 R
1427 F
1428 R
1429 F
1430 R
1431 F
1432 R
1433 F
1434 R
1435 F
1436 R
1437 F
1438 R
1439 F
1440 R
1441 F
1442 R
1443 F
1444 R
1445 F
1446 R
1447 F
1448 R
1449 F
1450 R
1451 F
1452 R
1453 F
1454 R
1455 F
1456 R
1457 F
1458 R
1459 F
1460 R
1461 F
1462 R
1463 F
1464 R
1465 F
1466 R
1467 F
1468 R
1469 F
1470 R
1471 F
1472 R
1473 F
1474 R
1475 F
1476 R
1477 F
1478 R
1479 F
1480 R
1481 F
1482 R
1483 F
1484 R
1485 F
1486 R
1487 F
1488 R
1489 F
1490 R
1491 F
1492 R
1493 F
1494 R
1495 F
1496 R
1497 F
1498 R
1499 F
1500 R
1501 F
1502 R
1503 F
1504 R
1505 F
1506 R
1507 F
1508 R
1509 F
1510 R
1511 F
1512 R
1513 F
1514 R
1515 F
1516 R
1517 F
1518 R
1519 F
1520 R
1521 F
1522 R
1523 F
1524 R
1525 F
1526 R
1527 F
1528 R
1529 F
1530 R
1531 F
1532 R
1533 F
1534 R
1535 F
1536 R
1537 F
1538 R
1539 F
1540 R
1541 F
1542 R
1543 F
1544 R
1545 F
1546 R
1547 F
1548 R
1549 F
1550 R
1551 F
1552 R
1553 F
1554 R
1555 F
1556 R
1557 F
1558 R
1559 F
1560 R
1561 F
1562 R
1563 F
1564 R
1565 F
1566 R
1567 F
1568 R
1569 F
1570 R
1571 F
1572 R
1573 F
1574 R
1575 F
1576 R
1577 F
1578 R
1579 F
1580 R
1581 F
1582 R
1583 F
1584 R
1585 F
1586 R
1587 F
1588 R
1589 F
1590 R
1591 F
1592 R
1593 F
1594 R
1595 F
1596 R
1597 F
1598 R
1599 F
1600 R
1601 F
1602 R
1603 F
1604 R
1605 F
1606 R
1607 F
1608 R
1609 F
1610 R
1611 F
1612 R
1613 F
1614 R
1615 F
1616 R
1617 F
1618 R
1619 F
1620 R
1621 F
1622 R
1623 F
1624 R
1625 F
1626 R
1627 F
1628 R
1629 F
1630 R
1631 F
1632 R
1633 F
1634 R
1635 F
1636 R
1637 F
1638 R
1639 F
1640 R
1641 F
1642 R
1643 F
1644 R
1645 F
1646 R
1647 F
1648 R
1649 F
1650 R
1651 F
1652 R
1653 F
1654 R
1655 F
1656 R
1657 F
1658 R
1659 F
1660 R
1661 F
1662 R
1663 F
1664 R
1665 F
1666 R
1667 F
1668 R
1669 F
1670 R
1671 F
1672 R
1673 F
1674 R
1675 F
1676 R
1677 F
1678 R
1679 F
1680 R
16

1
2
3 >GFIK01043228.1
4 -----RPFFCFNMSSFIACDLRAPAGRKHRGYRL YI-PRLF
5 ENGYIL-----PAKYDSVLNGCFL-----WKRKNLTNVGRG
6 LVIQVL-----CRDF--GFSKRFAEVFASRNIS-----EIRRIEQTIHAVLDC
7 LLL---YDTKLFIENRY---YSLLKHHIVRRIMSGTF---SIGLVISYWKDFTNDIWMTV
8 SEAE-----YQGEISIDKRNFFYEIERVLCCLKSLKGEKLDKV-----QLQRF
9 AHLTS-TRGLPCGDKQSEKRAILKFKEHVTPTQ---VSDYVVE-----DY
10 GRAAGLIARKCARRFEG-----FVRAAHISLATAGSLDTVSEGGRGEEIRR----
11 AIFPMLTFIPKESSIVKTPF-D---LNEVAG----VERWR--TWCRPEPYMDF----
12 -----PDKAFTGTPNGETLGGCE-----
13 -----VFIQGFDEAIGEQILACAYFA-----YQQDK
14 G-----DQVEILTRVLTIEEPGGKSRIVTAPWYVYVLEQ
15 PVAHIMRLVVSNHPSAESGLTRSDQAWAYLH-----LLERAKDHLQD--DF----
16 KGLSSDLVASTDLIPKP-ILEVLYTEFNKG AHL-----VGPLFEISLDLL---RRQR
17 FC--LFRD-G---SNCLTARGIFMGEPLTKALLLYN-----LCTEEIAIRKYLS--
18 -----YDFIQPVKVPW-RCF--AVAG-----DDHIAV-----GP
19 VEYLRLITQTHLAAEAQLSAGKHAVSSVALRYCEKMLD-LRKVKS-DWKV-KTINSSLEN
20 YL--NSPFVDSVKVRLLSPCSK---TNES-FNDRNTAIGKGKSLARTLSWLNSTL--FP
21 TKWVSMVRDRFFSRMGPLLPSRSSG--VYWHLLLPTLGGLGLGCP-----EDIPDL
22 LGKLPTLSKSFL--SEIAEGKITD--LYRAFRSLLSNTYRGYEMATTEIELAKS----
23 -----LLFDISK-----WLFPCKSMRELRE--SYPEIE-----EKPV
24 QNQLSYLM---KQG-W---MRYDRLED-----TFLRGFLFKEILT----
25 -----
26 --GNVVVKAFNSQKIKKRYSD-LWDR-FY-----DGDCVL--TEEQLKKAWKFRDFF
27 D-----FYQVKNEVKAWIRK-----KLVCTA-LQE
28 LTYGLPD--LTIRPSLIAMRYHDEELRDEDKVLTEALKASDITTWEDAITHEVLKDL---
29 --KAKL-----DSESQVEPKEALTNIVSESCDKKEG-----
30 >KX883482.1
31 -----MQSTNLNTSKLIQTGCPRAPAGKLGRGYRIFI-PHY
32 EN-----GDYLLPGRNHSLL-----KGVVVRNRMVKP
33 NLLISK-----LTLM--GFGKRFATLVSLRPIS-----HIERAEQSVLGVIDS
34 LLL---YDSELFITDTG---YKLLKFIVKKTFKVAIF---NCALVTKLWKEFSFYVWAHV
35 TETV-----TREKPVRKRENFFSLVSEQDI-EVLKLTLNK-----KTLTRL
36 AHLTS-SRQFPPGERQQEIKSLKDFESHVTEPYI---GNP-----VF
37 LRRIRLASRVVGRKARR----LGMKPLTDHSLSAAAGSFYTSVKDGGRAEELLN---
38 MLFKYLAYVPKESCEIKTPFY-T---LKDVEG----VERWR--TWCRPAVYDHF----
39 -----PNVSFGNL---LPETL-----
40 -----MGFETYQQFDEILGMQILACSYL-----QEEHL
41 K-----CNQNIPVRVLTITEPGSKARIVTGPCWLYTLQQ
42 PVSHVLRGFLGQHPSAAAGLTMSDQAQWQFLF-----LLEKAKSHFGD--DF----
43 SCLSSDLTAATDVIPLC-ILKELYEGFLEGRLI-----EGPFLNIVGKMI---QLPR
44 LC--SVEKIN----SYFLNSRGIFMGEPIAKVLLTLLN-----LSCEEIAIRDYLGV-
45 -----DFRASV-QVRW-RCF--SVAG-----DDHIAV-----GP
46 VGYLDKITETHIKAGSKISDLKHGHHSRIA VRYCEKILD-VRNFKG-SWSK-YTINDSTEA
47 YL--ASPFVDSIKVRLSPA AK---NVLS-FNEKNTAFGKGKSLGRTLQWLNKDC--FD
48 SKWISLVRDWFFKRMSSLLPDRSSG--VYWHLLLPEHLGGLGLGTE-----RDFEDL
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 IVRLPSPSRTLL---KSIEDGNPNM-GHIRLFSGFTSNTTYRGYKIPETEKEIAAL---
4 -----FFQGI-----LEEKSDKFENLVKLYRYE-----DLIP
5 NTAIRLLR---TQG-W---VRKDYLED-----KILRPFLFKEILT-----
6 -----
7 -----
8 --GRAKVAFNTEHLKSRYAR-LWDL-----TYRGEETI--SPQLQVCFKPAYH
9 L-----YYKMGEKLELPFRGRV-----LKVNLL-LEE
10 GLLGMPD--LSI-----PWEVIGDISGP
11 LAPKRG-----NHARSRIKSPRRRNS-----
12 >GFLH01017038.1
13 -----SIGTLMHGSQTAHLGTRVGSIHWTVVL-PQPH
14 DVGLKLR-----TSSKPFERISFKGEQI-----SSKKVFGKEKGRS
15 GLRAYL-----CTTI--GIVKHAEVIVSRKTS-----EIRRMRQWESIIDN
16 FLL---NPKFFELRF---QPAWKVLR AFLGVAPF---SVQCTTRMYKEFVEFLKQEY
17 FEPL-----EGEKISLPKDNLFKPLV-----EVVRSVLNKD---RKTDYVHF
18 AHICT-TRNLAPGDAKEQDNSIREFILETTTEFE---VSK-----LD
19 LSNLSLAMDRIVSICLN---SYESGGYLGAHMSISESAEYNVPQS KGGKA AAILA---
20 EAGEILNRIADKDQKVETPFR-S---YTHRQG---YPLWM---CI-----
21 -----ADVGVKDPV-EWLSPKE-----
22 -----MTIVDRPPPREGLDKTTGEQLFVA AWIA-----F
23 N-----ESRSIKVRAVSVPEPGGKSRVVTVSQWWLPVIQQ
24 PLSHLLSGVLSKHPSAYSFRRQDQA WQLSK-----LASRIKHTQTMS-GF-----
25 AWLHSDFTSATDAVPHK-VAYCIMERLIMRC-----FPNFKPVLRI---LGER
26 EV--EVKK-----DTFFT KRGIFMGEFSKLVLVGSI-----LAIEELCYSIEKSIS
27 L-----RRTFTK-PSNW-RFM--HVGG-----DDHRCY-----GP
28 VSYLKRLTRLCKQFGYIVSPTKHIMTRRAGMYTEKCIY-YHDRVI-NMKV-QQIND---N
29 VE--KSVWVDSIKLRLSPFTK---PLDG-RDDRNI AIGKAKTLSRSIQWFPMGT--EL
30 ETIARLAIERYKLRFSTMIPSSNRK-TMTAIIRLPDSVGGLGLVYQ-----PEVED
31 FRNLPEIFWWAL---NVIHAGHPKNYKVRSLQ--NTWQNFSGYEQ---LKYLS---
32 -----DFKSQI-----EDYPNMVG-VTFNALKE--EFNEHF-----NGDN
33 TEFVDFLR---EKN-I---IQIHEFMK-----LIERPHKFYKLMT-----
34 -----
35 -----RSRIKHFNLDNRTRIAR-TWEA-LENLKSEITDPIPLIK--EMEVIQGAAKLARHD
36 W-----FIKLDEPSEAAFISTE-LSEKQLDDEEYFTFMLDG TISMRTIPF-SEH
37 ILYNAPE--LKL-----
38 -----HNM-----KIQN-----
39 >HACV01002003.1
40 ILNIETGIPRDPAGVQDCSEGSSTLRLGSPSES VRAPLGPVPSTKSRQWGRGLLV-PTLS
41 PDSVLQN-----PVRPIYEVKWGGVSLSG-----GITAYGSRKWRQ
42 RLRKYL-----PTVW--GVSVRTAGILSGRSRS-----YLIQVFRNLDLSLIDN
43 LML---FDRTLFDQNNAES-RSVITRFVRKVLT VSAF---SIGDCIAKWKDFVNHFVTVS
44 SDATV-----SSTVSLKGNCFKFLLNE---GKFKRLVND---KPSTREDYMDI
45 AHVIS-TRNLANGGVKAQQAAVSQFRKT STEFK---VSP-----EN
46 LKKLRLVASRLGRIVTS----IRGSNKLRGHCSLNGSGTLDVPVNQGGRATDALL---
47 DLKKFLSHIPPSSGESRTYPWG-E---VWFPSG-----IPRWC--AFGKLENF-----
48 -----TDAEWMSLANREIIDRG-----
49 -----ALRNGCDSTTGRMVMAVAYEC-----YRPFE
50
51
52
53
54
55
56
57
58
59
60

1
2
3 E-----SGTPIPIRQAAVSEPGGKVRMVTGPPWWLAVLQQ
4 PLCHGLREILAYHPSAHSCLMRCDQAWQLSH-----VLERLGISKLED-GM-----
5 AVLSSDLKEATDAIPHS-VGKVLLSGFLEGLGR-----EDHDWITNF----IGER
6 TV--FCED-G---EFTLRRGVMMGEPLSKICLVLLG----LAMEELAFSEHEGLS
7 L-----HRRDTP-RCHW-RAF--HLGG-----DDHLAV----GP
8 VTYLRSITSNHKLFGSVISPVKHRMSSIYVVYTEKVLR-IKGAVL-NLKA-SQVDGNPG-
9 ---NCTFIDSVKVRLSPFSK---ATDS-QNDKNIAIGKVRGIAATLKYFPSES---
10 --IKRTVFDRVAYRFNHYIGSERHR-TIRAVMSLPSELGGLGFAFD-----DRY
11 LGRLPPIFNRAV---RTVINGGGVGFRVRRILGQIFANRCPRGLVSNALMVDDYVE---
12 -----QMIEQF-----LDYPGSVDMVSSGDLWA--RLDENG----LMSF
13 REFLEKAS---HEG-W---ISLHDIPR-----IAEKPFLMRKILE-----
14 -----
15 -----
16 --GSEVPEYFRTASNAQRIAECWNS-LEAIDD-VLDPTGGEL--SPLEREASRRAKHT
17 V-----FVNLMEMTCALVDPT-----SEAYDPGDPWISADFIDLSM-REL
18 LTYGEPSLTVP-----
19 -----
20 -----
21 >GFBP01062652.1
22 -----LTLRRYMPPTQLTAPPGRFGEGGVHSPGLLV-PNLT
23 PDHVVV-----NPVRPIWEHRFGGVCFASLGSK-----QKWRT
24 SFARTL-----RVEL--DMGRRASETLANRGLS-----YIKRVLKNLDSLIDN
25 LIL---FSPSMFDRSTPGVIYDTLKRILRVTLKYSTY---SVEDQISQWKDFVNHFITLV
26 SKSQV-----EKQICLRHNCYKWMLEQ---DPFKSLLSKI--DMTT-KDWTSI
27 AHLIS-TRNFAPGGESAKLKAIRAFKETTSGFS---VSE-----EN
28 LRKLKLATLRVTGISLH---LSEEKRLSGGHVSLNSSGTLDVSVAKGGRAADALL---
29 DLKQFLDEEPVQGETRVYPWG-E---LYYPPG---LRRWR--AFGPLQD-DDG---
30 -----PFMSKS-R--YDLI-----
31 -----DREAMLNGCGEYTGMIFTVAYEC-----YRVYL
32 N-----SGELIPIRQATVSEPGGKARIVTGPPWWLAVLQQ
33 PLCHGLREILGYHPSAHSCLRADQAWQLSH-----VWAKCSDNCIPE-GE-----
34 AVLSSDLKEATDAIPHS-VGQVLIRGCLEAIGA-----TKWLWMADL----IGPR
35 LV--LAED-K----DVFTLRRGVMMGEPLSKICLILLG----LTMEEIAFSEYSSL
36 L-----KRKFTP-AQGW-RAF--HLGG-----DDHLAK----GP
37 MRYLKSITSYHRLFGSIVSPYKHRISVRAVTYTEKILV-FENRIL-NMPV-EQVNSNID-
38 ---KSIFDSIKVRLSPFTK---ATDT-MNDKNVAIGKCIGIAGTLQWMHDRS---
39 --LKRTLIDRIHYRFRDFIAGPHHR-TLSAIQLPTTLGGGLGSLVD-----TKY
40 LENLPEPYNRAI--REIIKGGVLAYKARMVMNSTFVNKNLRGIKS---DDYVT---
41 -----ELITQW-----VEYPSSVNAMEKFKEAIA--VVDPEG----NNSF
42 RHNLWLLK---KRN-I---VSMNDIPQ-----IAESAYVMRRLL-----
43 -----
44 --QSQSGKYFRTERLDRVSK-TWEG-LTKLDPESGDVLSED--ELKTGTRLAKQALFI
45 D-----LTTMGTIAVADSTNPD-----YLPDDPWLAEEFIDIPF-GQS
46 IKYGQPN-----MGIDLDH
47 PFPEEK-----DLAC-----
48 -----
49 -----
50 -----
51 -----
52 -----
53 -----
54 -----
55 -----
56 -----
57 -----
58 -----
59 -----
60 -----