A picture containing table

Description automatically generated

**Supplemental Figure 1**. Average Nucleotide Identity (ANI) heatmap comparing AB26, AB19, and AB28 to other closely related *Rhodovulum* strains. The heatmap shows AB19 and AB26 share >99% ANI to each other, and 98% ANI to *Rhodovulum visakhapatnamense*. AB28 does not show increased ANI values to any related species, despite the relatively low (~94%) ANI to the other Woods Hole isolates (**Figure 2**). ANI values were calculated using JSpeciesWSd and visualized as a heatmap using Morpheus.

**Supplemental Table 1.** General Genome Features of Woods Hole isolates and publicly available *Rhodovulum visakhapatnamense* and *Rhodovulum sulfidophilum* genomes from NCBI and IMG/JGI databases.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Strain** | **Genome size (total bases)** | **Assembly level** | **GC%** | **Genes (total)** | **CDS (protein coding)** | **tRNAs** | **rRNAs** | **GenBank Accession** |
| **R. visakhapatnamense** | | | | | | | | |
| **AB26** | 4,380,746 | Scaffold | 67.9 | 4276 | 4095 | 43 | 3 | GCA\_001941715.1 |
| **AB19** | 4401286 | Contig | 67.8 | 4547 | 4214 | 45 | 3 | GCA\_016757185.1 |
| **JA181** | 4429538 | Scaffold | 67.5 | 4300 | 4243 | 45 | 3 | GCA\_004365965.1 |
| **R. sulfidophilum** | | | | | | | | |
| **AB14** | 4351599 | Scaffold | 66.8 | 4155 | 3972 | 44 | 3 | GCA\_001941695.1 |
| **AB15** | 4536182 | Contig | 66.7 | 4837 | 4586 | 46 | 3 | GCA\_021409265.1 |
| **AB16** | 4406360 | Contig | 66.7 | 4442 | 4237 | 40 | 3 | GCA\_016757145.1 |
| **AB17** | 4527006 | Contig | 66.7 | 4731 | 4511 | 40 | 3 | GCA\_021409205.1 |
| **AB18** | 4349103 | Contig | 66.6 | 4315 | 4078 | 44 | 3 | GCA\_021409185.1 |
| **AB20** | 4234138 | Contig | 66.8 | 4637 | 4473 | 45 | 3 | GCA\_016757195.1 |
| **AB21** | 4073787 | Contig | 66.6 | 5611 | 5365 | 34 | 3 | GCA\_021409225.1 |
| **AB22** | 4569598 | Contig | 67 | 4454 | 4255 | 44 | 3 | GCA\_016757115.1 |
| **AB23** | 4454817 | Contig | 66.8 | 4330 | 4147 | 45 | 3 | GCA\_016757055.1 |
| **AB28** | 5027343 | Contig | 66.9 | 5450 | 5199 | 49 | 3 | GCA\_021409165.1 |
| **AB30** | 4256231 | Contig | 67 | 4095 | 3903 | 43 | 3 | GCA\_001941745.1 |
| **AB33** | 4429719 | Scaffold | 66.4 | 4440 | 4111 | 48 | 3 | GCA\_021568715.1 |
| **AB35** | 4642509 | Contig | 66.9 | 4501 | 4286 | 46 | 3 | GCA\_016757045.1 |
| **DSM1374** | 4347929 | Complete | 66.9 | 4143 | 3944 | 52 | 3 | GCA\_001633165.1 |
| **DSM2351** | 4732772 | Complete | 66.9 | 4376 | 4146 | 50 | 3 | GCA\_001548075.1 |
| **IM796** | 3962372 | Scaffold | 67.3 | 3747 | 3574 | 43 | 3 | GCA\_016653205.1 |
| **SE1** | 4310535 | Contig | 67.1 | 4137 | 3954 | 44 | 3 | GCA\_010119435.1 |
| **SNK001** | 4196493 | Complete | 66.8 | 4027 | 3849 | 52 | 3 | GCA\_001633145.1 |

**Fasta Input Sequences for the 16S Phylogenetic Tree of Woods Hole Isolates**

>dsm2351

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCTTGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTTCTAA

>dsm1374

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>SNK001

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB14

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB26

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCGGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTCCGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCAAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB30

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAAC

>se1

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>IM796

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB35

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCTTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB23

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB22

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>aB16

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTTGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTGCGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB20

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB28

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB19

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAACAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTGTCGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACCATTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGAAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGTCGTGAAGATAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTCCGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCTTGCTTGTCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGAAGCATGCTTGTCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTACAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCTCAGTTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGTTAATCCCAAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCTTACGGGACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAACCTGCGGCTGGATCACCTCCTTTCTAA

>AB18

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB17

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB15

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCTTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB33

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>2\_4\_1

CAACTTGAGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGGAC

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCGGGAAACTGGGAGTAATACCGA

ATGTGCCCTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGGGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTAAA

GATCTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAAGTCAGAGGTGAAATCCCAGGG

CTCAACCCTGGAACTGCCTTTGAAACTCCCGATCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAGAT

GGTTCCTTCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTC

CGGCAACGAGCGCAACCCACGTCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGGAAACTGCCGGTGATAAGCCGGAG

GAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTT

AATCCCAAAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAAC

AGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGCGG

TGCGCCAACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>TIE1

CAACTTGAGAGTTTGATCCTGGCTCAGAGCGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAACGGGCGTAGCAATA

CGTCAGTGGCAGACGGGTGAGTAACGCGTGGGAACGTACCTTTTGGTTCGGAACAACACAGGGAAACTTGTGCTAATACC

GGATAAGCCCTTACGGGGAAAGATTTATCGCCGAAAGATCGGCCCGCGTCTGATTAGCTAGTTGGTGAGGTAATGGCTCA

CCAAGGCGACGATCAGTAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAG

GCAGCAGTGGGGAATATTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCCTAGGGTTGTA

AAGCTCTTTTGTGCGGGAAGATAATGACGGTACCGCAAGAATAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATA

CGAAGGGGGCTAGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTTTCTAAGTCAGAGGTGAAAGCCTGG

AGCTCAACTCCAGAACTGCCTTTGATACTGGAAGTCTTGAGTATGGCAGAGGTGAGTGGAACTGCGAGTGTAGAGGTGAA

ATTCGTAGATATTCGCAAGAACACCAGTGGCGAAGGCGGCTCACTGGGCCATTACTGACGCTGAGGCACGAAAGCGTGGG

GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGCCGTTAGTGGGTTTACTCACTAGTGG

CGCAGCTAACGCTTTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCAC

AAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGCAGAACCTTACCAGCCCTTGACATGTCCAGGACCGGTCGCAGA

GACGCGACCTTCTCTTCGGAGCCTGGAGCACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAG

TCCCGCAACGAGCGCAACCCCCGTCCTTAGTTGCTACCATTTAGTTGAGCACTCTAAGGAGACTGCCGGTGATAAGCCGC

GAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACACACGTGCTACAATGGCGGTGACAATGG

GAAGCTAAGGGGCGACCCTTCGCAAATCTCAAAAAGCCGTCTCAGTTCGGATTGGGCTCTGCAACTCGAGCCCATGAAGT

TGGAATCGCTAGTAATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATG

GGAGTTGGCTTTACCTGAAGACGGTGCGCTAACCAGCAATGGGGGCAGCCGGCCACGGTAGGGTCAGCGACTGGGGTGAA

GTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>geobactersulfurreducens

AAACTGGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGTGATCCATCCTT

CGGGGTGGTGAAAGTGGCGCACGGGTGAGTAACGCGTGGATAATCTGCCCGAGGATTTGGGATAACATCTCGAAAGGGGT

GCTAATACCGAATAAGCCCACGGGGTCTACGGATCTTGCGGGAAAAGGGGGGGACTTTCGGGCCTCCTGTCTTCGGATGA

GTCCGCGTACCATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGATCAGC

CACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCGCAATGGGGGAAACCCTGACG

CAGCAACGCCGCGTGGGTGATGAAGGCCTTCGGGTCGTAAAGCCCTGTCGGGAGGGAAGAAATGATTGAGAGCTAATACC

TCTTGGTCTTGACGGTACCTCCGAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCG

TTGTTCGGAATTATTGGGCGTAAAGCGCGTGTAGGCGGTCTTTTAAGTCTGATGTGAAAGCCCCGGGCTCAACCTGGGAA

GTGCATTGGAAACTGGGAGACTTGAGTACGGGAGAGGAGAGTGGAATTCCTAGTGTAGGAGTGAAATCCGTAGATATTAG

GAGGAACACCGGTGGCGAAGGCGGCTCTCTGGACCGATACTGACGCTGAGACGCGAAAGCGTGGGTAGCAAACAGGATTA

GATACCCTGGTAGTCCACGCCGTAAACGATGAGTACTAGGTGTTGCGGGTATTGACCCCTGCAGTGCCGTAGCTAACGCA

TTAAGTACTCCGCCTGGGAAGTACGGTCGCAAGACTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCA

TGTGGTTTAATTCGACGCAACGCGCAGAACCTTACCTGGGCTTGACATCCACGGAACCCTCCCGAAACGGAGGGGTGCCC

TTCGGGGAGCCGTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGC

GCAACCCTTATCATCAGTTGCCATCATTAAGTTGGGCACTCTGATGAAACTGCCGGTGTCAAACCGGAGGAAGGTGGGGA

TGACGTCAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGCCGGTACAAAGAGTCGCAATGCCGCG

AGGTGGAGCCAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGT

AATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGGGAGTCGACTGGT

CCCGAAGTGCGTGAGCTAACCAGCAATGGAGGCAGCGTCCTAAGGAATGGTCGGTGACTGGGGTGAAGTCGTAACAAGGT

AGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>Pseudomonas\_aeruginosa

GAACTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCAGCAGGGGCCTTCAACACATGCAAGTCGAG

CTTATGAAGGGAGCTTGCCTTGGATTCAGCGGCGGACGGGTGAGTAATGCCTAGGAATCTGCCTGGTAGT

GGGGGATAACGTCCGGAAACGGCCGCTAATACCGCATACGTCCTGAGGGAGAAAGTCGGGGATCTTCGGA

CCTCACGCTATCAGATGAGCCTAGGTCGGATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGACGAT

CCGTAACTGGTCTGAGAGGATGATCAGTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCA

GCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCG

GATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCAGTAAGTTAATACCTTGCTGTTTGACGTTACCAACAG

AATAAGCACCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTAC

TGGGCGTAAAGCGCGCGTAAGTGGTTCAGCAAGCTTGATGTGAAATCCCCGGGCTCAACCTGGGAACTGC

ATCCAAAAGCTACTGAGCTAGAGTACGGTAGAGGTGGTAGAATTTCCTGTGTAGCGGTGAAATGCGTAGA

TATAGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGACTGTACTGACACTGAGGTGCGAAAGCGTGGG

GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTCGACTAGCCGTTGGGATCCTTGA

GATCTTAGTGGCGCACGTAACGCGATAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAA

TGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACC

TGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAACAGAGACACAGGTGCTGCA

TGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGT

TACCAGCACCTCGGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTC

AAGTCATCATGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTCGGTACAAAGGGTTGCCAAGCC

GCGAGTGGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGT

CGGAATCGCTAGTAATCGTGAATCAGAATGTCACGGTGAATACGTCCCCGGGCCTTGTACACACCGCCCG

TCACACCATGGGAGTGGGTTGCTCCAGAAGTAGCTAGTCTAACCGCAAGGGGGACGGTTACCACGGAGTG

ATTCATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTA

>Nitrosococcus\_oceani

AGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCATTGACCACAC

CGTAGTAAGCGTCCTCCCGAAGGTTAAACTACCCACTTCTGGTGCAACCAACTCCCATGGTGTGACGGGC

GGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATTGCTGATCCGCGATTACTAGCGATTCCGACTT

CATGCAGTCGAGTTGCAGACTGCAATCCGGACTACGACCGGTTTTCTGAGATTAGCTCCGCATCGCTGCT

TGGCCACCCTCTGTACCGGCCATTGTAGCACGTGTGTAGCCCTGGCCATAAGGGCCATGATGACTTGACG

TCATCCCCACCTTCCTCCGTTTTGTCAACGGCAGTCCCTCTAGAGTGCCCGGCCGAACCGCTGGCAACTA

GAGATAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCA

GCACCTGTCTCCCGGCTCCCGAAGGCACCTTCGCCTCTCAGCAAAGTTCCGGGGATGTCAAGGCCAGGTA

AGGTTCTTCGCGTTGCATCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCATTTG

AGTTTCAACCTTGCGGCCGTACTCCCCAGGCGGAGAACTTAACGCGTTAGCTGCAACACTGAAAGGCTAT

CCCTCCCAACGTCTAGTTCTCATCGTTTAGGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCA

CGCTTTCGCACCTCAGCGTCAGTAGTGGTCCAGGAGGCCGCCTTCGCCACTGGTGTTCCTTCCGATATCT

ACGCATTTCACCGCTACACCGGAAATTCCGCCTCCCTCTACCATACTCTAGCCAGGCAGTATCAAGTGCC

ATTCCCAGGTTGAGCCCGGGGCTTTCACACCCAACTTACCAAACCGCCTACGCGCCCTTTACGCCCAGTA

ATTCCGATTAACGCTCGCACCCTCCGTATTACCGCGGCTGCTGGCACGGAGTTAGCCGGTGCTTCTTCTG

TAGGTAACGTCAACACGATGTGCTATTCACACATCGCTTTTCTTCCCCACTGAAAGGGCTTTACAACCCG

AAAGCCTTCTTCACCCACGCGGCATTGCTGGATCAGGGTTGCCCCCATTGTCCAATATTCCCCACTGCTG

CCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCTGATCGTCCTCTCAGACCAGCTACCGAT

CATCGCCTTGGTAGGCCCTTACCCCACCAACTAGCTAATCGGACGCAGGCTCATCCCCCAGCGCGAGGCC

TTTCGGTCCCCCGCTTTCCTCCTTAGAGATTATGCGGTATTAGCCCGAGTTTCCCCGAGTTATCCCCCTC

TAGAGGCCAGATTCCCACGCGTTACTCACCCGTCCGCCACTCGCCAGCCACCTAAGCAAGCTTAGGTGCT

GCTGCCGTTCGACTTGCATGTGTTAAGCATGCCGCCAGCGTTCAATCTGAGCCATGATCAAACTCTTCAG

TTTAA

**Input Fasta file for the PufM Woods Hole Isolates Phylogenetic Tree**

>AB35

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>DSM2351

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLVSVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWFIWAQEHNFVTEYTQPYGVDAYVGQGG

>DSM1374

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>IM796

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>SE1

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>SNK001

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB14

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB15

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB16

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB17

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB18

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB20

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB21

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB22

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB23

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB26

MAEYQNIFTQVQVQGPGDWGMDNENNMMEERAAGTGHFSILGWL

GNAQLGPIYLGYTGVISVIAGAFAFLIIGLNMLAQVDWSLVQFLRQGFWLALEPPSPE

YGLRIPPLKEGGWYMIASFFLLIAVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMV

LGFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA

MHGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVL

TPLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGIDAYVGQGG

>AB28

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB30

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>AB33

MAEYQNIFTQVQVQGPSDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY

GLRIPPLKEGGWYMVASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFMVL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFAM

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWFAVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAYVGQGG

>TIE\_1

MAQYQNIFTQVQVEGPAYAGVPLRPGSSPRETQTTFNYWLGKIGDAQVGPVYLGFTGVCSLLCGFVAIEI

IGLNMLASVDWSPIEFLRQFCWLALEPPKPEYGLTIPPLKEGGWWLMAGFFLTVSIALWWVRTYRRSRAL

GMGTHVSWAFASAILLYLALGFIQPLLMGSWSEAPPFGVFPHLDWTNNFSIKYGNLYYNPFHCLSIAFLY

GSALLFAMHGATILAVSRYGGEREIEQMLDRGTALERAALFWRWTMGFNATAESIHRWAWWFAVLCPLTG

AIGIILTGPVVDNWFDWGVKHGLAPPR

>Chloroflexus\_aggregans

MATMDMMPGDLELGRDRGRIGKPIEIPLLENFGFDSQLGPFYLGFWNTVAYISGGIFTFIWLFVMLSQVN

FNPVAFIKYFVVLQIDPPSSRYGLSFPPLNEGGWWLIATFFLTLSIFAWYMHIYTRAKALGIKPYLAYGF

TGAIALYLVIYIIRPVWMGDWSEAPAHGIKALLDWTNNVSVRYGNFYYNPFHMLSIFFLLGSTLLLAMHA

GTIWALEKYAAHEEWNEIQAPGTGTERAQLFWRWCIGFNANAYSIHLWAFWFAWLCGITGALGVFFSMPD

FVNNWFQWGVEAGIAYPQGPTPLP

**Fasta File Input for the Woods Whole EeuP Phylogenetic Tree**

>AB26

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAIFARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNTLGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLRGAFPRSGDGSALETAAAPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRALLEYLKSI

>AB33

MFTRSVIFSLLFSGAAAPVTAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLTWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIFSRPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNALGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQQDDKFMAEFDDLRGAFPRSGDGSALETAAAPCAYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTPGPAYNLEAVGMATEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRALLEYLKSI

>AB22

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAIFARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNTLGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLRGAFPRSGDGSALETAAAPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRALLEYLKSI

>AB35

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQD

QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTPNEGMGGTA

IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDEATFDAFADK

VAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRP

EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG

EFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIF

ARPTDQGGCVECHGKRQGAFRSIFHSTWATPILPVGTDERECEILTRTAKTGVLEGAK

VPLSSNTLGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDDKFMAEF

DDLRGAFPRSGDGSALETAAEPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER

VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA

LLEYLKSI

>AB20

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQD

QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTA

IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDEAVFDAFADK

VAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRP

EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG

EFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAIF

ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTARTGVLEGAK

VPLSSNTLGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDDKFMAEF

DDLRGAFPRSGDGSALETAAAPCAYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER

VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA

LLEYLKSI

>IM796

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQD

QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTA

IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADK

VAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRP

EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG

EFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIF

ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAK

VPLSSNALGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDDKFMAEF

DDLRGAFPRSGDGSALETAAAPCAYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER

VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA

LLEYLKSI

>AB30

MNEKIIPILFWEESMFTRSVIFSLLFSGAAAPVTAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLTWME

ALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDG

GPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGETGDRAGLRAELETWSLRFHTLIDRSLPDPGWG

PVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLAR

NLGEVYGVFGEFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIFAR

PTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNTLGAEAPA

FDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQQDDKFMAEFDDLRGAFPRSGDGSALETAAEPCAYE

ARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSS

GNSRCGHDYGTALSDDDKRALLEYLKSI

>AB23

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQD

QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTA

IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQNFLADLDAAMQAVLDDDATFDAFADK

VAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGLDLGRP

EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG

EFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAIF

ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAK

VPLSSNTLGAEAPAFDLLGASVIGAIIQNAVGFRVTGGLEATGAAAALQKDGKFMAEF

DDLRGAFPRSGDGSALETAAEPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER

VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA

LLEYLKSI

>AB16

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQD

QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTA

IGMTCAACHTRQIEVSGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDEAVFDAFADK

VAGETGDRAGLRAELETWSLRFHTLIDRSLPDPGWGPVRLDAVSMIFNRLTGLDLGRP

EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG

EFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIF

ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAK

VPLSSNALGAEAPAFDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDGKFMAEF

DDLRGAFPRSGDGSALETAAAPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER

VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDGDKRA

LLEYLKSI

>AB14

MNEKITPILFWEESMFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWME

ALAQPDGTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDG

GPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWG

PVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLAR

NLGEVYGVFGEFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDARLVEQGGAIFAR

PTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTARTGVLEGAKVPLSSNALGAEAPA

FDLLGASVIGAIIQHAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLRGAFPRSGDGSALETAAEPCAYE

ARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSS

GNSRCGHDYGTALSDDDKRALLEYLKSI

>AB15

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPDGTPFLADG

LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADL

DAAMQAVLDDEATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRL

TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHP

IKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIFARPTDQGGCVECHGKR

QGAFRSIFHSTWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNTLGAEAPAFDLLGASVIGAIIQ

HAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLRGAFPRSGDGSALETAAEPCKYEARVLDGIWAAAPYL

HNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALS

DDDKRALLEYLKSI

>AB17

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPDGTPFLADG

LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADL

DAAMQAVLDDDATFDAFADKVAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRL

TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHP

IKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAIFARPTDQGGCVECHGKR

QGAFRSIFHATWATPILHVGTDERECEILTRTARTGVLEGAKVPLSSNALGAEAPAFDLLGASVIGAIIQ

NAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLRGAFPRSGDGSALETAAEPCKYEARVLDGIWAAAPYL

HNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALS

DDDKRALLEYLKSR

>AB18

MFTRSVIFSLLFSGAAAPVTAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLTWMEALAQPDGTPFLADG

LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADL

DAAMQAVLDDDATFDAFADKVAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRL

TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHP

IKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIFSRPTDQGGCVECHGKR

QGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNALGAEAPAFDLLGASVIGAIIQ

HAVGFRVTGGLEATGAAAALQQDDKFMAEFDDLRGAFPRSGDGSALETAAAPCAYEARVLDGIWAAAPYL

HNGSVPTLRELLTPPDERVADFTPGPAYNLEAVGMATEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALS

DDDKRALLEYLKSI

>AB28

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWMEALAQTDGTPFLADG

LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADL

DAAMQAVLDDEAVFDAFADKLAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRL

TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFGEFHP

IKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAIFARPTDQGGCVECHGKR

QGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNALGAEAPASDLLGASVIGAIIQ

HAAGFRVTGGLEATGAAAALQKDDKFMAEFDDLRGAFPRSGDGSALETAAAPCAYEARVLDGIWAAAPYL

HNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALS

DDDKRALLEYLKSI

**Input Fasta File for the 16S New Taxonomic Reassignment Tree**

>Rhodovulum\_viride\_JA756

ATGCAGTCGAGCGACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGC

TCGGGAAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGG

TAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGAC

ACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAG

CGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTC

CGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTAT

TAAGTCGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTCCGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTG

GAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGA

CGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGT

CGGCAAGCATGCTTGTCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAA

AGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGA

CATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCT

CGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTA

AGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACG

TGCTACAATGGCAGTGACAATGGGTTAATCCCAAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA

GTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCA

TGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCTACGGGGGCAGCTGCCA

>Rhodobacter\_viridis\_JA737

CTGCCATGCGGTCGAGCGAGACCTTCGGGTCTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTACGGA

ATAGCCCCGGGAAACTGGGAGTAATACCGTATGTGCCCTTCGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGG

ATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGACGATCCATAGCTGGTTTGAGAGGAGATCAGCCACACTGGGACT

GAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTAGACAATGGGGGAAACCCTGATCTAGCCATGCCGC

GTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCT

AACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGG

ATCAGAAAGTCAGAGGTGAAATCCCAGGGCTCAACCTTGGAACTGCCTTTGAAACTCCTGGTCTTGAGGTCGAGAGAGGT

GAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGAT

ACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCA

GTCGTCGGCAGGCATGCCTGTCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAA

CTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACC

CTTGACATCGAGATCGCGGTTACCAGAGATGGTTTCCTTCAGTTCGGCTGGATCTTAGACAGGTGCTGCATGGCTGTCGT

CAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTTTCAGTTGCCATCATTCAGTTGGGCA

CTCTGGAAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTAC

ACACGTGCTACAATGGTGGTGACAATGGGCCAATCCCAAAAAGCCATCTCAGTTCGGATTGGGGTCTGCAACTCGACCCC

ATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCA

CACCATGGGAATTGGGTCTACCCTAAGATGGTGCGCCAACCTGCAAAGGAGGCAGCCAGCCACGGTAGGCTCAGTGACTG

GGGTGAAGTCTAACCAACGTGCCCCACCCCC

>Rhodovulum\_robiginosum\_N2

GAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGTCTTTCGGACTTAGCGGCGGACGGGTGAGTAACGCGTGG

GAACGTGCCCTTCACTACGGAATAGTCCCGGGAAACTGGGTTTAATACCGTATACNCCCTTCGGGGGAAAGAATTTCGGT

GAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGG

ATGATCAACCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGCA

ACCCTGATCCAACCATTCCGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCGCCTGGGAAGATAATGACGGTA

CCAGGTAAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGG

GCGTAAAGCGCGCGTAGGCGGACTGTTAAGTCAGAGGTGAAATCCCAGGGCTCAACCCTGGAACGGCCTTTGATACTGGC

AGTCTAGAGGTCTGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCG

AAGGCGGCTCACTGGCCAGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA

CGCCGTAAACGATGAATGCCAGACGTCGGGCAGCATGCTGTTCGGTGTCACACCTAACGGATTAAGCATTCCGCCTGGGG

AGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCA

ACGCGCAGAACCTTACCAACCCTTGACATGGGTATCGCGGTTTTCCGAGAGGGATTCCTTCAGTTCGGCTGGATACCACA

CAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGACGCGAACCCACACCCTTAG

TTGCCAGCATTCAGTTGGGCACTCTAGGGGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCAT

GGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCATCTACAGTGGGTTAATCCCCAAAAGATGTCTCAGTTCGGAT

TGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGG

CCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTCACCCGAAGGCCGTGCGCCAACCTTTCGAGGGGGCAGCGGA

CCACGGTGAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAA

>Rhodovulum\_iodosum\_N1

GAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGACTTTCGGTCTTAGCGGCGGACGGGTGAGTAACACGTGG

GAACATGCCCTTCACTACGGAATAGTCCTGGGAAACCGGGTTTAATACCGTATACNCCCTTCGGGGGAAAGAATTTCGGT

GAAGGATTGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGACGATCAGTAGCTGGTTTGAGAGG

ATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAA

ACCCTGATCCAACCATGCCGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCACCGGTGAAGATAATGACGGTA

GCCGGAGAAGAAACCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGTTAGCGTTGTTCGGAATTACTGG

GCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTTTGATACTGGT

AGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCG

AAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA

CGCCGTAAACGATGAATGCCAGACGTCGGGCAGCATGCTGTTCGGTGTCACACCTAACGGATTAAGCATTCCGCCTGGGG

AGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCA

ACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTCTCGAGAGAGATTCCTTCAGTTCGGCTGGATCANCNN

CAGGGGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAG

TTGCCAGCAGTTCGGCTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCA

TGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCCAAAAACTGTCTCAGTTCGGA

TTGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGG

GCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCTTTCGAGGGGGCAGCTG

GCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAA

>Rhodovulum\_steppense\_DSM21153

TAACACATGCAAGTCGAGCGAGGACTTCGGTCCTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTGTGCG

GAATAGCCTCGGGAAACTGAGAGTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGCAGAAGGATCGGCCCGCGTC

TGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGG

ACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGCAACCCTGATCCAGCCATGC

CGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCG

GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGG

CGGACTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCCTTGATACTGGTAGTCTAGAGGTCGAGAGA

GGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTC

GATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATG

CCAGACGTCGGCAGGCATGCCTGTCGGTGTCACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTA

AAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCA

ACCCTTGACATGTGTATCGCGGTTCCAGAGATGGTTCCTTCAGTTCGGCTGGATACAACACAGGTGCTGCATGGCTGTCG

TCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGC

ACTCTAGGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTA

CACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCAAAAAGCTGTCTCAGTTCGGATTGTCCTCTGCAACTCGAGGG

CATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTC

ACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCTTTCGAGGAGGCAGCTGGCCACGGTAA

>Rhodovulum\_bhavnagarense\_JA738

ACCATGGCAGTCGAGCGAAGACTTCGGTCTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTACGGAA

CAGTCCCGGGAAACTGGGTTTAATACCGTATACGCCCTTCGGGGGAAAGAATTTCGGCGAAGGATCGGCCCGCGTCTGAT

TAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTG

AGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGCAACCCTGATCCAGCCATGCCGCG

TGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCGCCAGGGATGATAATGACAGTACCTGGTAAAGAAACCCCGGCTA

ACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGTTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGG

CTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCCTTGATACTGGTAGTCTAGAGATCTGGAGAGGTG

AGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAAGAACACCAGTGGCGAAGGCGGCTCACTGGCCAGATA

CTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAG

TCGTCGGGTAGCATGCTATTCGGTGACACACCTAACGGAGTAAGCGTTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAAC

TCACAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCRACGCGCAGAACCTTACCAACCC

TTGACATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTC

AGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACCCTTAGTTGCCAGCATTCAGTTGGGCAC

TCTAAGGGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACA

CACGTGCTACAATGGCAGTGACAATGGGTTAATCCCCAAAAACTGTCTCAGTTCGGATTGTCCTCTGCAACTCGAGGGCA

TGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCAC

ACCATGGGAGTTGGTTACCCGAAGACGGTGCGCCAACCTTCGAGGAGGCAGCTGGCAACGTAA

>Rhodovulum\_euryhalinum\_DSM4868

AACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAGGACTTCGGTCCTAGCGGCGGACGGGTGAGTAACGCGT

GGGAACGTGCCCAAAGGTACGGAATAGCCTCGGGAAACTGAGAGTAATACCGTATGAGCCCTTCGGGGGAAAGATTTATC

GCCTTTGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTNAG

AGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGG

GAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGTCGTGAAGATAATGACG

GTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTAC

TGGGCGTAAAGCGCGCGTAGGCGGATTGGAAAGTTGGGGGTGAAATCCCAGGGCTCAACCCTGGAACTGCCTCNAAAACT

ACCAGTCTAGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTG

GCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGT

CCACGCCGTAAACGATGAATGCCAGACGTCGGGCAGCATGCTGTTCGGTGTCACACCTAACGGATTAAGCATTCCGCCTG

GGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAA

GCAACGCGCAGAACCTTACCAACCCTTGACATGTGTATCGCGGTTCCAGAGATGGTTCCTTCAGTTCGGCTGGATACAAC

ACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTA

GTTGCCAGANTAAGTTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCAT

GGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCNCCAAAAGCTGTCTCAGTTCGGAT

TGTCCTCTGCAACTCGAGGGCATGAAGTNGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGG

CCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCTGTTCGCAGGAGGCAGCC

GGCCACGGTAAGCTCAGCGACTGGGGTG

>Rhodobacter\_azotoformans\_KA25

CTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGGACTTAGCGGCGGACGGGTGA

GTAACGCGTGGGAACATGCCCAAAGGTACGGAATAGCCCCGGGAAACTGGGAGTAATACCGTATGTGCCCTTCGGGGGAA

AGATTTATCGCCTTTGGATTGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGACGATCCATAGC

TGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTAG

ACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTAAAGATCTTTCAGGTGGGAAG

ATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTATT

CGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTGGAAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCC

TTTGAAACTCCCAGTCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGA

ACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATAC

CCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACACACCTAACGGATTAAGCA

TTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTT

TAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAGATGGTTCCTTCAGTTCGGCT

GGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCC

ACGTCCTCAGTTGCCAGCATTCAGTTGGGCACTCTGGGGAAACTGCCGGTGATAAGCCGGAGGAAGGTGTGGATGACGTC

AAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCAAAAAGCTGTCT

CAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATA

CGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGCGGTGCGCCAACCTCGCAAGA

GGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGGTGAAGTCGTACAAAGGTAACCA

>Rhodovulum\_imhoffii\_DSM18064

NANCCGGCTCAGAACGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGGCTTCGGCCTTAGCGGCGGACGGG

TGAGTAACGCGTGGGAACGTGCCCTTCACTACGGGATAGTCCTGGGAAACTGGGTTTAATACCGTATACGCCCTTTTGGG

GAAAGAATTTCGGTGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGACGATCAGTA

GCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGAKTCCTACGGGAGGCAGCAGTGAGGAATCTT

GGACAATGGGGGCAACCCTGATCCAGCCATGCCGCGTGAACGATGAAGGCTTTAGGGTTGTAAAGTTCTTTCACCGGGAA

GATAATGACGGTAGCCGGAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGT

TCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGC

CTTTGATACTGGTAGTCTAKAGTTCTGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGG

AACACCAGTGGCGAAGGCGGCTCACTGGCCAGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATA

CCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGASGTCGGGCASCATGCTGTTCGGTGTCCACCTAACGGATTAAGCA

TTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTT

TAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGG

CTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAAC

CCACGTCTTTAGTTGCCAGCAGTTCGGCTGGGCACTCTAAAGAAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGAC

GTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCGATGACAATGGGTTAATCCCTAAAAGTCG

TCTCAGTTCGGATTGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGA

ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCTGCTC

GCAGGAGGCAGCCGGCCACGGTADGTCAGGAKGGGGGA

>Roseibaca\_ekhonensis\_CECT\_7235

TGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGCCTTCGGGCTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTG

CCCTTTGCTACGGAATAGTCCCTGGAAACGGGGTTTAATACCGTATGTGCTCTTCGGAGGAAAGAATTTCGGCAAAGGAT

CGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAAAGGCCTACCAAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCA

GCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTAGACAATGGGGGCAACCCTGA

TCTAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCGCTGGGGAAGATAATGACGGTACCCAGTA

AAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAA

GCGCGCGTAGGCGGATTGGAAAGTTGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTCCAAAACTATCAGTCTTG

AGTTCGAGAGAGGTGAGTGGAATTGCGAGTGTAGAGGTGAAATTCGTAGATATTCGCAGGAACACCAGTGGCGAAGGCGG

CTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA

AACGATGAATGCCAGACGTCGGGCAGCATGCTGTTCGGTGTCACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGG

TCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCA

GAACCTTACCAACCCTTGACATGGATATCGCGGGACCGGAGACGGTTCTTTCAGTTCGGCTGGATATCACACAGGTGCTG

CATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACGTCCCCAGTTGCCAGCA

TTCAGTTGGGCACTCTGAGGAGACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTAC

GGGTTGGGCTACACACGTGCTACAATGGTGGTGACAGAGGGATAATCCCTAAAAGCCATCTCAGTTCGGATTGGGGTCTG

CAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGGAACAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACA

CACCGCCCGTCACACCATGGGAGTTGGGTCTACCCGAAGGTGGTGCGCTAACCAGCAATGGAGGCAGCCAACCACGGTAG

GCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCC

>Rhodovulum\_marinum\_DSM\_18063

TCGAGCGAAGCCTTCGGGCTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGCCACTGG

AAACGGTGAGTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTT

GGTGGGGTAATGGCCTACCAAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGC

CCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGCAACCCTGATCCAGCCATGCCGCGTGAACGATG

AAGGCCTTAGGGTTGTAAAGTTCTTTCAGCKGGAAGATAAGACGGTACCAGCAGAAGAAGCCCCGGCTAACTCCGTGCCA

GCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGATTGGAAAGTTG

GGGGTGAAATCCCAGGGCTCAACCCTGGAACTGCCTCCAAAACTACCAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCC

GAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAG

GTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGATGAATGCCAGACGTCGGGCA

GCATGCTGTTCGGTGTCCACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATT

GACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTG

ATCGCGGTTTCTCGAGAGAGATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCG

TGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCAGTTCGGCTGGGCACTCTAAGAGAA

CTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTAC

AATGGCAGTGACAATGGGTTAATCCCCAAAAGCTGTCTCAGTTCGGATTGTCCTCTGCAACTCGAGGGCATGAAGTCGGA

ATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAG

TTGGGTTTACCCGAAGACGGTGCGCCAACCTTTCGAGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTRAAGTCG

TAACAAGGGTAACCAA

>Rhodovulum\_adriaticum\_DSM\_2781

CCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGCCTTCGGGCTTAGCGGCGGACGGGTG

AGTAACGCGTGGGAACGTGCCCTTCACTACGGAATAGTCCCGGGAAACTGGGTTTAATACCGTATACGCCCTTCGGGGGA

AAGAATTTCGGTGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCTATAGC

TGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTAG

ACTAGATGGGGGCAACCCTGATCTAGCCTGCCGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCGCCTGGGAA

GATAATGACGGTACCAGGTAAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGT

TCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGGCTGTTAAGTCAGAGGTGAAATCCCAGGGCTCAACCCTGGAACTGC

CTTTGATACTGGCAGCCTGGAGGTCTGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAAG

AACACCAGTGGCGAAGGCGGCTCACTGGCCAGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATA

CCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACACACCTAACGGAGTAAGC

ATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT

TTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCCTGTCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCG

GCTGGACAGGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAA

CCCACGTCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGGAAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGAC

GTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCCAAAAACTG

TCTCAGTTCGGATTGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGA

ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCTGTTC

GCAGGAGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTC

>Rhodovulum\_kholense\_DSM\_19783

GGCCTACACATGCAAGTCGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCT

GCGGAATAGGCTCGGGAAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGC

GTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACT

GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCA

TGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGCTGGGAAGATAATGACGGTACCAGCAGAAGAAGCC

CCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGT

AGGCGGATTGGAAAGTTGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTCCAAAACTCCCAGTCTTGAGGTCGAG

AGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGG

CTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGA

ATGCCAGTCGTCGGGTAGCATGCTATTCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGG

TTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTA

CCAACCCTTGACATCCTGATCGCGGTTACCCGAGAGGGTTTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGC

TGTCGTCAGCTCATGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGT

TGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTG

GGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCAAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTC

GAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGC

CCGTCACACCATGGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCA

GCGACTGGGG

>Rhodobacter\_capsulatus\_SB1003

AGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAGACCTTCGGGTCTAGCGG

CGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTACGGAATAGCCCCGGGAAACTGGGAGTAATACCGTATGTGCC

CTTCGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGA

CGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTG

GGGAATCTTAGACAATGGGGGAAACCCTGATCTAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTT

CAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGG

CTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCAGAAAGTCAGAGGTGAAATCCCAGGGCTCAACC

TTGGAACTGCCTTTGAAACTCCTGGTCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGA

TATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACA

GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAGGCATGCCTGTCGGTGACACACCTAA

CGGATTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGG

AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCGAGATCGCGGTTACCAGAGATGGTTTC

CTTCAGTTCGGCTGGATCTTAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCA

ACGAGCGCAACCCACACTTTCAGTTGCCATCATTCAGTTGGGCACTCTGGAAGAACTGCCGATGATAAGTCGGAGGAAGG

TGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGTGGTGACAATGGGCCAATCC

CAAAAAGCCATCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCAT

GACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGGTCTACCCTAAGATGGTGCGC

CAACCAGCAATGGAGGCAGCCAGCCACGGTAGGCTCAGTGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCT

GCGGCTGGATCACCT

>Rhodospirillum\_rubrum\_ATCC\_11170

AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGCATCCTTCGGGATGAGTGG

CGCACGGGTGAGTAACACGTGGGAACGTACCTTGGAGTGCGGAATAATCTTTGGAAACGAGGACTAATACCGCATACGCC

CTTAGGGGGAAAGATTTATCGCTCCAAGATCGGCCCGCGTCCGATTAGCTAGTTGGCGGGGTAATGGCCCACCAAGGCGA

CGATCGGTAGCTGGTCTGAGAGGATGGCCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTG

GGGAATATTGCGCAATGGGGGCAACCCTGACGCAGCCATGCCGCGTGAGTGAAGAAGGCCTTCGGGTTGTAAAGCTCTTT

CGGGTGTGAAGATGATGACGGTAACACCAGAAGAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGG

CAAGCGTTGTTCGGAATTACTGGGCGTAAAGAGCGCGTAGGCGGTCTGATTAGTCAGAGGTGAAATCCCAGAGCTCAACT

TTGGAACTGCCTTTGATACTGTTAGACTAGAATCCGTGAGAGGGTGGTGGAATTCCCAGTGTAGAGGTGAAATTCGTAGA

TATTGGGAGGAACACCAGTGGCGAAGGCGGCCACCTGGCGCGGTATTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACA

GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAGATGTCGGGGTACATGTACCTCGGTGTCGCAGCTAA

CGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGG

AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAGCCCTTGACATCCCGTGACACTTCCAGAGATGGAAGGTT

CCCTTCGGGGACACGGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGA

GCGCAACCCTCATCTTCAGTTGCCAGCAAGTAACGTTGGGCACTCTGAAGAGACTGCCGGTGACAAGCCGGAGGAAGGTG

GGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACACACGTGCTACAATGGCGCCTACAATGGGCAGCGACCT

CGCGAGGGGAAGCTAATCTCCAAAAGGCGTCTCAGTTCGGATTGCACTCTGCAACTCGGGTGCATGAAGTCGGAATCGCT

AGTAATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTT

CTACCCGAAGACGGTACGCTAACCGCAAGGAGGCAGCCGGCCACGGTAGGGTCAGCGACTGGGGTGAAGTCGTAACAAGG

TAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>Rhodobacter\_sphaeroides\_2.4.1

CAACTTGAGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGGAC

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCGGGAAACTGGGAGTAATACCGA

ATGTGCCCTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGGGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTAAA

GATCTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAAGTCAGAGGTGAAATCCCAGGG

CTCAACCCTGGAACTGCCTTTGAAACTCCCGATCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAGAT

GGTTCCTTCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTC

CGGCAACGAGCGCAACCCACGTCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGGAAACTGCCGGTGATAAGCCGGAG

GAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTT

AATCCCAAAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAAC

AGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGCGG

TGCGCCAACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>Rhodobacter\_sphaeroides\_KD131

AGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGGACTTAGCGG

CGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCGGGAAACTGGGAGTAATACCGAATGTGCC

CTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGA

CGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTG

GGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTAAAGATCTTT

CAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGG

CTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAAGTCAGAGGTGAAATCCCAGGGCTCAACC

CTGGAACTGCCTTTGAAACTCCCGATCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGA

TATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACA

GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACACACCTAA

CGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGG

AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAGATGGTTCCT

TCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAAC

GAGCGCAACCCACGTCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGGAAACTGCCGGTGATAAGCCGGAGGAAGGTG

TGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCA

AAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGA

CGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGCGGTGCGCCA

ACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCT

GCGGCTGGATCACCT

>Rhodobacter\_sphaeroides\_ATCC17025

AGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGGACTTAGCGG

CGGACGGGTGAGTAACGCGTGGGAACATGCCCAAAGGTACGGAATAGCCCCGGGAAACTGGGAGTAATACCGTATGTGCC

CTTCGGGGGAAAGATTTATCGCCTTTGGATTGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGA

CGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTG

GGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTAAAGATCTTT

CAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGG

CTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTGGAAAGTCAGGGGTGAAATCCCGGGGCTCAACC

CCGGAACTGCCTTTGAAACTCCCAGTCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGA

TATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACA

GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACACACCTAA

CGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGG

AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAGATGGTTCCT

TCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAAC

GAGCGCAACCCACGTCCTCAGTTGCCAGCATTCAGTTGGGCACTCTGGGGAAACTGCCGGTGATAAGCCGGAGGAAGGTG

TGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCA

AAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGA

CGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGCGGTGCGCCA

ACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCT

GCGGCTGGATCACCTCCTTT

>DSM2351

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCTTGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTTCTAA

>DSM1374

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>SNK001

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB14

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB26

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCGGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTCCGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCAAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB30

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAAC

>SE1

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>IM796

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB35

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCTTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB23

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB22

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB16

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTTGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTGCGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB20

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB28

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB18

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB17

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB15

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCTTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>AB33

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAACCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAA

GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA

GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTA

ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGAC

GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>2\_4\_1

CAACTTGAGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGGAC

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCGGGAAACTGGGAGTAATACCGA

ATGTGCCCTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACC

AAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC

AGCAGTGGGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTAAA

GATCTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAAGTCAGAGGTGAAATCCCAGGG

CTCAACCCTGGAACTGCCTTTGAAACTCCCGATCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAGAT

GGTTCCTTCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTC

CGGCAACGAGCGCAACCCACGTCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAGGGAAACTGCCGGTGATAAGCCGGAG

GAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTT

AATCCCAAAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTAAC

AGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGCGG

TGCGCCAACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGG

GGAACCTGCGGCTGGATCACCTCCTTT

>TIE1

CAACTTGAGAGTTTGATCCTGGCTCAGAGCGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAACGGGCGTAGCAATA

CGTCAGTGGCAGACGGGTGAGTAACGCGTGGGAACGTACCTTTTGGTTCGGAACAACACAGGGAAACTTGTGCTAATACC

GGATAAGCCCTTACGGGGAAAGATTTATCGCCGAAAGATCGGCCCGCGTCTGATTAGCTAGTTGGTGAGGTAATGGCTCA

CCAAGGCGACGATCAGTAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAG

GCAGCAGTGGGGAATATTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCCTAGGGTTGTA

AAGCTCTTTTGTGCGGGAAGATAATGACGGTACCGCAAGAATAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATA

CGAAGGGGGCTAGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTTTCTAAGTCAGAGGTGAAAGCCTGG

AGCTCAACTCCAGAACTGCCTTTGATACTGGAAGTCTTGAGTATGGCAGAGGTGAGTGGAACTGCGAGTGTAGAGGTGAA

ATTCGTAGATATTCGCAAGAACACCAGTGGCGAAGGCGGCTCACTGGGCCATTACTGACGCTGAGGCACGAAAGCGTGGG

GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGCCGTTAGTGGGTTTACTCACTAGTGG

CGCAGCTAACGCTTTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCAC

AAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGCAGAACCTTACCAGCCCTTGACATGTCCAGGACCGGTCGCAGA

GACGCGACCTTCTCTTCGGAGCCTGGAGCACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAG

TCCCGCAACGAGCGCAACCCCCGTCCTTAGTTGCTACCATTTAGTTGAGCACTCTAAGGAGACTGCCGGTGATAAGCCGC

GAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACACACGTGCTACAATGGCGGTGACAATGG

GAAGCTAAGGGGCGACCCTTCGCAAATCTCAAAAAGCCGTCTCAGTTCGGATTGGGCTCTGCAACTCGAGCCCATGAAGT

TGGAATCGCTAGTAATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATG

GGAGTTGGCTTTACCTGAAGACGGTGCGCTAACCAGCAATGGGGGCAGCCGGCCACGGTAGGGTCAGCGACTGGGGTGAA

GTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>geobactersulfurreducens

AAACTGGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGTGATCCATCCTT

CGGGGTGGTGAAAGTGGCGCACGGGTGAGTAACGCGTGGATAATCTGCCCGAGGATTTGGGATAACATCTCGAAAGGGGT

GCTAATACCGAATAAGCCCACGGGGTCTACGGATCTTGCGGGAAAAGGGGGGGACTTTCGGGCCTCCTGTCTTCGGATGA

GTCCGCGTACCATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGATCAGC

CACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCGCAATGGGGGAAACCCTGACG

CAGCAACGCCGCGTGGGTGATGAAGGCCTTCGGGTCGTAAAGCCCTGTCGGGAGGGAAGAAATGATTGAGAGCTAATACC

TCTTGGTCTTGACGGTACCTCCGAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCG

TTGTTCGGAATTATTGGGCGTAAAGCGCGTGTAGGCGGTCTTTTAAGTCTGATGTGAAAGCCCCGGGCTCAACCTGGGAA

GTGCATTGGAAACTGGGAGACTTGAGTACGGGAGAGGAGAGTGGAATTCCTAGTGTAGGAGTGAAATCCGTAGATATTAG

GAGGAACACCGGTGGCGAAGGCGGCTCTCTGGACCGATACTGACGCTGAGACGCGAAAGCGTGGGTAGCAAACAGGATTA

GATACCCTGGTAGTCCACGCCGTAAACGATGAGTACTAGGTGTTGCGGGTATTGACCCCTGCAGTGCCGTAGCTAACGCA

TTAAGTACTCCGCCTGGGAAGTACGGTCGCAAGACTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCA

TGTGGTTTAATTCGACGCAACGCGCAGAACCTTACCTGGGCTTGACATCCACGGAACCCTCCCGAAACGGAGGGGTGCCC

TTCGGGGAGCCGTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGC

GCAACCCTTATCATCAGTTGCCATCATTAAGTTGGGCACTCTGATGAAACTGCCGGTGTCAAACCGGAGGAAGGTGGGGA

TGACGTCAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGCCGGTACAAAGAGTCGCAATGCCGCG

AGGTGGAGCCAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGT

AATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGGGAGTCGACTGGT

CCCGAAGTGCGTGAGCTAACCAGCAATGGAGGCAGCGTCCTAAGGAATGGTCGGTGACTGGGGTGAAGTCGTAACAAGGT

AGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>Pseudomonas\_aeruginosa

GAACTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCAGCAGGGGCCTTCAACACATGCAAGTCGAG

CTTATGAAGGGAGCTTGCCTTGGATTCAGCGGCGGACGGGTGAGTAATGCCTAGGAATCTGCCTGGTAGT

GGGGGATAACGTCCGGAAACGGCCGCTAATACCGCATACGTCCTGAGGGAGAAAGTCGGGGATCTTCGGA

CCTCACGCTATCAGATGAGCCTAGGTCGGATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGACGAT

CCGTAACTGGTCTGAGAGGATGATCAGTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCA

GCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCG

GATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCAGTAAGTTAATACCTTGCTGTTTGACGTTACCAACAG

AATAAGCACCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTAC

TGGGCGTAAAGCGCGCGTAAGTGGTTCAGCAAGCTTGATGTGAAATCCCCGGGCTCAACCTGGGAACTGC

ATCCAAAAGCTACTGAGCTAGAGTACGGTAGAGGTGGTAGAATTTCCTGTGTAGCGGTGAAATGCGTAGA

TATAGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGACTGTACTGACACTGAGGTGCGAAAGCGTGGG

GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTCGACTAGCCGTTGGGATCCTTGA

GATCTTAGTGGCGCACGTAACGCGATAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAA

TGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACC

TGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAACAGAGACACAGGTGCTGCA

TGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGT

TACCAGCACCTCGGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTC

AAGTCATCATGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTCGGTACAAAGGGTTGCCAAGCC

GCGAGTGGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGT

CGGAATCGCTAGTAATCGTGAATCAGAATGTCACGGTGAATACGTCCCCGGGCCTTGTACACACCGCCCG

TCACACCATGGGAGTGGGTTGCTCCAGAAGTAGCTAGTCTAACCGCAAGGGGGACGGTTACCACGGAGTG

ATTCATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTA

>Nitrosococcus\_oceani

AGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCATTGACCACAC

CGTAGTAAGCGTCCTCCCGAAGGTTAAACTACCCACTTCTGGTGCAACCAACTCCCATGGTGTGACGGGC

GGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATTGCTGATCCGCGATTACTAGCGATTCCGACTT

CATGCAGTCGAGTTGCAGACTGCAATCCGGACTACGACCGGTTTTCTGAGATTAGCTCCGCATCGCTGCT

TGGCCACCCTCTGTACCGGCCATTGTAGCACGTGTGTAGCCCTGGCCATAAGGGCCATGATGACTTGACG

TCATCCCCACCTTCCTCCGTTTTGTCAACGGCAGTCCCTCTAGAGTGCCCGGCCGAACCGCTGGCAACTA

GAGATAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCA

GCACCTGTCTCCCGGCTCCCGAAGGCACCTTCGCCTCTCAGCAAAGTTCCGGGGATGTCAAGGCCAGGTA

AGGTTCTTCGCGTTGCATCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCATTTG

AGTTTCAACCTTGCGGCCGTACTCCCCAGGCGGAGAACTTAACGCGTTAGCTGCAACACTGAAAGGCTAT

CCCTCCCAACGTCTAGTTCTCATCGTTTAGGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCA

CGCTTTCGCACCTCAGCGTCAGTAGTGGTCCAGGAGGCCGCCTTCGCCACTGGTGTTCCTTCCGATATCT

ACGCATTTCACCGCTACACCGGAAATTCCGCCTCCCTCTACCATACTCTAGCCAGGCAGTATCAAGTGCC

ATTCCCAGGTTGAGCCCGGGGCTTTCACACCCAACTTACCAAACCGCCTACGCGCCCTTTACGCCCAGTA

ATTCCGATTAACGCTCGCACCCTCCGTATTACCGCGGCTGCTGGCACGGAGTTAGCCGGTGCTTCTTCTG

TAGGTAACGTCAACACGATGTGCTATTCACACATCGCTTTTCTTCCCCACTGAAAGGGCTTTACAACCCG

AAAGCCTTCTTCACCCACGCGGCATTGCTGGATCAGGGTTGCCCCCATTGTCCAATATTCCCCACTGCTG

CCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCTGATCGTCCTCTCAGACCAGCTACCGAT

CATCGCCTTGGTAGGCCCTTACCCCACCAACTAGCTAATCGGACGCAGGCTCATCCCCCAGCGCGAGGCC

TTTCGGTCCCCCGCTTTCCTCCTTAGAGATTATGCGGTATTAGCCCGAGTTTCCCCGAGTTATCCCCCTC

TAGAGGCCAGATTCCCACGCGTTACTCACCCGTCCGCCACTCGCCAGCCACCTAAGCAAGCTTAGGTGCT

GCTGCCGTTCGACTTGCATGTGTTAAGCATGCCGCCAGCGTTCAATCTGAGCCATGATCAAACTCTTCAG

TTTAA

>Rhodovulum\_visakhapatnamense\_JA181

GGCTCAGAACGAACGCTGGCGGCASGCCTAACCATGCAASTCGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGASTA

ACGCGTGGGAACGTGCCCTTCTCTGCGGGATAGGCTCGGGAAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGA

TTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGG

TTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACA

ATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGTCGTGAAGATA

ATGACGGTAGCGACAGAAGAAGCCCCGGCTAAYTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGG

AATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTCC

GATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACA

CCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCT

GGTAGTCCACGCCGTAAACGGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACACACCTAACGGAYTAAGCRYT

CCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAACTGACGGGGGCCCGCACAAGCCGGTGGAGCATGTGGTTT

AATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGGC

TTGGATCAGTTGACAGGTGCTTGCATGGCTTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGC

AACCCACACTCTTAGTTGCCAGCATTCAGTTTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGAT

GACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCAAAAAA

CTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGG

TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCT

TACGGGGGGCAGGGGCCACGGTAKGTCAGCGRTGGGTGAG