



Supplemental Figure S1. 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 S1. 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

```
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGCGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCTTGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGCAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTGAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTTCTAA
```

>dsm1374

```
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGCGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
```

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGTAAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTT
>SNK001
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTTCGGAGGAACACCACTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCTGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTAACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTT
>AB14
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTTCGGAGGAACACCACTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCTGGTGACA

CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB26
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCTTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGCGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTCTGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCGGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTCCGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AAGAGAACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB30
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCACTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTGCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCA
AAGGAATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAAC

>se1

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTGCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>IM796

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCACTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB35

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCTTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB23

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB22
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCAGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>aB16

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCTTCTCTGCGGAATAG
GCTTGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTGCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTGCGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGG
GTGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB20

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGG
CTCAACCCCGGAACCTGCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTT

>AB28

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC

AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCCGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCACTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTT

>AB19

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAACAACCTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTA
ACACATGCAAGTCGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACG
CGTGGGAACGTGCCCTGTGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGA
GTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCGGGAAACTGGGTTTA
ATACCGCATACGCCCTTCGGGGGAAAGTCTCTGCGGAATAGGCTCGGGAAACTGG
GTTTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCC
GCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTACCATTATCGGAGAAGGATC
GGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCT
ATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGAAGCCTAC
GATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCC
AGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGAT
CCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAAC
CCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTT
CAGTCGTGAAGATAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGC
TCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCC

GTGCCAGCAGCCGCGGTAATACGAATGACGGTAGCGACAGAAGAAGCCCCGGCT
AACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCCGAATTA
CTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGGAGGGGGCTAGCGTTGTTCCG
GAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCC
CGGGGCTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGGGGGTGA
AATCCCGGGGCTCAACCCCGGAACTGCCTCCGATACTGGTAGTCTAGAGTTCGAG
AGAGGTGAGTGGAATTCAGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAAC
TCGAGAGAGGTGAGTGGAATTCAGAGTGTAGAGGTGAAATTCGTAGATATTCGGA
GGAACACCACTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGA
AAGCGTGGGGAACCACTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAG
GTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAA
ACGATGAATGCCAGTCGTCGGCGCAAACAGGATTAGATACCCTGGTAGTCCACGC
CGTAAACGATGAATGCCAGTCGTCGGCAAGCTTGCTTGTCGGTGACACACCTAAC
GGATTAAGCATTCCGCCTGGGGAGTACGGCCGAAGCATGCTTGTCGGTGACACAC
CTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAG
GAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTACAAGGTTAAACT
CAAAGGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAAATCGAAG
CAACGCGCAGAACCCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGAATTC
GAAGCAACGCGCAGAACCCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAG
AGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAG
CTCGTGGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTC
GTCAGCTCGTGTCGTGAGATGTTCCGGTTAAGTCCGGCAACGAGCGCAACCCACAC
TCTTAGTTGCCAGCATTCTGTGAGATGTTCCGGTTAAGTCCGGCAACGAGCGCAACC
CACACTCTTAGTTGCCAGCATTCTGTGAGATGTTCCGGTTAAGTCCGGCAACGAGCGCAACC
GTCGGAGGAAGGTGTGGATGACGTCTCAGTTGGGCACTCTAAGAGAACTGCCGATGATAA
GATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTG
GGCTACACACGTGCTACAATGGCAGTGACAATGGGAAGTCCTCATGGCCCTTACG
GGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCTAAAACT
GTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGTTAATCCCAAAA
AACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCT
AGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACGAA
TCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTAC
ACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACC
CTTACGGGACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCG
CCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGT
CGTAACAAGGTAGCCGTAGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTTCT
AAGGAACCTGCGGCTGGATCACCTCCTTTCTAA

>AB18

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCA
AAGGAATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGCAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTCAAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB17

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB15

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA

GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCTTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB33
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAACCTTCCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCTTCTCTGCGGAATAG
GCTCGGGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGA
GCGATGAAGGCCTTAGGGTTGTAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACT
ATTAAGTCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGT
GGAATTCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGCAAGCATGCTTGTCCGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCT
AGGAGAACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACAC
GTGCTACAATGGCAGTGACAATGGG

TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>2_4_1
CAACTTGAGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACA
TGCAAGTCGAGCGAAGTCTTCGGAC
TTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCTTTGCTTCGGAATAG
CCCCGGGAACTGGGAGTAATACCGA
ATGTGCCCTTTGGGGGAAAGATTTATCGGCCAAAGGATCGGCCCGCGTTGGATTAG
GTAGTTGGTGGGGTAATGGCCTACC
AAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGA
CACGGCCCAGACTCCTACGGGAGGC
AGCAGTGGGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGA
TCGATGAAGGCCTTAGGGTTGTAA
GATCTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACT
CCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATC
GGAAAGTCAGAGGTGAAATCCCAGGG
CTCAACCCTGGAAGTGCCTTTGAACTCCCGATCTTGAGGTCGAGAGAGGTGAGT
GGAATTCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTG
ACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCG
TCGGGCAGCATGCTGTTCCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCA
AAGGAATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTG
ACATGGCGATCGCGTTCCAGAGAT
GGTTCCTTCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTC
GTGTCGTGAGATGTTCCGGTTAAGTC
CGGCAACGAGCGCAACCCACGTCCTTAGTTGCCAGCATTGAGTTGGGCACTCTAG
GGAACTGCCGGTGATAAGCCGGAG
GAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGT
GCTACAATGGCAGTGACAATGGGTT
AATCCCAAAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGT
CGGAATCGCTAGTAATCGCGTAAC
AGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCAT
GGGAATTGGTTCTACCCGAAGGCGG
TGCGCCAACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGACTGGGG
TGAAGTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>TIE1

CAACTTGAGAGTTTGTATCCTGGCTCAGAGCGAACGCTGGCGGCAGGCTTAACACA
TGCAAGTCGAACGGGCGTAGCAATA
CGTCAGTGGCAGACGGGTGAGTAACGCGTGGGAACGTACCTTTTGGTTCGGAACA
ACACAGGGAACTTGTGCTAATACC
GGATAAGCCCTTACGGGGAAAGATTTATCGCCGAAAGATCGGCCCGCGTCTGATT
AGCTAGTTGGTGAGGTAATGGCTCA
CCAAGGCGACGATCAGTAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGA
GACACGGCCCAAACCTCCTACGGGAG
GCAGCAGTGGGGAATATTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCG
TGAGTGATGAAGGCCCTAGGGTTGTA
AAGCTCTTTTGTGCGGGAAGATAATGACGGTACCGCAAGAATAAGCCCCGGCTAA
CTTCGTGCCAGCAGCCGCGGTAATA
CGAAGGGGGCTAGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGG
TTTCTAAGTCAGAGGTGAAAGCCTGG
AGCTCAACTCCAGAACTGCCTTTGATACTGGAAGTCTTGAGTATGGCAGAGGTGAG
TGGAAGTGCAGTGATAGAGGTGAA
ATTCGTAGATATTCGCAAGAACACCAAGTGGCGAAGGCGGCTCACTGGGCCATTAC
TGACGCTGAGGCACGAAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGC
CGTTAGTGGGTTTACTCACTAGTGG
CGCAGCTAACGCTTTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAACT
CAAAGGAATTGACGGGGGCCCCGCAC
AAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGCAGAACCTTACCAGCCCT
TGACATGTCCAGGACCGGTGCGCAGA
GACGCGACCTTCTCTTCGGAGCCTGGAGCACAGGTGCTGCATGGCTGTCGTCAGC
TCGTGTCGTGAGATGTTGGGTAAAG
TCCCGCAACGAGCGCAACCCCCGTCCTTAGTTGCTACCATTTAGTTGAGCACTCTA
AGGAGACTGCCGGTGATAAGCCGC
GAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACAC
ACGTGCTACAATGGCGGTGACAATGG
GAAGCTAAGGGGCGACCCTTCGCAAATCTCAAAAAGCCGTCTCAGTTCGGATTGG
GCTCTGCAACTCGAGCCCATGAAGT
TGGAATCGCTAGTAATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCT
TGACACACCGCCCGTCACACCATG
GGAGTTGGCTTTACCTGAAGACGGTGCGCTAACCAGCAATGGGGGCAGCCGGCC
ACGGTAGGGTCAGCGACTGGGGTGAA
GTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>geobactersulfurreducens

AAACTGGAGAGTTTGTATCCTGGCTCAGAACGAACGCTGGCGGCGTGCCTAACACA
TGCAAGTCGAACGTGATCCATCCTT
CGGGGTGGTGAAAGTGGCGCACGGGTGAGTAACGCGTGGATAATCTGCCCCGAGG
ATTTGGGATAACATCTCGAAAGGGGT
GCTAATACCGAATAAGCCCACGGGGTCTACGGATCTTGCGGGAAAAGGGGGGGA
CTTTCCGGGCCTCCTGTCTTCGGATGA
GTCCGCGTACCATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGATGGT
TAGCTGGTCTGAGAGGATGATCAGC

CACACTGGAAGTGAAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
TTTGCGCAATGGGGGAAACCCTGACG
CAGCAACGCCGCGTGGGTGATGAAGGCCTTCGGGTCGTAAAGCCCTGTCGGGAG
GGAAGAAATGATTGAGAGCTAATACC
TCTTGGTCTTGACGGTACCTCCGAAGGAAGCACCGGCTAACTCCGTGCCAGCAGC
CGCGGTAATACGGAGGGTGCAAGCG
TTGTTCCGAATTATTGGGCGTAAAGCGCGTGTAGGCGGTCTTTTAAGTCTGATGTG
AAAGCCCCGGGCTCAACCTGGGAA
GTGCATTGGAAACTGGGAGACTTGAGTACGGGAGAGGAGAGTGGAAATTCCTAGTG
TAGGAGTGAAATCCGTAGATATTAG
GAGGAACACCGGTGGCGAAGGCGGCTCTCTGGACCGATACTGACGCTGAGACGC
GAAAGCGTGGGTAGCAAACAGGATTA
GATACCCTGGTAGTCCACGCCGTAAACGATGAGTACTAGGTGTTGCGGGTATTGA
CCCCTGCAGTGCCGTAGCTAACGCA
TTAAGTACTCCGCCTGGGAAGTACGGTCGCAAGACTAAAACCTCAAAGGAATTGACG
GGGGCCCGCACAAGCGGTGGAGCA
TGTGGTTTAATTGACGCAACGCGCAGAACCTTACCTGGGCTTGACATCCACGGA
ACCCTCCCGAAACGGAGGGGTGCCC
TTCGGGGAGCCGTGAGACAGGTGCTGCATGGCTGTCTGTCAGCTCGTGTCTGAG
ATGTTGGGTTAAGTCCCGCAACGAGC
GCAACCCTTATCATCAGTTGCCATCATTAAAGTTGGGCACTCTGATGAAACTGCCGG
TGTCAAACCGGAGGAAGGTGGGGA
TGACGTCAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGC
CGGTACAAAGAGTCGCAATGCCGCG
AGGTGGAGCCAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGAGTCTGCAACTCG
ACTCCATGAAGTCGGAATCGCTAGT
AATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCC
CGTCACACCACGGGAGTCGACTGGT
CCCGAAGTGCGTGAGCTAACCAGCAATGGAGGCAGCGTCCTAAGGAATGGTCGG
TGACTGGGGTGAAGTCGTAACAAGGT
AGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>*Pseudomonas_aeruginosa*

GAAGTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCAGCAGGGGCCTTCAA
CACATGCAAGTCGAG
CTTATGAAGGGAGCTTGCCTTGGATTGAGCGGCGGACGGGTGAGTAATGCCTAGG
AATCTGCCTGGTAGT
GGGGGATAACGTCCGGAAACGGCCGCTAATACCGCATACGTCCTGAGGGAGAAA
GTCGGGGATCTTCGGA
CCTCACGCTATCAGATGAGCCTAGGTGCGATTAGCTAGTTGGTGGGGTAAAGGCC
TACCAAGGCGACGAT
CCGTAAGTGGTCTGAGAGGATGATCAGTCACACTGGAAGTGAAGACACGGTCCAGA
CTCCTACGGGAGGCA
GCAGTGGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGT
GTGAAGAAGGTCTTCG
GATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCAGTAAGTTAATACCTTGCTGTTT
GACGTTACCAACAG

AATAAGCACCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGC
GTTAATCGGAATTAC
TGGGCGTAAAGCGCGCGTAAGTGGTTCAGCAAGCTTGATGTGAAATCCCCGGGCT
CAACCTGGGAACTGC
ATCCAAAAGCTACTGAGCTAGAGTACGGTAGAGGTGGTAGAATTTCTGTGTAGCG
GTGAAATGCGTAGA
TATAGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGACTGTACTGACACTGAG
GTGCGAAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTGCGACTAGC
CGTTGGGATCCTTGA
GATCTTAGTGGCGCACGTAACGCGATAAGTCGACCGCCTGGGGAGTACGGCCGC
AAGGTTAAAACTCAAA
TGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAAC
GCGAAGAACCTTACC
TGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAACAGA
GACACAGGTGCTGCA
TGGCTGTCGTCAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCA
ACCCTTGTCCTTAGT
TACCAGCACCTCGGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGA
AGGTGGGGATGACGTC
AAGTCATCATGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTCGGTACA
AAGGGTTGCCAAGCC
GCGAGTGGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAAC
TCGACTGCGTGAAGT
CGGAATCGCTAGTAATCGTGAATCAGAATGTCACGGTGAATACGTCCCCGGGCCT
TGACACACCGCCCCG
TCACACCATGGGAGTGGGTTGCTCCAGAAGTAGCTAGTCTAACCGCAAGGGGGAC
GGTTACCACGGAGTG
ATTCATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGG
ATCACCTCCTTA
>Nitrosococcus_oceani
AGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCC
AGTCATTGACCACAC
CGTAGTAAGCGTCCTCCCGAAGGTTAACTACCCACTTCTGGTGCAACCAACTCCC
ATGGTGTGACGGGC
GGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATTGCTGATCCGCGATTAC
TAGCGATTCCGACTT
CATGCAGTCGAGTTGCAGACTGCAATCCGGACTACGACCGGTTTTCTGAGATTAG
CTCCGCATCGCTGCT
TGGCCACCCTCTGTACCGGCCATTGTAGCACGTGTGTAGCCCTGGCCATAAGGGC
CATGATGACTTGACG
TCATCCCCACCTTCCTCCGTTTTGTCAACGGCAGTCCCTCTAGAGTGCCCGGCCG
AACCGCTGGCAACTA
GAGATAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCT
GACGACAGCCATGCA

GCACCTGTCTCCCGGCTCCCGAAGGCACCTTCGCCTCTCAGCAAAGTTCCGGGGA
TGTCAAGGCCAGGTA
AGGTTCTTCGCGTTGCATCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCC
CCGTCAATTCATTTG
AGTTTCAACCTTGCGGCCGTA CTCCCCAGGCGGAGAACTTAACGCGTTAGCTGCA
ACACTGAAAGGCTAT
CCCTCCCAACGTCTAGTTCTCATCGTTTAGGGCGTGGACTACCAGGGTATCTAATC
CTGTTTGCTCCCCA
CGCTTTCGCACCTCAGCGTCAGTAGTGGTCCAGGAGGCCGCTTCGCCACTGGTG
TTCCTTCGGATATCT
ACGCATTTACCGCTACACCGGAAATTCCGCCTCCCTCTACCATACTCTAGCCAGG
CAGTATCAAGTGCC
ATTCCCAGGTTGAGCCCCGGGGCTTTACACCCAACTTACCAAACCGCCTACGCGC
CCTTTACGCCAGTA
ATTCCGATTAACGCTCGCACCTCCGTATTACCGCGGCTGCTGGCACGGAGTTAG
CCGGTGCTTCTTCTG
TAGGTAACGTCAACACGATGTGCTATTCACACATCGCTTTTCTTCCCCACTGAAAG
GGCTTTACAACCCG
AAAGCCTTCTTCACCCACGCGGCATTGCTGGATCAGGGTTGCCCCCATTTGTCCAAT
ATTCCCCACTGCTG
CCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCTGATCGTCCTCTC
AGACCAGCTACCGAT
CATCGCCTTGGTAGGCCCTTACCCACCAACTAGCTAATCGGACGCAGGCTCATC
CCCCAGCGCGAGGCC
TTTCGGTCCCCCGCTTTCCTCCTTAGAGATTATGCGGTATTAGCCCGAGTTTCCCC
GAGTTATCCCCCTC
TAGAGGCCAGATTCCACGCGTTACTCACCCGTCCGCCACTCGCCAGCCACCTAA
GCAAGCTTAGGTGCT
GCTGCCGTTGCACTTGCATGTGTTAAGCATGCCGCCAGCGTTCAATCTGAGCCAT
GATCAAACCTCTTCAG
TTTAA

Input Fasta file for the PufM Woods Hole Isolates Phylogenetic Tree

>AB35

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWIWAQEHNFVTEYTQPYGVDAVVGQGG

>DSM2351

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLF
MVL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWIWAQEHNFVTEYTQPYGVDAVVGQGG

>DSM1374

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWIWAQEHNFVTEYTQPYGVDAVVGQGG

>IM796

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWIWAQEHNFVTEYTQPYGVDAVVGQGG

>SE1

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWWYIWAQEHNFVTEYTQPYGVDAAYVGQGG
>SNK001
MAEQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWWYIWAQEHNFVTEYTQPYGVDAAYVGQGG
>AB14
MAEQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWWYIWAQEHNFVTEYTQPYGVDAAYVGQGG
>AB15
MAEQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWWYIWAQEHNFVTEYTQPYGVDAAYVGQGG
>AB16
MAEQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG
NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAAYVGQGG

>AB17

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA

M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF

AVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAAYVGQGG

>AB18

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA

M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF

AVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAAYVGQGG

>AB20

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA

M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF

AVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAAYVGQGG

>AB21

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA

M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF

AVLT

PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAAYVGQGG

>AB22

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT

PLTGGIGILLTGTVVDNWWYIWAQEHNFTVEYTQPYGVDAAYVGQGG

>AB23

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT

PLTGGIGILLTGTVVDNWWYIWAQEHNFTVEYTQPYGVDAAYVGQGG

>AB26

MAEYQNIFTQVQVQGPGDWGMDNENNMMEERAAGTGHFSILGWL

GNAQLGPIYLGYTGVISVIAGAFALIIIGLNMLAQVDWSLVQFLRQGFWLAEPPSPE

YGLRIPPLKEGGWYMIASFFLLIAVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLF

MV

LGFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA

MHGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWW

FAVL

TPLTGGIGILLTGTVVDNWWYIWAQEHNFTVEYTQPYGIDAYVGQGG

>AB28

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT

PLTGGIGILLTGTVVDNWWYIWAQEHNFTVEYTQPYGVDAAYVGQGG

>AB30

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLAEPPSPEY

GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM

VL

GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYYNPFHALSIAFLYGSALLFA
M

HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT

PLTGGIGILLTGTVVDNWWYIWAQEHNFTVEYTQPYGVDAAYVGQGG

>AB33

MAEYQNIFTQVQVQGSPDWGMDNENNMMEERAGMGHFSILGWLG

NAQLGPIYLGYTGVVSLIAGCFAFLIIGLNMLAQVDWSLVQLLRQGFWLALEPPSPEY
GLRIPPLKEGGWYMWASFFLLISVWAWWARTYMLAVEHKMGKHIAWAFLSAIWLFM
VL
GFFRPILMGSWSEMVPYGIFPHLDWTTAFSIRYGNLYNPFHALSIAFLYGSALLFA
M
HGGTILAVTRFGGDRELEQIYDRGTASERAALFWRWTMGFNATMEGIHRWAWWF
AVLT
PLTGGIGILLTGTVVDNWYIWAQEHNFVTEYTQPYGVDAVVGQGG
>TIE_1
MAQYQNIFTQVQVEGPAYAGVPLRPGSSPRETQTTFNYWLGKIGDAQVGPVYLGF
TGVCSLLCGFVAIEI
IGLNMLASVDWSPIEFLRQFCWLALEPPKPEYGLTIPPLKEGGWWLMAGFFLTVSIA
LWWVRTYRRSRAL
GMGTHVSWAFASAILLYLALGFIQPLLMSWSEAPPFGVFPPLDWTNNFSIKYGNL
YYNPFHCLSIAFLY
GSALLFAMHGATILAVSRYGGEREIEQMLDRGTALERAALFWRWTMGFNATAESIH
RWAWWFVAVLCPLTG
AIGIILTGPVVDNWFDWGVKHGLAPPR
>Chloroflexus_aggregans
MATMDMMPGDLELGRDRGRIGKPIEIPLENFGFDSQLGPFYLGFWNTVAYISGGIF
TFIWL FVMLSQVN
FNPVAFIKYFVVLQIDPPSSRYGLSFPLNEGGWWLIATFFLTL SIFAWYMHIYTRAK
ALGIKPYLAYGF
TGAIALYLVIYIIRPVWMGDWSEAPAHGIKALLDWTNNVSVRYGNFYYPFHMLSIF
LLGSTLLLAMHA
GTIWALEKYAAHEEWNEIQAPGTGTERAQLFWRWCIGFNANAYSIDLWAFWFAWL
CGITGALGVFFSMPD
FVNNWFQWGV EAGIAYPQGPTPLP

Fasta File Input for the Woods Whole EeuP Phylogenetic Tree

>AB26

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPD
GTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRID
GGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDG
SLPDPGWGPPVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQ
WPGFAHNGNDLLGLARNLGEVYGVFGEFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWD
IGAPRWPWELDAGLVEQGGAI FARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDER
ECEILTRAKTGVLEGAKVPLSSNTLGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAA
LQKDDKFMAEFDDL RGA FPRSGDGSAL ETA AAPCKYE ARVLDGIWAAAPYLHNGSVPTLRELL
TPPDERVADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDK
RALLEYL KSI

>AB33

MFTRSVIFSLLFSGAAAPVTAQTVFVDQGDWSWSPSLRAAFYTQDQGSRIMPLTWMEALAQPD
GTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRID
GGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGETGDRAGLRAELETWSLRFHTLIDG
SLPDPGWGPPVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQ
WPGFAHNGNDLLGLARNLGEVYGVFGEFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWD
IGAPRWPWGLDAGLVEQGGAI FSRPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDER
ECEILTRAKTGVLEGAKVPLSSNALGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAA
LQQDDKFMAEFDDL RGA FPRSGDGSAL ETA AAPCAYE ARVLDGIWAAAPYLHNGSVPTLRELL
TPPDERVADFTP GPAYNLEAVGMATEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDK
RALLEYL KSI

>AB22

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPD
GTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRID
GGPAIVDFQSFLADLDAAMQAVLDDDATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDG
SLPDPGWGPPVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQ
WPGFAHNGNDLLGLARNLGEVYGVFGEFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWD
IGAPRWPWELDAGLVEQGGAI FARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDER
ECEILTRAKTGVLEGAKVPLSSNTLGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAA
LQKDDKFMAEFDDL RGA FPRSGDGSAL ETA AAPCKYE ARVLDGIWAAAPYLHNGSVPTLRELL
TPPDERVADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDK
RALLEYL KSI

>AB35

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQD
QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTPNEGMGGTA
IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDEATFDAFADK
VAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPPVRLDAVSMIFNRLTGLDLGRP
EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG
EFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAI F
ARPTDQGGCVECHGKRQGAFRSIFHSTWATPILPVGTDERECEILTRAKTGVLEGAK
VPLSSNTLGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAALQKDDKFMAEF
DDL RGA FPRSGDGSAL ETA AEPCKYE ARVLDGIWAAAPYLHNGSVPTLRELLTPPDER
VADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA
LLEYL KSI

>AB20

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQD
QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTA
IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDEAVFDAFADK

VAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGDLGRP
EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG
EFHPIKKPGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAI
ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTARTGVLEGAK
VPLSSNTLGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAALQKDDKFMAEF
DDLRGAFPRSGDGSAALETAAAPCAYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER
VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA
LLEYLKSI

>IM796

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQD
QGSRIPLAWMEALAQPDPGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTA
IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDDATFADFADK
VAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGDLGRP
EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG
EFHPIKKQGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAI
ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAK
VPLSSNALGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAALQKDDKFMAEF
DDLRGAFPRSGDGSAALETAAAPCAYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER
VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA
LLEYLKSI

>AB30

MNEKIIPILFWEESMFTRSVIFSLLFSGAAAPVTAQTVFVDQGDWSWSPSLRAAFYTQDQGSRI
PLTWME
ALAQPDPGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVR
GTAYRIDG
GPAIVDFQSFLADLDAAMQAVLDDDATFADFADKVAGETGDRAGLRAELETWSLRFHTLIDRSL
PDPGWG
PVRLDAVSMIFNRLTGDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNG
NDLLGLAR
NLGEVYGVFGEFHPIKKQGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVE
QGGAI
PTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNTL
GAEAPA
FDLLGASVIGAIQHAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLGRGAFPRSGDGSAALETAA
EPCAYE
ARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTPGPAYDLEAVGMAVEQTAFDHVIETTGC
DDLSS
GNSRCGHDYGTALSDDDKRALLEYLKSI

>AB23

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQD
QGSRIPLAWMEALAQPDPGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTA
IGMTCAACHTRQIEVRGTAYRIDGGPAIVDFQNFADLDAAMQAVLDDDATFADFADK
VAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDAVSMIFNRLTGDLGRP
EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG
EFHPIKKQGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAI
ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAK
VPLSSNTLGAEAPAFDLLGASVIGAIQNAVGFRTGGLEATGAAAALQKDGKFMAEF
DDLGRGAFPRSGDGSAALETAAEPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER
VADFTPGPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDDDKRA
LLEYLKSI

>AB16

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQD
QGSRIMPLAWMEALAQPDGTPFLADGLARYGYLPMPDRDDGAALPVGFTTNEGMGGTA
IGMTCAACHTRQIEVSGTAYRIDGGPAIVDFQSFLADLDAAMQAVLDDEAVFDAFADK
VAGETGDRAGLRAELETWSLRFHTLIDRSLPDPGWGPVRLDAVSMIFNRLTGLDLGRP
EDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVYGVFG
EFHPIKKQGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAI
ARPTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAK
VPLSSNALGAEAPAFDLLGASVIGAIQHAVGFRVTGGLEATGAAAALQKDGKFMAEF
DDLRGAFPRSGDGSALETAAPCKYEARVLDGIWAAAPYLHNGSVPTLRELLTPPDER
VADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGHDYGTALSDGDKRA
LLEYLKSI

>AB14

MNEKITPILWEESMFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQDQGSRIM
PLAWME
ALAQPDGTPFLADGLARYGYLPMPDRDDGATLPVGFTTNEGMGGTAIGMTCAACHTRQIEVR
GTAYRIDG
GPAIVDFQSFLADLDAAMQAVLDDATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDGS
LPDPGWG
PVRLDAVSMIFNRLTGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNG
NDLLGLAR
NLGEVYGVFGEFHPIKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDARLVE
QGGAI
PTDQGGCVECHGKRQGAFRSIFHATWATPILPVGTDERECEILTRTARTGVLEGAKVPLSSNAL
GAEAPA
FDLLGASVIGAIQHAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLGRGAFPRSGDGSALETA
EPCAYE
ARVLDGIWAAAPYLHNGSVPTLRELLTPPDERVADFTP GPAYDLEAVGMAVEQTAFDHVIETT
GCDDLSS
GNSRCGHDYGTALSDDDKRALLEYLKSI

>AB15

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPD
GTPFLADG
LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDF
QSFLADL
DAAMQAVLDDEATFDAFADKVAGGTGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDA
VSMIFNRL
TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFAHNGNDLLGLARNLGEVY
GVFGEFHP
IKKPGLLFNRDYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAI
ARPTDQGG
CVECHGKR
QGAFRSIFHSTWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNTLGAEAPAFDLLGASVI
GAIQ
HAVGFRVTGGLEATGAAAALQKDDKFMAEFDDLGRGAFPRSGDGSALETAAPCKYEARVLDGI
WAAAPYL
HNGSVPTLRELLTPPDERVADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGH
DYGTALS
DDDKRALLEYLKSI

>AB17

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDWSWSPSLRAAFYTQDQGSRIMPLAWMEALAQPD
GTPFLADG

LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDF
QSFLADL
DAAMQAVLDDDDATFADFADKVAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDA
VSMIFNRL
TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFHNGNDLLGLARNLGEVY
GVFGEFHP
IKKPGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWELDAGLVEQGGAI FARPTDQGG
CVECHGKR
QGAFRSIFHATWATPILHVGTDERECEILTRTARTGVLEGAKVPLSSNALGAEAPAFDLLGASVI
GAIQ
NAVGFVRTGGLEATGAAAALQKDDKFMAEFDDL RGAFPRSGDGSAL ETAAEPCKYEARVLDGI
WAAAPYL
HNGSVPTLRELLTPPDERVADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGH
DYG TALS
DDDKRALLEYLKS R

>AB18

MFTRSVIFSLLFSGAAAPVTAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLTWMEALAQPD
GTPFLADG
LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDF
QSFLADL
DAAMQAVLDDDDATFADFADKVAGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDA
VSMIFNRL
TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFHNGNDLLGLARNLGEVY
GVFGEFHP
IKKQGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAI FSRPTDQGG
CVECHGKR
QGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNALGAEAPAFDLLGASVI
GAIQ
HAVGFVRTGGLEATGAAAALQQDDKFMAEFDDL RGAFPRSGDGSAL ETAAAPCAYEARVLDG
IWAAAPYL
HNGSVPTLRELLTPPDERVADFTP GPAYNLEAVGMATEQTAFDHVIETTGCDDLSSGNSRCGH
DYG TALS
DDDKRALLEYLKS I

>AB28

MFTRSVIFSLLFSGAAAPVAAQTVFVDQGDSWSPSLRAAFYTQDQGSRIMPLAWMEALAQTD
GTPFLADG
LARYGYLPMPDRDDGAALPVGFTTNEGMGGTAIGMTCAACHTRQIEVRGTAYRIDGGPAIVDF
QSFLADL
DAAMQAVLDDEAVFADFADKLGETGDRAGLRAELETWSLRFHTLIDGSLPDPGWGPVRLDA
VSMIFNRL
TGLDLGRPEDAYLIPENVMLADAPTRYPFLWNAARQDMTQWPGFHNGNDLLGLARNLGEVY
GVFGEFHP
IKKQGLLFNRDYYVTNNSANWGGLKALEEWIWDIGAPRWPWGLDAGLVEQGGAI FARPTDQGG
CVECHGKR
QGAFRSIFHATWATPILPVGTDERECEILTRTAKTGVLEGAKVPLSSNALGAEAPASDLLGASVI
GAIQ
HAAGFVRTGGLEATGAAAALQKDDKFMAEFDDL RGAFPRSGDGSAL ETAAAPCAYEARVLDGI
WAAAPYL
HNGSVPTLRELLTPPDERVADFTP GPAYDLEAVGMAVEQTAFDHVIETTGCDDLSSGNSRCGH
DYG TALS
DDDKRALLEYLKS I

Input Fasta File for the 16S New Taxonomic Reassignment Tree

>Rhodovulum_viride_JA756

ATGCAGTCGAGCGACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGT
GCCCTTCTCTGCGGAATAGGC
TCGGGAAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGA
TCGGCCCCGCGTCTGATTAGG
TAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATC
AGCCACACTGGGACTGAGAC
ACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGAAACCCT
GATCCAGCCATGCCGCGTGAG
CGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAG
AAGAAGCCCCGGCTAACTC
CGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTA
AAGCGCGCGTAGGCGGACTAT
TAAGTCGGGGGTGAAATCCCGGGGCTCAACCCCGGAAGTGCCTCCGATACTGGTAGTCTA
GAGTTCGAGAGAGGTGAGTG
GAATTCGAGTGAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCG
GCTCACTGGCTCGATACTGA
CGCTGAGGTGCGAAAGCGTGCGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCG
TAAACGATGAATGCCAGTCGT
CGGCAAGCATGCTTGTGCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGAGTACG
GCCGCAAGGTTAAACTCAA
AGGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCG
CAGAACCCTTACCAACCCTTGA
CATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTG
CTGCATGGCTGTGTCAGCT
CGTGTCGTGAGATGTTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCA
GCATTCAGTTGGGCACTCTA
AGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCT
TACGGGTTGGGCTACACACG
TGCTACAATGGCAGTGACAATGGGTAAATCCCAAAAACTGTCTCAGTTCGGATTGTTCTCT
GCAACTCGAGAGCATGAA
GTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGT
ACACACCGCCCGTCACACCA
TGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCTACGGGGGCAGCTGCCA

>Rhodobacter_viridis_JA737

CTGCCATGCGGTGAGCGAGACCTTCGGGTCTAGCGGCGGACGGGTGAGTAACGCGTGG
GAACGTGCCCTTTGCTACGGA
ATAGCCCCGGGAACTGGGAGTAATACCGTATGTGCCCTTCGGGGGAAAGATTTATCGGC
AAAGGATCGGCCCGCGTTGG
ATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGACGATCCATAGCTGGTTTGAGAGG
AGATCAGCCACACTGGGACT
GAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTAGACAATGGGGGAA
ACCCTGATCTAGCCATGCCGC
GTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGGTGGGAAGATAATGACGGTAC
CACCAGAAGAAGCCCCGGCT
AACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCCGAATTACTGG
GCGTAAAGCGCACGTAGGCGG

ATCAGAAAGTCAGAGGTGAAATCCCAGGGCTCAACCTTGAACTGCCTTTGAACTCCTGG
TCTTGAGGTGAGAGAGGT
GAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGA
AGGCGGCTCACTGGCTCGAT
ACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCAC
GCCGTAAACGATGAATGCCA
GTCGTCGGCAGGCATGCCTGTCGGTGACACACCTAACGGATTAAGCATTCCGCCTGGGGA
GTACGGTCGCAAGATTAAAA
CTCAAAGGAATTGACGGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAA
CGCGCAGAACCTTACCAACC
CTTGACATCGAGATCGCGGTTACCAGAGATGGTTTCCTTCAGTTCGGCTGGATCTTAGACA
GGTGCTGCATGGCTGTCGT
CAGCTCGTGTCTGAGATGTTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTTTTCAGT
TGCCATCATTCAAGTTGGGCA
CTCTGGAAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATG
GCCCTTACGGGTTGGGCTAC
ACACGTGCTACAATGGTGGTGACAATGGGCCAATCCCAAAAAGCCATCTCAGTTCGGATTG
GGGTCTGCAACTCGACCCC
ATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGC
CTTGTACACACCGCCCGTCA
CACCATGGGAATTGGGTCTACCCTAAGATGGTGCGCCAACCTGCAAAGGAGGCAGCCAGC
CACGGTAGGCTCAGTGAAGT
GGGTGAAGTCTAACCAACGTGCCCCACCCCC
>Rhodovulum_robiginosum_N2
GAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGTCTTTCGGACTTAGCGGGC
GACGGGTGAGTAACGCGTGG
GAACGTGCCCTTCACTACGGAATAGTCCCGGGAACTGGGTTTAATACCGTATACNCCCTT
CGGGGGAAAGAATTTTCGGT
GAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGA
TCAGTAGCTGGTTTGAGAGG
ATGATCAACCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAG
GAATCTTGGACAATGGGGGCA
ACCCTGATCCAACCATTCGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCGCC
TGGGAAGATAATGACGGTA
CCAGGTAAAGAACCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCT
AGCGTTGTTTCGGAATTACTGG
GCGTAAAGCGCGCGTAGGCGGACTGTAAAGTCAGAGGTGAAATCCCAGGGCTCAACCCTG
GAACGGCCTTTGATACTGGC
AGTCTAGAGGTCTGGAGAGGTGAGTGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATT
CGGAGGAACACCAGTGGCG
AAGGCGGCTCACTGGCCAGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGG
ATTAGATACCCTGGTAGTCCA
CGCCGTAAACGATGAATGCCAGACGTCGGGCAGCATGCTGTTTCGGTGTACACCTAACGG
ATTAAGCATTCCGCCTGGGG
AGTACGGTCGCAAGATTAAACTCAAAGGAATTGACGGGGGGCCCGCACAAAGCGGTGGAGC
ATGTGGTTTAATTCGAAGCA
ACGCGCAGAACCTTACCAACCCTTGACATGGGTATCGCGGTTTTCCGAGAGGGATTCTTC
AGTTCGGCTGGATACCACA
CAGGTGCTGCATGGCTGTCTGTCAGCTCGTGTCTGAGATGTTTCGGTTAAGTCCGGCAACG
ACGCGAACCCACACCCTTAG

TTGCCAGCATTCAAGTTGGGCACTCTAGGGGAACTGCCGATGATAAGTCGGAGGAAGGTGT
GGATGACGTCAAGTCCTCAT
GGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCATCTACAGTGGGTTAATCCCCA
AAAGATGTCTCAGTTCGGAT
TGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGAC
GCGGTGAATACGTTCCCGGG
CCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTACCCGAAGGCCGTGCGCCAA
CCTTTCGAGGGGGCAGCGGA
CCACGGTGAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAA
>Rhodovulum_iodosum_N1
GAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGACTTTCGGTCTTAGCGGCG
GACGGGTGAGTAACACGTGG
GAACATGCCCTTCACTACGGAATAGTCCTGGGAAACCGGGTTTAATACCGTATACNCCCTT
CGGGGGAAAGAATTTTCGGT
GAAGGATTGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCGACGA
TCAGTAGCTGGTTTGAGAGG
ATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAG
GAATCTTGACAATGGGGGAA
ACCCTGATCCAACCATGCCGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCACC
GGTGAAGATAATGACGGTA
GCCGGAGAAGAAACCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGTT
AGCGTTGTTTCGGAATTACTGG
GCGTAAAGCGCGCGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCC
GGAAGTGCCTTTGATACTGGT
AGTCTAGAGTTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATT
CGGAGGAACACCAAGTGGCG
AAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGG
ATTAGATACCTTGGTAGTCCA
CGCCGTAAACGATGAATGCCAGACGTCGGGCAGCATGCTGTTTCGGTGTACACCTAACGG
ATTAAGCATTCCGCCTGGGG
AGTACGGCCGCAAGGTTAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAG
CATGTGGTTTAATTCGAAGCA
ACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTCTCGAGAGAGATTCTTC
AGTTCGGCTGGATCANCNN
CAGGGGCTGCATGGCTGTGCTCAGCTCGTGTGCTGAGATGTTTCGGTTAAGTCCGGCAACG
AGCGCAACCCACACTCTTAG
TTGCCAGCAGTTCGGCTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTG
TGGATGACGTCAAGTCCTCA
TGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTTAATCCCC
AAAACTGTCTCAGTTCGGA
TTGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGAC
GCGGTGAATACGTTCCCGG
GCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCA
ACCTTTCGAGGGGGCAGCTG
GCCACGGTAAGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAA
>Rhodovulum_steppense_DSM21153
TAACACATGCAAGTCGAGCGAGGACTTCGGTCTTAGCGGCGGACGGGTGAGTAACGCGT
GGGAACGTGCCCTTCTGTGCG
GAATAGCCTCGGGAACTGAGAGTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCG
CAGAAGGATCGGCCCGCGTC

TGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGA
GGATGATCAGCCACACTGGG
ACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGG
GCAACCCTGATCCAGCCATGC
CGCGTGAACGATGAAGGCCTTAGGGTTGTAAAGTTCTTTCAGTCGTGAAGATAATGACGGT
AGCGACAGAAGAAGCCCCG
GCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCGGAATTACT
GGGCGTAAAGCGCGCGTAGG
CGGACTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCCTTGATACTG
GTAGTCTAGAGGTCGAGAGA
GGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGG
CGAAGGCGGCTCACTGGCTC
GATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC
CACGCCGTAAACGATGAATG
CCAGACGTCGGCAGGCATGCCTGTCGGTGTACACCTAACGGATTAAGCATTCCGCCTGG
GGAGTACGGCCGCAAGGTTA
AAACTCAAAGGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTGGAAG
CAACGCGCAGAACCTTACCA
ACCCTTGACATGTGTATCGCGGTTCCAGAGATGGTTCCTTCAGTTCGGCTGGATACAACAC
AGGTGCTGCATGGCTGTGC
TCAGCTCGTGTCTGAGATGTTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAG
TTGCCAGCATTTCAGTTGGGC
ACTCTAGGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCAT
GGCCCTTACGGGTTGGGCTA
CACACGTGCTACAATGGCAGTGACAATGGGTAAATCCCAAAAAGCTGTCTCAGTTCGGATT
GTCCTCTGCAACTCGAGGG
CATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGG
CCTTGTACACACCGCCCCGTC
ACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCTTTCGAGGAGGCAGCTG
GCCACGGTAA
>Rhodovulum_bhavnagarensense_JA738
ACCATGGCAGTCGAGCGAAGACTTCGGTCTTAGCGGCGGACGGGTGAGTAACGCGTGGG
AACGTGCCCTTTGCTACGGAA
CAGTCCCGGGAACTGGGTTTAATACCGTATACGCCCTTCGGGGGAAAGAATTCGGCGA
AGGATCGGCCCGCGTCTGAT
TAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGAGAGGAT
GATCAGCCACACTGGGACTG
AGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATGGGGGCAA
CCCTGATCCAGCCATGCCGCG
TGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCGCCAGGGATGATAATGACAGTACCT
GGTAAAGAAACCCCGGCTA
ACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGTTAGCGTTGTTCGGAATTACTGGGC
GTAAAGCGCGCGTAGGCGGG
CTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCCTTGATACTGGTAG
TCTAGAGATCTGGAGAGGTG
AGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAAGAACACCAGTGGCGAA
GGCGGCTCACTGGCCAGATA
CTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACG
CCGTAAACGATGAATGCCAG
TCGTCCGGTAGCATGCTATTTCGGTGACACACCTAACGGAGTAAGCGTTCCGCCTGGGGAG
TACGGCCGCAAGGTTAAAAC

TCACAGGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCRAC
GCGCAGAACCTTACCAACCC
TTGACATCCTGATCGCGGTTTTCCGAGAGGAATTCCTTCAGTTCGGCTGGATCAGTGACAG
GTGCTGCATGGCTGTCGTC
AGCTCGTGTCTGAGATGTTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACCCTTAGTT
GCCAGCATTCAAGTTGGGCAC
TCTAAGGGAAGTCCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGG
CCCTTACGGGTTGGGCTACA
CACGTGCTACAATGGCAGTGACAATGGGTAAATCCCCAAAACTGTCTCAGTTCGGATTGT
CCTCTGCAACTCGAGGGCA
TGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCC
TTGTACACACCGCCCGTCAC
ACCATGGGAGTTGGTTACCCGAAGACGGTGCGCCAACCTTCGAGGAGGCAGCTGGCAAC
GTAA

>Rhodovulum_euryhalinum_DSM4868

AACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAGGACTTCGGTCCTAGCG
GCGGACGGGTGAGTAACGCGT
GGGAACGTGCCCCAAAGGTACGGAATAGCCTCGGGAACTGAGAGTAATACCGTATGAGCC
CTTCGGGGGAAAGATTTATC
GCCTTTGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTA
CGATCAGTAGCTGGTTTNAAG
AGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGT
GAGGAATCTTGACAATGGGG
GAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTT
CAGTCGTGAAGATAATGACG
GTAGCGACAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGG
GCTAGCGTTGTTTCGGAATTAC
TGGGCGTAAAGCGCGCGTAGGCGGATTGGAAAGTTGGGGGTGAAATCCCAGGGCTCAAC
CCTGGAAGTGCCTCNAAAAGT
ACCATCTAGAGGTTCGAGAGAGGTGAGTGGAATTCGAGTGTAGAGGTGAAATTCGTAGA
TATTCGGAGGAACACCAAGT
GCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAAC
AGGATTAGATACCCTGGTAGT
CCACGCCGTAAACGATGAATGCCAGACGTCGGGCAGCATGCTGTTCCGGTGTACACCTAA
CGGATTAAGCATTCCGCCTG
GGGAGTACGGCCGCAAGGTTAAAGTCAAAGGAATTGACGGGGGCCCCGCACAAGCGGTG
GAGCATGTGGTTTAATTCGAA
GCAACGCGCAGAACCTTACCAACCCTTGACATGTGTATCGCGGTTCCAGAGATGGTTCCTT
CAGTTCGGCTGGATACAAC
ACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCTGAGATGTTCCGGTTAAGTCCGGCAAC
GAGCGCAACCCACACTCTTA
GTTGCCAGANTAAGTTGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGT
GGATGACGTCAAGTCCTCAT
GGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGGTAAATCNCCA
AAAGCTGTCTCAGTTCGGAT
TGTCTCTGCAACTCGAGGGCATGAAGTNGGAATCGCTAGTAATCGCGTAACAGCATGAC
GCGGTGAATACGTTCCCGGG
CCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAA
CCTGTTTCGAGGAGGCAGCC
GGCCACGGTAAGCTCAGCGACTGGGGTG

>Rhodobacter_azotoformans_KA25

CTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGTCTTCGG
ACTTAGCGGCGGACGGGTGA
GTAACGCGTGGGAACATGCCCAAAGGTACGGAATAGCCCCGGGAACTGGGAGTAATACC
GTATGTGCCCTTCGGGGGAA
AGATTTATCGCCTTTGGATTGGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTAC
CAAGCCGACGATCCATAGC
TGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAG
GCAGCAGTGGGGAATCTTAG
ACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCCTTAGGGTTGTA
AAGATCTTTCAGGTGGGAAG
ATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAA
CGGAGGGGGCTAGCGTTATT
CGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTGGAAAGTCAGGGGTGAAATCCCG
GGGCTCAACCCCGGAACTGCC
TTTGAAACTCCCAGTCTTGAGGTGAGAGAGGTGAGTGAATTCCGAGTGTAGAGGTGAA
ATTCGTAGATATTCGGAGGA
ACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGG
GGAGCAAACAGGATTAGATAC
CCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCCGGTGA
CACACCTAACGGATTAAGCA
TTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGAATTGACGGGGGGCCCGCA
CAAGCGGTGGAGCATGTGGTT
TAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGCGGTTCCAGAG
ATGTTTCCTTCAGTTCGGCT
GGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTGTCGTGAGATGTTCCGGTTAAG
TCCGGCAACGAGCGCAACCC
ACGTCCTCAGTTGCCAGCATTTCAGTTGGGCACTCTGGGGAACTGCCGGTGATAAGCCGG
AGGAAGGTGTGGATGACGTC
AAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAATGGG
TTAATCCCCAAAAGCTGTCT
CAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATA
CGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTCTACCCGAAGGC
GGTGCGCCAACCTCGCAAGA
GGAGGCAGCCGACCACGGTAGGATCAGTGAAGTGGGGTGAAGTCGTACAAAGGTAACCA
>Rhodovulum_imhoffii_DSM18064
NANCCGGCTCAGAACGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGGCT
TCGGCCTTAGCGGCGGACGGG
TGAGTAACGCGTGGGAACGTGCCCTTCACTACGGGATAGTCCTGGGAACTGGGTTTAAT
ACCGTATACGCCCTTTTGGG
GAAAGAATTTCCGGTGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCC
TACCAAGCCGACGATCAGTA
GCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGAKTCTACGGG
AGGCAGCAGTGAGGAATCTT
GGACAATGGGGGCAACCCTGATCCAGCCATGCCGCGTGAACGATGAAGGCTTTAGGGTTG
TAAAGTTCTTTCACCGGGAA
GATAATGACGGTAGCCGGAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAA
TACGGAGGGGGCTAGCGTTGT
TCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCG
GGGCTCAACCCCGGAACTGC

CTTTGATACTGGTAGTCTAKAGTTCTGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAA
ATTCGTAGATATTCGGAGG
AACACCAAGTGGCGAAGGCGGCTCACTGGCCAGATACTGACGCTGAGGTGCGAAAGCGTG
GGGAGCAAACAGGATTAGATA
CCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGASGTCGGGCASCATGCTGTTCCGGTG
TCCACCTAACGGATTAAGCA
TTCCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAAGGAATTGACGGGGGCCCCGCAC
AAGCGGTGGAGCATGTGGTT
TAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAG
AGGAATTCCTTCAGTTCGG
CTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCCGGTTA
AGTCCGGCAACGAGCGCAAC
CCACGTCTTTAGTTGCCAGCAGTTCGGCTGGGCACTCTAAAGAACTGCCGATGATAAGTC
GGAGGAAGGTGTGGATGAC
GTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCGATGACAAT
GGGTAAATCCCTAAAAGTCG
TCTCAGTTCGGATTGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCG
TAACAGCATGACGCGGTGA
ATACGTTCCCGGGCCTTGACACACCGCCGTCACACCATGGGAGTTGGGTTTACCCGAA
GACGGTGCGCAACCTGCTC
GCAGGAGGCAGCCGGCCACGGTADGTCAGGAKGGGGGA
>Roseibaca_ekhonensis_CECT_7235
TGGCGGCAGGCTTAACACATGCAAGTCGAGCGAAGCCTTCGGGCTTAGCGGCGGACGGG
TGAGTAACGCGTGGGAACGTG
CCCTTTGCTACGGAATAGTCCCTGGAAACGGGGTTTAATACCGTATGTGCTCTTCGGAGGA
AAGAATTTCCGCAAAGGAT
CGGCCCCGCGTTGGATTAGGTAGTTGGTGGGGTAAAGGCCTACCAAGCCGACGATCCATAG
CTGGTTTGAGAGGATGATCA
GCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTT
AGACAATGGGGGCAACCCTGA
TCTAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCGCTGGGGAA
GATAATGACGGTACCCAGTA
AAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTG
TTCGGAATTACTGGGCGTAAA
GCGCGCGTAGGCGGATTGGAAAGTTGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTG
CCTCCAAAATATCAGTCTTG
AGTTCGAGAGAGGTGAGTGGAATTGCGAGTGTAGAGGTGAAATTCGTAGATATTCGCAGG
AACACCAAGTGGCGAAGGCGG
CTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTAGATA
CCCTGGTAGTCCACACCGTA
AACGATGAATGCCAGACGTCGGGCAGCATGCTGTTCCGGTGTCACACCTAACGGATTAAGC
ATTCCGCCTGGGGAGTACGG
TCGCAAGATTAAACTCAAAGGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGT
TTAATTCGAAGCAACGCGCA
GAACCTTACCAACCCTTGACATGGATATCGCGGGACCGGAGACGGTTCTTTCAGTTCGGC
TGGATATCACACAGGTGCTG
CATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCCGGTTAAGTCCGGCAACGAGCGCAACC
CACGTCCCCAGTTGCCAGCA
TTCAGTTGGGCACTCTGAGGAGACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGT
CAAGTCCTCATGGCCCTTAC

GGGTTGGGCTACACACGTGCTACAATGGTGGTGACAGAGGGATAATCCCTAAAAGCCATC
TCAGTTCGGATTGGGGTCTG
CAACTCGACCCCATGAAGTCGGAATCGCTAGTAATCGCGGAACAGCATGCCGCGGTGAAT
ACGTTCCCGGGCCTTGTACA
CACCGCCCGTCACACCATGGGAGTTGGGTCTACCCGAAGGTGGTGCGCTAACCAGCAATG
GAGGCAGCCAACCACGGTAG
GCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATC
ACCTCC

>Rhodovulum_marinum_DSM_18063

TCGAGCGAAGCCTTCGGGCTTAGCGGCGGACGGGTGAGTAACGCGTGGAACGTGCCCT
TCTCTGCGGAATAGCCACTGG
AAACGGTGAGTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGC
CCGCGTTAGATTAGGTAGTT
GGTGGGGTAATGGCCTACCAAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCA
CACTGGGACTGAGACACGGC
CCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGACAATGGGGGCAACCCTGATCCA
GCCATGCCGCGTGAACGATG
AAGGCCTTAGGGTTGTAAAGTTCTTTCAGCKGGAAGATAAGACGGTACCAGCAGAAGAAG
CCCCGGCTAACTCCGTGCCA
GCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCCGAATTACTGGGCGTAAAGCGCG
CGTAGGCGGATTGGAAAGTTG
GGGGTGAAATCCCAGGGCTCAACCCTGGAAGTGCCTCCAAAACCTACCAGTCTAGAGTTCG
AGAGAGGTGAGTGGAATTCC
GAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACT
GGCTCGATACTGACGCTGAG
GTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGA
TGAATGCCAGACGTCGGGCA
GCATGCTGTTCCGGTGTCCACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAA
GGTAAAACTCAAAGGAATT
GACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTGAAGCAACGCGCAGAACC
TTACCAACCCTTGACATCCTG
ATCGCGGTTTCTCGAGAGAGATTCTTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATG
GCTGTCGTGACGCTCGTGTG
TGAGATGTTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCAGTTC
GGCTGGGCACTCTAAGAGAA
CTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGG
TTGGGCTACACACGTGCTAC
AATGGCAGTGACAATGGGTAAATCCCCAAAAGCTGTCTCAGTTCGGATTGTCCTCTGCAAC
TCGAGGGCATGAAGTCGGA
ATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACAC
CGCCCGTCACACCATGGGAG
TTGGGTTTACCCGAAGACGGTGCGCCAACCTTTCGAGGGGGCAGCTGGCCACGGTAAGC
TCAGCGACTGGGGTRAAGTCG
TAACAAGGGTAACCAA

>Rhodovulum_adriaticum_DSM_2781

CCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAAGCCTTC
GGGCTTAGCGGCGGACGGGTG
AGTAACGCGTGGGAACGTGCCCTTCACTACGGAATAGTCCCGGGAAACTGGGTTTAATAC
CGTATACGCCCTTCGGGGGA
AAGAATTTCCGGTGAAGGATCGGCCCCGCGTTAGATTAGGTAGTTGGTGGGGTAATGGCCTA
CCAAGCCTACGATCTATAGC

TGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAG
GCAGCAGTGAGGAATCTTAG
ACTAGATGGGGGCAACCCTGATCTAGCCTGCCGCGTGAACGATGAAGGCCTTAGGGTTGT
AAAGTTCTTTCGCCTGGGAA
GATAATGACGGTACCAGGTAAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAAT
ACGGAGGGGGCTAGCGTTGT
TCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGGCTGTTAAGTCAGAGGTGAAATCCCA
GGGCTCAACCCTGGAAGTGC
CTTTGATACTGGCAGCCTGGAGGTCTGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGA
AATTCGTAGATATTCGGAAG
AACACCAAGTGGCGAAGGCGGCTCACTGGCCAGATACTGACGCTGAGGTGCGAAAGCGTG
GGGAGCAAACAGGATTAGATA
CCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATGCTGTTCCGGTG
ACACACCTAACGGAGTAAGC
ATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGGCCCCG
ACAAGCGGTGGAGCATGTGGT
TTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCCTGTCGCGGTTTTCCGA
GAGGAATTCCTTCAGTTCTG
GCTGGACAGGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTGTCGTGAGATGTTCCGGT
AAGTCCGGCAACGAGCGCAA
CCCACGTCCTTAGTTGCCAGCATTGAGTTGGGCACTCTAAGGAACTGCCGATGATAAGTC
GGAGGAAGGTGTGGATGAC
GTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGACAAT
GGGTTAATCCCCAAAACTG
TCTCAGTTTCGGATTGTCCTCTGCAACTCGAGGGCATGAAGTCGGAATCGCTAGTAATCGCG
TAACAGCATGACGCGGTGA
ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCCGAA
GACGGTGCGCCAACCTGTTC
GCAGGAGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAGTC
>Rhodovulum_kholense_DSM_19783
GGCCTACACATGCAAGTCGAGCGAACCCTTCGGGGTTAGCGGCGGACGGGTGAGTAACG
CGTGGGAACGTGCCCTTCTCT
GCGGAATAGGCTCGGGAACTGGGTTTAATACCGCATACGCCCTTCGGGGGAAAGATTTA
TCGGAGAAGGATCGGCCCGC
GTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACGATCAGTAGCTGGTTTGA
GAGGATGATCAGCCACACT
GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGGACAATG
GGGGAAACCCTGATCCAGCCA
TGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGCTGGGAAGATAATGA
CGGTACCAGCAGAAGAAGCC
CCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTCCGGAAT
TACTGGGCGTAAAGCGCACGT
AGGCGGATTGAAAAGTTGGGGGTGAAATCCCGGGGCTCAACCCCGGAACTGCCTCCAAA
ACTCCAGTCTTGAGGTGAG
AGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAG
TGGCGAAGGCGGCTCACTGG
CTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTA
GTCCACGCCGTAAACGATGA
ATGCCAGTCGTCGGGTAGCATGCTATTCGGTGACACACCTAACGGATTAAGCATTCCGCCT
GGGGAGTACGGCCGCAAGG

TTAAAACTCAAAGGAATTGACGGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCG
AAGCAACGCGCAGAACCTTA
CCAACCCTTGACATCCTGATCGCGGTTACCCGAGAGGGTTTCCTTCAGTTCGGCTGGATCA
GTGACAGGTGCTGCATGGC
TGTCGTCAGCTCATGTGCTGAGATGTTTCGGTTAAGTCCGGCAACGAGCGCAACCCACACT
CTTAGTTGCCAGCATTCACT
TGGGCACTCTAAGAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGTC
CTCATGGCCCTTACGGGTTG
GGCTACACACGTGCTACAATGGCAGTGACAATGGGTAAATCCCAAAAAACTGTCTCAGTTC
GGATTGTTCTCTGCAACTC
GAGAGCATGAAGTCGGAATCGCTAGTAATCGCGTAACAGCATGACGCGGTGAATACGTTC
CCGGGCCTTGTACACACCGC
CCGTACACCATGGGGAGTTGGGTTTACCCGAAGACGGTGCGCCAACCCTTACGGGGGG
CAGCTGGCCACGGTAAGCTCA
GCGACTGGGG

>Rhodobacter_capsulatus_SB1003

AGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
GAGACCTTCGGGTCTAGCGG
CGGACGGGTGAGTAACGCGTGGAACGTGCCCTTTGCTACGGAATAGCCCCGGGAAACT
GGGAGTAATACCGTATGTGCC
CTTCGGGGGAAAGATTTATCGGCAAAGGATCGGCCCCGCGTTGGATTAGGTAGTTGGTGGG
GTAATGGCCTACCAAGCCGA
CGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGAC
TCCTACGGGAGGCAGCAGTG
GGGAATCTTAGACAATGGGGGAAACCCTGATCTAGCCATGCCGCGTGAGCGATGAAGGCC
TTAGGGTTGTAAAGCTCTTT
CAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCA
GCCGCGGTAATACGGAGGGGG
CTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCAGAAAGTCAGAGG
TGAAATCCCAGGGCTCAACC
TTGGAAGTGCCTTTGAAACTCCTGGTCTTGAGGTCGAGAGAGGTGAGTGGAATTCCGAGT
GTAGAGGTGAAATTCGTAGA
TATTCGGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTG
CGAAAGCGTGGGGAGCAAACA
GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGCAGGCATG
CCTGTCGGTGACACACCTAA
CGGATTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAAGGAATTGACG
GGGGCCCCGCACAAGCGGTGG
AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCGAGATCGC
GGTTACCAGAGATGGTTTC
CTTCAGTTCGGCTGGATCTTAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCTGAG
ATGTTTCGGTTAAGTCCGGCA
ACGAGCGCAACCCACACTTTCACTTCCATCATTCACTTGGGCACTCTGGAAGAACTGCC
GATGATAAGTCGGAGGAAGG
TGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGG
TGGTGACAATGGGCCAATCC
CAAAAAGCCATCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGC
TAGTAATCGCGTAACAGCAT
GACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGG
TCTACCCTAAGATGGTGCGC

CAACCAGCAATGGAGGCAGCCAGCCACGGTAGGCTCAGTGAAGTCGTAAC
AAGGTAGCCGTAGGGGAACCT
GCGGCTGGATCACCT
>Rhodospirillum_rubrum_ATCC_11170
AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAAC
GCATCCTTCGGGATGAGTGG
CGCACGGGTGAGTAACACGTGGGAACGTACCTTGGAGTGCGGAATAATCTTTGGAAACGA
GGACTAATACCGCATACGCC
CTTAGGGGGAAAGATTTATCGCTCCAAGATCGGCCCCGCGTCCGATTAGCTAGTTGGCGGG
GTAATGGCCCACCAAGGCGA
CGATCGGTAGCTGGTCTGAGAGGATGGCCAGCCACACTGGGACTGAGACACGGCCCAGA
CTCCTACGGGAGGCAGCAGTG
GGGAATATTGCGCAATGGGGGCAACCCTGACGCAGCCATGCCGCGTGAGTGAAGAAGGC
CTTCGGGTTGTAAAGCTCTTT
CGGGTGTGAAGATGATGACGGTAACACCAGAAGAAGCCCCGGCTAACTTCGTGCCAGCAG
CCGCGGTAATACGAAGGGGG
CAAGCGTTGTTTCGGAATTACTGGGCGTAAAGAGCGCGTAGGCGGTCTGATTAGTCAGAGG
TGAAATCCCAGAGCTCAACT
TTGGAAGTGCCTTTGATACTGTTAGACTAGAATCCGTGAGAGGGTGGTGGAAATCCCAGTG
TAGAGGTGAAATTCGTAGA
TATTGGGAGGAACACCAGTGGCGAAGGCGGCCACCTGGCGCGGTATTGACGCTGAGGCG
CGAAAGCGTGGGGAGCAAACA
GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAGATGTCGGGGTACATG
TACCTCGGTGTCGCAGCTAA
CGCATTAAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACG
GGGGCCCGCACAAAGCGGTGG
AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAGCCCTTGACATCCCGTGACA
CTTCCAGAGATGGAAGGTT
CCCTTCGGGGACACGGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGT
TGGGTTAAGTCCCGCAACGA
GCGCAACCCTCATCTTCAGTTGCCAGCAAGTAACGTTGGGCACTCTGAAGAGACTGCCGG
TGACAAGCCGGAGGAAGGTG
GGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACACACGTGCTACAATGGCG
CCTACAATGGGCAGCGACCT
CGCGAGGGGAAGCTAATCTCCAAAAGGCGTCTCAGTTCGGATTGCACTCTGCAACTCGGG
TGCATGAAGTCGGAATCGCT
AGTAATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCTTGACACACCGCCCCG
TCACACCATGGGAGTTGGTT
CTACCCGAAGACGGTACGCTAACCGCAAGGAGGCAGCCGGCCACGGTAGGGTCAGCGAC
TGGGGTGAAGTCGTAACAAGG
TAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT
>Rhodobacter_sphaeroides_2.4.1
CAACTTGAGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAAGTCTTCGGAC
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCG
GGAAACTGGGAGTAATACCGA
ATGTGCCCTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCCGCGTTGGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC

AGCAGTGGGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGAT
GAAGGCCTTAGGGTTGTAAA
GATCTTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAA
GTCAGAGGTGAAATCCCAGGG
CTCAACCCTGGAAGTGCCTTTGAAACTCCCGATCTTGAGGTGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGG
CAGCATGCTGTTCCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATG
GCGATCGCGGTTCCAGAGAT
GGTTCCTTCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTC
GTGAGATGTTCCGGTTAAGTC
CGGCAACGAGCGCAACCCACGTCCTTAGTTGCCAGCATTGAGTTGGGCACTCTAGGGAAA
CTGCCGGTGATAAGCCGGAG
GAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTAC
AATGGCAGTGACAATGGGTT
AATCCCAAAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAA
TCGCTAGTAATCGCGTAAC
AGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGGAA
TTGGTTCTACCCGAAGGCGG
TGCGCCAACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGAAGTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>Rhodobacter_sphaeroides_KD131
AGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
GAAGTCTTCGGACTTAGCGG
CGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCGGGAACT
GGGAGTAATACCGAATGTGCC
CTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGTTGGTGGG
GTAATGGCCTACCAAGCCGA
CGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGAC
TCCTACGGGAGGCAGCAGTG
GGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCC
TTAGGGTTGTAAAGATCTTT
CAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCA
GCCGCGGTAATACGGAGGGGG
CTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAAGTCAGAGG
TGAAATCCCAGGGGCTCAACC
CTGGAAGTGCCTTTGAAACTCCCGATCTTGAGGTGAGAGAGGTGAGTGGAATTCCGAGT
GTAGAGGTGAAATTCGTAGA
TATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTG
CGAAAGCGTGGGGAGCAAACA
GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATG
CTGTTCCGGTGACACACCTAA

CGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACCTCAAAGGAATTGACG
GGGGCCCGCACAAAGCGGTGG
AGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGC
GGTTCCAGAGATGGTTCCCT
TCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGAT
GTTTCGGTTAAGTCCGGCAAC
GAGCGCAACCCACGTCCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGGAAACTGCCGG
TGATAAGCCGGAGGAAGGTG
TGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCA
GTGACAATGGGTAAATCCCA
AAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTA
GTAATCGCGTAACAGCATGA
CGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTC
TACCCGAAGGCGGTGCGCCA
ACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGAAGTGGGGTGAAGTCGTAAC
AAGGTAGCCGTAGGGGAACCT
GCGGCTGGATCACCT
>Rhodobacter_sphaeroides_ATCC17025
AGAGTTTGATCCTGGCTCAGAATGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
GAAGTCTTCGGAAGTTAGCGG
CGGACGGGTGAGTAACGCGTGGGAACATGCCCAAAGGTACGGAATAGCCCCGGGAAACT
GGGAGTAATACCGTATGTGCC
CTTCGGGGGAAAGATTTATCGCCTTTGGATTGGCCCGCGTTGGATTAGGTAGTTGGTGGG
GTAATGGCCTACCAAGCCGA
CGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGAC
TCCTACGGGAGGCAGCAGTG
GGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGATGAAGGCC
TTAGGGTTGTAAAGATCTTT
CAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCA
GCCGCGGTAATACGGAGGGGG
CTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTGGAAAGTCAGGGG
TGAAATCCCGGGGCTCAACC
CCGGAAGTGCCTTTGAAACTCCCAGTCTTGAGGTGCGAGAGAGGTGAGTGGAAATCCGAGT
GTAGAGGTGAAATTCGTAGA
TATTCGGAGGAACACAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTG
CGAAAGCGTGGGGAGCAAACA
GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGCAGCATG
CTGTTCCGGTGACACACCTAA
CGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTAAAACCTCAAAGGAATTGACG
GGGGCCCGCACAAAGCGGTGG
AGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATGGCGATCGC
GGTTCCAGAGATGGTTCCCT
TCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGAT
GTTTCGGTTAAGTCCGGCAAC
GAGCGCAACCCACGTCCTCAGTTGCCAGCATTAGTTGGGCACTCTGGGGAAACTGCCGG
TGATAAGCCGGAGGAAGGTG
TGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCA
GTGACAATGGGTAAATCCCA
AAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAATCGCTA
GTAATCGCGTAACAGCATGA

CGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAATTGGTTC
TACCCGAAGGCGGTGCGCCA
ACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGAAGTGGGGTGAAGTCGTAAC
AAGGTAGCCGTAGGGGAACCT
GCGGCTGGATCACCTCCTTT
>DSM2351
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACCTGCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCTTGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTGAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTTCTAA
>DSM1374
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC

AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTTCGGAGGAACACCAGTGGCGAAGGCGGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGCAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTGAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>SNK001
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTTCGGAGGAACACCAGTGGCGAAGGCGGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGA
ATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB14
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGTTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGA
ATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG

TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB26

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCAGTAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCGGGGGTGAAATCCCGGGG
CTCAACCCCCGGAAGTGCCTCCGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAAT
TCCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAAGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCAAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB30

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAAC
>SE1
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA

GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>IM796
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGTTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG

TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB35

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGAGGAACACCAAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB23

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC

ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB22
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAAGTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT

TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB16

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCTTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTTG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG

AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTGCGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAA
GTCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB20

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACCTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCACTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB28

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCCTTCGGGG

TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGATAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTGAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>AB18
CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG

CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB17

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG

TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB15

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGGG
CTCAACCCCGGAACCTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCTGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTCCGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAAGCTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCCAACCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>AB33

CAACTTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAACCCTTCGGGG
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTCTCTGCGGAATAGGCTCG
GGAAACTGGGTTTAATACCGC
ATACGCCCTTCGGGGGAAAGATTTATCGGAGAAGGATCGGCCCGCGTTAGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCTACGATCTATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGAGGAATCTTGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGAT
GAAGGCCTTAGGGTTGTAAA
GCTCTTTCAGTCGTGAAGATAATGACGGTAGCGACAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAG
TCAGGGGTGAAATCCCGGG
CTCAACCCCGGAACCTGCCTTTGATACTGGTAGTCTAGAGTTCGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGC
AAGCATGCTTGTCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGGAAGCAACGCGCAGAACCTTACCAACCCTTGACATC
CTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGCTGGATCAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTTCGGTTAAG
TCCGGCAACGAGCGCAACCCACACTCTTAGTTGCCAGCATTGAGTTGGGCACTCTAGGAG
AACTGCCGATGATAAGTCGG
AGGAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCT
ACAATGGCAGTGACAATGGG
TTAATCCCTAAAACTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGG
AATCGCTAGTAATCGCGTA
ACAGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGG
AGTTGGGTTTACCCGAAGAC
GGTGCGCAACCCTTACGGGGGGCAGCTGGCCACGGTAAGCTCAGCGACTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT

>2_4_1

CAACTTGAGAGTTTGATCCTGGCTCAGAAATGAACGCTGGCGGCAGGCCTAACACATGCAA
GTCGAGCGAAGTCTTCGGAC
TTAGCGGCGGACGGGTGAGTAACGCGTGGGAACGTGCCCTTTGCTTCGGAATAGCCCCG
GGAAACTGGGAGTAATACCGA
ATGTGCCCTTTGGGGGAAAGATTTATCGGCAAAGGATCGGCCCGCGTTGGATTAGGTAGT
TGGTGGGGTAATGGCCTACC
AAGCCGACGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGC
AGCAGTGGGGAATCTTAGACAATGGGCGCAAGCCTGATCTAGCCATGCCGCGTGATCGAT
GAAGGCCTTAGGGTTGTAAA
GATCTTTCAGGTGGGAAGATAATGACGGTACCACCAGAAGAAGCCCCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACG

GAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGGAAA
GTCAGAGGTGAAATCCAGGG
CTCAACCCTGGAAGTGCCTTTGAAACTCCCGATCTTGAGGTGAGAGAGGTGAGTGGAATT
CCGAGTGTAGAGGTGAAAT
TCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCT
GAGGTGCGAAAGCGTGGGA
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGG
CAGCATGCTGTTCCGGTGACA
CACCTAACGGATTAAGCATTCCGCCTGGGGAGTACGGCCGCAAGGTTAAACTCAAAGGA
ATTGACGGGGGCCCCGCACAA
GCGGTGGAGCATGTGGTTTAATTGCAAGCAACGCGCAGAACCTTACCAACCCTTGACATG
GCGATCGCGGTTCCAGAGAT
GGTTCCTTCAGTTCGGCTGGATCGCACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTC
GTGAGATGTTCCGGTTAAGTC
CGGCAACGAGCGCAACCCACGTCCTTAGTTGCCAGCATTGAGTTGGGCACTCTAGGGAAA
CTGCCGGTGATAAGCCGGAG
GAAGGTGTGGATGACGTCAAGTCCTCATGGCCCTTACGGGTGGGCTACACACGTGCTAC
AATGGCAGTGACAATGGGT
AATCCCAAAAAGCTGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCCATGAAGTCGGAA
TCGCTAGTAATCGCGTAAC
AGCATGACGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGGAA
TTGGTTCTACCCGAAGGCGG
TGCGCCAACCTCGCAAGAGGAGGCAGCCGACCACGGTAGGATCAGTGAAGTGGGGTGAAG
TCGTAACAAGGTAGCCGTAGG
GGAACCTGCGGCTGGATCACCTCCTTT
>TIE1
CAACTTGAGAGTTTGATCCTGGCTCAGAGCGAACGCTGGCGGCAGGCTTAACACATGCAA
GTCGAACGGGCGTAGCAATA
CGTCAGTGGCAGACGGGTGAGTAACGCGTGGGAACGTACCTTTTGGTTCGGAACAACACA
GGGAACTTGTGCTAATACC
GGATAAGCCCTTACGGGGAAAGATTTATCGCCGAAAGATCGGCCCCGCGTCTGATTAGCTA
GTTGGTGAGGTAATGGCTCA
CCAAGGCGACGATCAGTAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGAGACAC
GGCCCAAACCTCCTACGGGAG
GCAGCAGTGGGGAATATTGGACAATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGT
GATGAAGGCCCTAGGGTTGTA
AAGCTCTTTTGTGCGGGAAGATAATGACGGTACCGCAAGAATAAGCCCCGGCTAACTTCGT
GCCAGCAGCCGCGGTAATA
CGAAGGGGGCTAGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTTTCT
AAGTCAGAGGTGAAAGCCTGG
AGCTCAACTCCAGAAGTGCCTTTGATACTGGAAGTCTTGAGTATGGCAGAGGTGAGTGGA
CTGCGAGTGTAGAGGTGAA
ATTCGTAGATATTCGCAAGAACACCAGTGGCGAAGGCGGCTCACTGGGCCATTACTGACG
CTGAGGCACGAAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGCCGTTA
GTGGGTTTACTCACTAGTGG
CGCAGCTAACGCTTTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAAG
GAATTGACGGGGGCCCCGCAC
AAGCGGTGGAGCATGTGGTTTAATTGACGCAACGCGCAGAACCTTACCAGCCCTTGACA
TGTCCAGGACCGGTGCGAGA

GACGCGACCTTCTCTTCGGAGCCTGGAGCACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTGGGTTAAG
TCCCGCAACGAGCGCAACCCCCGTCCTTAGTTGCTACCATTTAGTTGAGCACTCTAAGGAG
ACTGCCGGTGATAAGCCGC
GAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACACACGTG
CTACAATGGCGGTGACAATGG
GAAGCTAAGGGGCGACCCTTCGCAAATCTCAAAAAGCCGTCTCAGTTCGGATTGGGCTCT
GCAACTCGAGCCCATGAAGT
TGGAATCGCTAGTAATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCTTGAC
ACACCGCCCGTCACACCATG
GGAGTTGGCTTTACCTGAAGACGGTGCGCTAACCAGCAATGGGGGCAGCCGGCCACGGT
AGGGTCAGCGACTGGGGTGAA
GTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT
>geobactersulfurreducens
AAACTGGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCGTGCCTAACACATGCAA
GTCGAACGTGATCCATCCTT
CGGGGTGGTGAAAGTGGCGCACGGGTGAGTAACGCGTGGATAATCTGCCCCGAGGATTG
GGATAACATCTCGAAAGGGGT
GCTAATACCGAATAAGCCACGGGTCTACGGATCTTGCGGGAAAAGGGGGGGACTTTG
GGCCTCCTGTCTTCGGATGA
GTCCGCGTACCATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGATGGTTAGCT
GGTCTGAGAGGATGATCAGC
CACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGC
CAATGGGGGAAACCCTGACG
CAGCAACGCCGCGTGGGTGATGAAGGCCTTCGGGTGCTAAAGCCCTGTCGGGAGGGAAG
AAATGATTGAGAGCTAATACC
TCTTGGTCTTGACGGTACCTCCGAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGG
TAATACGGAGGGTGCAAGCG
TTGTTGGAATTATTGGGCGTAAAGCGCGTGTAGGCGGTCTTTTAAGTCTGATGTGAAAGC
CCCGGGCTCAACCTGGGAA
GTGCATTGGAACTGGGAGACTTGAGTACGGGAGAGGAGAGTGGAATTCCTAGTGTAGGA
GTGAAATCCGTAGATATTAG
GAGGAACACCGGTGGCGAAGGCGGCTCTCTGGACCGATACTGACGCTGAGACGCGAAAG
CGTGGGTAGCAAACAGGATTA
GATACCCTGGTAGTCCACGCCGTAAACGATGAGTACTAGGTGTTGCGGGTATTGACCCCT
GCAGTGCCGTAGCTAACGCA
TTAAGTACTCCGCCTGGGAAGTACGGTCGCAAGACTAAAACCTCAAAGGAATTGACGGGGG
CCCGCACAAGCGGTGGAGCA
TGTGGTTTAATTCGACGCAACGCGCAGAACCTTACCTGGGCTTGACATCCACGGAACCCTC
CCGAAACGGAGGGGTGCC
TTCGGGGAGCCGTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTGCTGAGATGTTG
GGTTAAGTCCCGCAACGAGC
GCAACCCTTATCATCAGTTGCCATCATTAAAGTTGGGCACTCTGATGAAACTGCCGGTGTCA
AACC GGAGGAAGGTGGGGA
TGACGTCAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGCCGGTA
CAAAGAGTCGCAATGCCGCG
AGGTGGAGCCAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGAGTCTGCAACTCGACTCC
ATGAAGTCGGAATCGCTAGT
AATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCA
CACCACGGGAGTCGACTGGT

CCCGAAGTGCGTGAGCTAACCAGCAATGGAGGCAGCGTCCTAAGGAATGGTCCGGTGACT
GGGGTGAAGTCGTAACAAGGT
AGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT
>Pseudomonas_aeruginosa
GAACTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCAGCAGGGGCCTTCAACACAT
GCAAGTCGAG
CTTATGAAGGGAGCTTGCCTTGGATTGAGCGGCGGACGGGTGAGTAATGCCTAGGAATCT
GCCTGGTAGT
GGGGGATAACGTCCGGAACGGCCGCTAATACCGCATACGTCCTGAGGGAGAAAGTCGG
GGATCTTCGGA
CCTCACGCTATCAGATGAGCCTAGGTGCGATTAGCTAGTTGGTGGGGTAAAGGCCTACCA
AGGCGACGAT
CCGTAAGTGGTCTGAGAGGATGATCAGTCACACTGGAAGTGAAGACACGGTCCAGACTCCT
ACGGGAGGCA
GCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGTGTGAA
GAAGGTCTTCG
GATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCAGTAAGTTAATACCTTGCTGTTTGACGT
TACCAACAG
AATAAGCACCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAA
TCGGAATTAC
TGGGCGTAAAGCGCGCGTAAGTGGTTCAGCAAGCTTGATGTGAAATCCCCGGGCTCAACC
TGGGAAGTGC
ATCCAAAAGCTACTGAGCTAGAGTACGGTAGAGGTGGTAGAATTCCTGTGTAGCGGTGAA
ATGCGTAGA
TATAGGAAGGAACACCAAGTGGCGAAGGCGACCACCTGGACTGTACTGACACTGAGGTGCG
AAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTGCGACTAGCCGTTG
GGATCCTTGA
GATCTTAGTGGCGCACGTAACGCGATAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGT
TAAAACTCAAA
TGAATTGACGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACC
TGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAACAGAGACAC
AGGTGCTGCA
TGGCTGTCGTCAGCTCGTGTCTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCT
TGTCCTTAGT
TACCAGCACCTCGGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTG
GGGATGACGTC
AAGTCATCATGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTCGGTACAAAGGG
TTGCCAAGCC
GCGAGTGGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGAC
TGCGTGAAGT
CGGAATCGCTAGTAATCGTGAATCAGAATGTCACGGTGAATACGTCCCCGGGCCTTGTACA
CACCGCCCG
TCACACCATGGGAGTGGGTTGCTCCAGAAGTAGCTAGTCTAACCGCAAGGGGGACGGTTA
CCACGGAGTG
ATTCATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCAC
CTCCTTA
>Nitrosococcus_oceani
AGGAGGTGATCCAGCCGAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCA
TTGACCACAC

CGTAGTAAGCGTCCTCCCGAAGGTTAAACTACCCACTTCTGGTGCAACCAACTCCCATGGT
GTGACGGGC
GGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATTGCTGATCCGCGATTACTAGCG
ATTCCGACTT
CATGCAGTCGAGTTGCAGACTGCAATCCGGACTACGACCGGTTTTCTGAGATTAGCTCCG
CATCGCTGCT
TGGCCACCCTCTGTACCGGCCATTGTAGCACGTGTGTAGCCCTGGCCATAAGGGCCATGA
TGACTTGACG
TCATCCCCACCTTCCTCCGTTTTGTCAACGGCAGTCCCTCTAGAGTGCCCGGCCGAACCG
CTGGCAACTA
GAGATAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGA
CAGCCATGCA
GCACCTGTCTCCCGGCTCCCGAAGGCACCTTCGCCTCTCAGCAAAGTTCCGGGGATGTCA
AGGCCAGGTA
AGGTTCTTCGCGTTGCATCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTC
AATTCATTTG
AGTTTCAACCTTGCGGCCGTACTCCCCAGGCGGAGAACTTAACGCGTTAGCTGCAACACT
GAAAGGCTAT
CCCTCCCAACGTCTAGTTCTCATCGTTTAGGGCGTGGACTACCAGGGTATCTAATCCTGTT
TGCTCCCCA
CGCTTTCGCACCTCAGCGTCAGTAGTGGTCCAGGAGGCCGCTTCGCCACTGGTGTTCCT
TCCGATATCT
ACGCATTTACACGCTACACCGGAAATTCCGCCTCCCTCTACCATACTCTAGCCAGGCAGTA
TCAAGTGCC
ATTCCCAGGTTGAGCCCCGGGGCTTTCACACCCAACCTTACCAAACCGCCTACGCGCCCTTTA
CGCCCAGTA
ATTCCGATTAACGCTCGCACCCCTCCGTATTACCGCGGCTGCTGGCACGGAGTTAGCCGGT
GCTTCTTCTG
TAGGTAACGTCAACACGATGTGCTATTACACATCGCTTTTCTTCCCCACTGAAAGGGCTTT
ACAACCCG
AAAGCCTTCTTCACCCACGCGGCATTGCTGGATCAGGGTTGCCCCATTGTCCAATATTCC
CCACTGCTG
CCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCTGATCGTCCTCTCAGACC
AGCTACCGAT
CATCGCCTTGGTAGGCCCTTACCCACCAACTAGCTAATCGGACGCAGGCTCATCCCCCA
GCGCGAGGCC
TTTCGGTCCCCCGCTTTCCTCCTTAGAGATTATGCGGTATTAGCCCGAGTTTCCCCGAGTT
ATCCCCCTC
TAGAGGCCAGATTCCACGCGTTACTCACCCGTCCGCCACTCGCCAGCCACCTAAGCAAG
CTTAGGTGCT
GCTGCCGTTTCTGACTTGCATGTGTTAAGCATGCCGCCAGCGTTCAATCTGAGCCATGATCAA
ACTCTTCAG
TTAA

>Rhodovulum_visakhapatnamense_JA181

GGCTCAGAACGAACGCTGGCGGCASGCCTAACCATGCAASTCGAGCGAACCCCTTCGGGG
TTAGCGGCGGACGGGTGASTA
ACGCGTGGGAACGTGCCCTTCTCTGCGGGATAGGCTCGGGAAACTGGGTTTAATACCGCA
TACGCCCTTCGGGGGAAAGA
TTTATCGGAGAAGGATCGGCCCGCGTCTGATTAGGTAGTTGGTGGGGTAATGGCCTACCA
AGCCTACGATCAGTAGCTGG

TTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCA
GCAGTGAGGAATCTTGGACA
ATGGGGGAAACCCTGATCCAGCCATGCCGCGTGAGCGATGAAGGCCTTAGGGTTGTAAAG
CTCTTTCAGTCGTGAAGATA
ATGACGGTAGCGACAGAAGAAGCCCCGGCTAAYTCCGTGCCAGCAGCCGCGGTAATACG
GAGGGGGCTAGCGTTGTTCCG
AATTACTGGGCGTAAAGCGCGCGTAGGCGGACTATTAAGTCGGGGGTGAAATCCCGGGG
CTCAACCCCGGAACCTGCCTCC
GATACTGGTAGTCTAGAGTTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATT
CGTAGATATTCGGAGGAACA
CCAGTGCGGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGA
GCAAACAGGATTAGATAACCT
GGTAGTCCACGCCGTAAACGGATGAATGCCAGTCGTCGGCAAGCATGCTTGTCGGTGACA
CACCTAACGGAYTAAGCRYT
CCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAACTGACGGGGGGCCCGCACA
AGCCGGTGGAGCATGTGGTTT
AATTCGAAGCAACGCGCAGAACCCTTACCAACCCTTGACATCCTGATCGCGGTTTTCCGAGA
GGAATTCCTTCAGTTCGGC
TTGGATCAGTTGACAGGTGCTTGCATGGCTTGTCTGTCAGCTCGTGTCGTGAGATGTTCCGT
TAAGTCCGGCAACGAGCGC
AACCCACACTCTTAGTTGCCAGCATTGAGTTTGGGCACTCTAAGAGAACTGCCGATGATAA
GTCGGAGGAAGGTGTGGAT
GACGTCAAGTCCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGCAGTGAC
AATGGGTAAATCCCAAAAAA
CTGTCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGTCGGAATCGCTAGTAATC
GCGTAACAGCATGACGCGG
TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTTTACCC
GAAGACGGTGCGCCAACCCT
TACGGGGGGCAGGGGCCACGGTAKGTCAGCGRTGGGTGAG