

# FinalExamPythonforGenomicDataScience-AB

February 6, 2024

```
[1]: from Bio import SeqIO
```

```
[4]: fasta_sequences = SeqIO.parse(open(dna.example.fasta), 'fasta')
with open(output_file) as out_file:
    for fasta in fasta_sequences:
        name, sequence = fasta.id, str(fasta.seq)
        new_sequence = some_function(sequence)
        write_fasta(out_file)
```

```

      □
↳ -----

NameError                                Traceback (most recent call↳
↳ last)

<ipython-input-4-566c304df787> in <module>
----> 1 fasta_sequences = SeqIO.parse(open(dna.example.fasta), 'fasta')
      2 with open(output_file) as out_file:
      3     for fasta in fasta_sequences:
      4         name, sequence = fasta.id, str(fasta.seq)
      5         new_sequence = some_function(sequence)

NameError: name 'dna' is not defined
```

```
[5]: from pysam import FastFile
```

```
fasta = "dna.example.fasta"
# read FASTA file
sequences_object = FastFile(fasta)
```

```

      □
↳ -----
```

```
ModuleNotFoundError                                Traceback (most recent call_
↳last)
```

```
<ipython-input-5-58d82b588e80> in <module>
----> 1 from pysam import FastaFile
      2
      3 fasta = "dna.example.fasta"
      4 # read FASTA file
      5 sequences_object = FastaFile(fasta)
```

```
ModuleNotFoundError: No module named 'pysam'
```

```
[9]: sequence = open(r'dna.example.fasta', 'r')

a = sequence.read()

sequence.close()
```

```
[20]: sequence2 = open(r'dna2.fasta', 'r')

a2 = sequence2.read()

sequence2.close()
```

```
[21]: a2
```

```
[21]: '>gi|142022655|gb|EQ086233.1|91 marine metagenome JCVI_SCAF_1096627390048
genomic scaffold, whole genome shotgun sequence\nCTCGCGTTGCAGGCCGGCGTGTGCGCGCAACG
ACGTGTGGGGCCTGACGGGCAGGAGGATCTCGGCGGCG\nCCAACTATGCGGTCTTTTCGGCTCGAAAGCCAGTTCAGAG
CCTCCGACGGCGCGCTGACCGTGCCCGGCTC\nCGCATTTCAGTTCGCAAGCCTACGTCTGGGCTCGGCGGCGACTGGGGGA
CCGTGACGCTCGGGCGCCAGTTC\nGATTTTCGTGCGGATCTGATGCCGGCTTTTCGCGATCGGCGCGAACACGCCGGCCG
GCCTGCTCGCGTGGG\nGCTTGCCGGCGAATGCGTCGGCGGGCGGTGCGCTCGACAACCGCGTGTGGGGCGTCCAGGTGA
ACAATGC\nGGTGAAGTACGTGAGCCCGACGTTTCGGCGGATTGTCGTTTCGGCGGCCTGTGGGGCTTCGGCAACGTGCCC\
nGGCACGGTCGCGCGCAGCAGCGTGCAAAGCGCGATGCTGTCTACACGCAAGGCGCGTTCAGCGCCGCGC\nTCGCTTA
TTTCGGCCAGCACGATGTAACGCGGTGGCAATCTGCGCAATTTCTCGGGCGGTGCAGGCTA\nCAACGTCGGGCAGTT
CCGCGTCTTCGGCATGGTGTGCGACGTGCGGATCAGCGCCGCCGCGCGCTGCGG\nGCCACGACCTATGACGGCGGCTT
GACCTATGCGGTACGCGCGCGTTGCAGCTCGGCGGCGGCTTCAGT\nACCAGCAGCGCGGCGCGACATCGGCTCGGC
CAACCAGGTCACGTTGAGCGCCGACTATTCGCTGTGCGAA\nGCGTACCGGCCTTTACGTGGTATTCGCACGCGGGCACGA
CAGTGCGTATGGCGCGCAGGTCGAGGCGGCG\nCTCGGCGGGGCGGCGTCCGGCTCGACGCAGACCGCGGTCCGGCTCGG
GCTGCGGCATCAGTTCTGACGAT\nGCGCGAGAAACACGGGCTGCCGCGTACGCCGCGCGGAGCCCGTGTTCGCGC
GGATTTCAGAACCGAT\nGCATCATCCCGACGCGCAACGCCAGCTGGTTGCGGCCCCGACGACTGCCCGGCCGTGCCGAGCA
CGTGCGC\nGTTAGTCGAAATCGGTGCCGTATGGCCGTTGCATGCTGGTACGCGCCCTGGATGTACGTCGAGGTTTCGC\
nTTGCTCAGATCGTAGTCGAGCATCATCGACACCTGGTGCCAGTTCGGCGACGCATCGCCCGCCCGCGTCG\nnCGACATG
CGCATGCGTGTAGATATAGGCGGCGCCGAGCCAGAGGTCCGGCCGAAGAAATACTGGCCGTT\nGACCTCGAAGTTGTC
GAATTTCCACGCGTTCTGCGAGCCTGCCGCCGGCTGGTTCGCGAAGTACAGATTC\nGACTCGGGGTGGTTCAGTCGAC
```

GTGCGAATACGCGGCGGACAGCGTCAGCTTGTCTGCTGAACTTGTACG\nATGCGCCCGCATCGATGTTCTGCTGCGCGCG  
 GCGCTGAAGACCGTATCGTCGCTCGACAACGCACCGCC\nCGCGGTGCGCGCGCGCTTGTGGCCTTCAGGTACGCGAC  
 GGCGGCGGAGAACGCGCCCGCTGGTACTGC\nACGGCGGCGCTATAGACGCGGTTCCGCGCGAAGCCCCCGCTGGTT  
 CGCGAGCCCGTAGAGCACTTCGC\nCCTTGAAGCCGGCGTAGGAGGGCGACACGTACTTGACGCTGTTCTGGATGCGGTA  
 GTCCCAATCGGCGTT\nGTCTGTTGCTGACGGGTGCGCGCGGAAAGTCGCGGAGCCAGTTGCCGGTGGCGGTGAAGGTACT  
 CCACATG\nTCGAGCGTGGGGTCTGACTGGCGGCGCAACGTGAGCGTGCCGAACCGGTGCGAATCGAGGCCGACGTAGG\  
 nCTTGCTGCGCAACATCCGCGACGTACGCGGAGCGCGCGGTGCTCGCGTTGAAGCCGTTCTCGAGCTG\nnGAACAGC  
 GCTCTGATCCCGCGCGGAGATCCTCGGTGCCTTTAGCCCCCAGTTGCTGCGGACGGTATCG\nnCCGCTGACCATCTGG  
 TAAGCCTTGCTGCGCGCGCGTTGCTCGTGAAATTGACGCCTTCGTGATCAGGC\nnCGTACAGCGTCACGGTGCTTTGC  
 GCATGGGCGGTGACGGCGAATCCCATGGCGCTCGCGGCGGATACGGC\nnCGCCATCATCTTCTTCATCGTGCATCTCCAG  
 GTGTGGGACGCGCGCGCGGTGCGGTTTCCGACGG\nnCATACTGACGACCGGACGCGTGCAGCGAGTCCGTTTGT  
 CGTTAGTAAAACGAAATATGAATTACGAAGA\nnAAAGAGGCCCGCGCGATCGGGCGGACCGGATGACGAACCGCGCGCG  
 GCGCGCGGTTCCGCGTGGCGGCG\nnGCGTTACAGCATCCGTGCGAGCAGTCGTTTTCCGGGTTGCGGTGGAAGAACCGG  
 TGTTTCAGCGCGCCA\nnAGCACATGCAGCAGCATGAGCGGAGCAACGTATAGGCGAGCGTTTTGTGACGCCACTGAAAC  
 GCCTCGA\nnACCAATGGTCGTTCTTCGGCAGCAGCTCGGGTACGCGGAAAAAGAAGAACGGCAGCGCGTCTGCGT\  
 nATACGTGCTGGACATCGAATAGCCCATCAGCGGCACGATGATCGCGAGCGCATACAGCAGGTAATGGCCG\nnATCTTCG  
 CGAGCGTGCGATCGAGCGTCGGCAGCGTGGAGGGCAGCGCGGCAACGGCTTCATCGCGCGAA\nnTCGCCAGTTGCGTGA  
 CGACCAACAGCAGCGTGAGCACGCCGAACCTCTTGTCGTCGGGTAGTACCAGTC\nnGAACCTGGCCGCGAGTTGTCGT  
 CGAGCCGACCATCGTCCAGCCGGTCCACAGCTGCGCGCGGATCAGG\nnATCGCCCTGACCCAGTGCAGCAACCGAAGCG  
 CGAGCGGGTATTTCTCCGCGGGGAATGCAATGGCCTGGT\nnTGTCATTTGACTTGTCTTCTTTGAAACGACATTGCGA  
 CAGGCGGGATGCCGATCGTCGACTTCGGCGT\nnCCC CGCGGATCTACGTGCAACCGAGGCCGGATCGCGATACGCACGG  
 ACCTTCGACGAAAAGCCCGTGCG\nnTCGGGCTGTTTCGCGCGGTGAGCGGATCGCGCGCAATCCGCGCAGCACTCCTGC  
 AACGCCTCGCGCAGT\nnGTCACGGCCATGAAGTCGGCTTCGCCAGGCTGCAGCGTGAGCGCGCGGAGATCACCACTTG  
 TTCGCGA\nnTCGCCCCGACACGACGCCGCGCGGACGCGCGCATCGGCCACGCGGGCGGCGAACCCTTGCCCGGATC\  
 nGAGCGGCGTGCGGCTGGCCTTGTCGGTCACCAGATCGAGCGCCAGCATCAGCCCCCTTGCCGCGGACGTGCG\nnCCGACCG  
 TCTCGAAGTCGGCCTTCAGCGGAGCAATTGCTCGAGCAGGCGTGTCCGACCCGCGCGCGT\nnTGCCCGGAGATCCT  
 CGGCCTCGACGATGTGAGCACCGGAGCGACGCGGCGCACGCGAGCGGATGGCC\nnGCTGTACGTATAGCCGTGCATGA  
 CCACGTGCGAGAAGCCCGCGCATGTTGATCGCGTCGGCGATCCGC\nnCCGTTGTACAGCGTCGCGCGGAGCGGCACGT  
 AGCCGGCGGAGATGCCCTTGCGGACGCACAGGATGTGCG\nnCGGCGACGCCCCAGCCGCGGTGCCGAGCAGGCTGCCGG  
 TGCGGCCGAAGCCGGTGACCACTCGTCGGC\nnGATCAGCAGGATGCCGTTGCGGTGCGACACGGCCCGACGCGCGCCC  
 AGTAATCGGCCGGCGGACGATC\nnAGGCCGCGCGGCCCTGCACCGGTTGCGGACGAATGCGGCAATGGTCTGCGGGC  
 CGTGGAACGCGATCG\nnTTTCTCGAGCTGCCGATGCGTGTGCGAGCTGGTGCGGATCGTCGCAATTCCACGGAT  
 TGCGATA\nnGAGCCACGGCGTGTGAGCAGCACGCAGCCGGGACGAGCGGGCCATGATTGTAATGGTAGACCGCGTTG\  
 nCCGCCGATCGACGTGCCGCCATGTGCACGCCGTGATAGCCGTTGCGCAGCGACAGGAATTTCTGTCGCG\nnACGGCTC  
 GCGCCGCGGACCCAGTACTGGCGCGCATCTTGAGCGCGTCTCGATCGCGTCGGAGCCGCC\nnGCTGCCGAACATCAC  
 CCGTGCATGTCTTCTGCGGAACATCGCGCACAGGCGGTGCGCCAGATCGTAG\nnACGCGCGGATGGGCGACGCCGTC  
 GAAGGTCTGGTAGTACGACAGCTCGTCCATCTGCCGCGGATCGCGC\nnTTTTCACTTCGGCGCGGTTGTGGCCGACGTT  
 GACGTTCCACAGTCCGCCGACGCCGTCGAGCATGCGGTG\nnGCCGTCGACGTGTCACGTAGTTGCCGTCGCCGCGCTC  
 GATGATCACGGTGGGAGTCGTATGCCCGGCG\nnGCCCGCGAGCTCATCGGGTGCCAGAATCGTTTTTTGACGCTTTTGATA  
 GTCCATGGTGTCTCTGCTCTCTC\nnTTACGAAGTGCTTGAACAGGGGGCGGCGCGGTGTGCGCGCCCGGATCACAGGGCT  
 GCGGTGATGATCTTT\nnTCCTCGAGGTACGAATCGAGGGCGGCTTCGACAGGTCGCGCCCGATCCCGCTTTGCTTGGTG  
 CCGCCCC\nnACGGCAGGCGTGCGTCGATCGGGCTCCACGATTGATCGCGACCGCGCGGATGTGAGCCGCTCGGCGAC\  
 nGCGGTGCGCGGTGCC\nn>gi|142022655|gb|EQ086233.1|304 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCGCGATC  
 CAGTAGCGCTTGTAGCCGAGCGCTTCGGCACGCTTCGCGAGCGCGATCGTGAAATTCAGTGCG\nnTCGTGCGCGCTCGCC  
 CTTCTGCAATCGGGCTTTTGTGCGAGAATCGAAAGGAATAAGACATGAAGCTCA\nnACCTTTTAGAACGGCTGACCCGC  
 CGGCACGAAGCGGCGTCCGGCGGGTGCCGATACAGCGGAAAACACCG\nnGGCAGACCCGTTATTATCGATAACGGATCAA

CGTCCCGTTACGGTCATTTCTGCTATGGATATGGTCGTT\ncGATCAGAACAGATAGCGCGCGCGCAGATAATAAAATCC  
GCCGTTGACGCCGATCGTGCTGCCGCCGCTG\nTACTGAATACCTTCCAGTTGCGTCTCGAAAGGCAGCTTGCTCGGGCG  
AACGTGCAACAGGTTTTGCGCGC\nCCGCCGCGATCTGGAACGCTTGGTCACGTCTAGCGCACTTCCAGGTCGGTCAC  
GTACCGCGCGGCCTC\nCTCGTAGTGAACGAACTGCGACGTCGAGAATGCGTTTCGGGCCGACGAAGTACGTCCCTCGCC  
CGAGGTC\nTTCCCGTAGCGGTTTCATGCAGCGTGACCGCCCACTTGTGCGAGACGCCAGGTGCCGCCGATGATGATCT\  
nTGTTCTTCGGTGTCGTCTGGACAGCGACGCGATCTGCTGTCCGTTACGCGACGGCTTCCCGTTTCGATCC\nnGTTTCGCC  
ACGTGCGTGACCGACGTCGTATTAGGTTACGCCGAAATCCCAGTCCACCTGGCCGTAGCGA\nCCGAACCCGGTGTA  
TAGGTGCCCGCGAGGTCGATCCCGCGCTGCGCGTGTTCGCGCCGTTGGTGAAGT\nAACTTGCGCTGACCGCCGATGCC  
GGCACCGAGCCCGCACGCTCAGGCCCGCCCTTCGAGCGCGCGAT\nngGCCGCGCGCCCGACGAGTTCGCCAAGA  
ACGATCCGGTTGCGGATATTGATCTGGTACGCGTCGATC\nnGTCAGCTGCAGGTTGCGAACGGGATTGAGCACCAATCCG  
AAGTTGTAGCTGGTCGACTGCTCCGGCTTCA\nnGCGCCTGCGCGCCGATCAGGCGCGCGCCGGCGACGTCGTTGCCAGG  
ATGCCGCCACGGAGGCCGGCGA\nCGTCGTACGCTGGTATAGGACTCTGGGCC\nn>gi|142022655|gb|EQ086  
233.1|255 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole  
genome shotgun sequence\nCTCGACGCGCTCCGCGTCGAGGTCGCCCCGACGTCTCGCGCAGCAACTGATTCAAA  
AACAGGCCGCCGCTC\nATGCCGATCTTGCGGTGGATGCGCCACACCGACAGTTCGATGCCTTCGGCATCGAGCGCTTCC  
TTCCACG\nnCAAGCACGTGCTGGTAGACGCTGTGACGAGCGTGCCGTCGAGATCGAACAGAAAAGACGTTTCAATGCG\  
nCATGTGTATCTCCTGGCTCGAAAGGGGGCGAGCGAACGGGTCGTGAAGCGTGTCGCGACATTATCGGCGC\nnGCGGCGA  
TGTCATGACCATGTCCCGCGGCCCGCGACGCGACCCCGTGCCGCGCCGTGTATGTG\nCCGCTGGTACAATCG  
CGGCGATCGCCGGGCGGGCTCTCCGCGCCGCGCGCCCCAACCTCGTCTCGCC\nnGATTCCAGGTATGGCTACACCGG  
ACGCCGTACGTTCCAAGCACTCGTGGTGGGGTGTCTGGCCCTGGCA\nnCTCACCGCCTTCATCTTCAATACACCGAAT  
TCGTGCCCCGTGCGCTGCTCAGCGCGATCGGCGACAGCC\nnTGACATGCAGCCGACCGACGTCGGCCTGATGCTGACGA  
TCTACGCGTGGGCCGTGGCCGTGCTGCTT\nnGCCGCTGACGTGGCCACGCGCCACGTCGAGCGCCGAAGCTGCTGA  
CGGGGGCATTGCTGGTATTATC\nnGCGAGCCACGTGCTGACCGGTGTGCGTGGAATTTGCGGTGCTGATGGTCGGCC  
GGCTGGGCATCGCAT\nnGTGCGCATGCGGTGTTCTGGTCGATTTCCGTGCCGCTGGCCGTGCGGCTCGCGCCGAGCGACC  
GGAAAAG\nnCCGCGCGCTCAGCCTGCTGGCGATGGGCACGGCGATCGCGATGGTGGCCGGCATTCCGCTCGGGCGCGTG\  
nGTCGGCGAGAGGTTCCGCTGGCGCGTCACGTTCTGATCATTGCCGGCGCGGCCGGCTCGCGTTGCTGC\nnTGCTGCG  
CGCGAGTTGCCGATCGCCGAGCCAGGGCGCCGGTTCGCTCGGCAGCATCGGCGTGTTCCT\nnGCGCAAGCCCGCGCT  
GGTGGCGTTGTACGCGATACCGTGCTCGTGTGTCGCGCACTTCACGTCGTAC\nnACGTACATCGAGCCTTTCTGTCCA  
GAGCGTCAACCACGCGAGCAGCAGCCGATCACGTATGTGCTGATCC\nnTGTTTCGGCGTCGCCGGCATAACGGCCCGGAT  
CTGCTTCAACCGCATCTATCCGACCGGCCGACGACTT\nnCCTGCTCGGGTCGATCGTCGCGTGGCGGGATGCCTGCT  
GATCCTGTTCCCGTGCGCGCTGAACATCGTC\nnACGCTGTCCGTGCATACGCTGGTGTGGGGCGGGCGATCGTCTGCTT  
CGGGCTGGCGATGCAGGCATGGG\nnTGCTGAAGCTGGCGCCGAGGGGACCGACCTCGCGGTGTGATCTTCTCCGGGCT  
GTACAACTCGGGAT\nnCGGCGCGGGCGCGCTGATCGGCAACCATATCGCCGGCGACTTCGGGCTGCCGTGGATCGGCAC  
GTTCCGGC\nnGGCGTGGTCGGCGCGTGGCCGTGGGATCGCGTGGCTGGCGTGCACGCGAGGCAGGCGGGCG\  
nCCTAGCCAGCGCCGGCTGGCGCGTCCGGCATGGGGCGAGGTCGCCACGGCCCGCCTCGCGCTTCGCG\nnCTTCGCG  
GGTCATGCCACCGAGTAGCCGCCGTGCGCCACGAGCGCATGACCTGACACATAGCTGGAGTCA\nnTCCGATGCGAGGAAC  
GCGACGATCGTCGCATCTCCTCGGACAGCCCCAGCGTCCGGCCGCTGTCGATG\nnCGGCAAACGCTTGCTGCGCTTCT  
TCCGACGGCCAGATGCCGCGATGATGGAACGGCGTTTCGATCAGCCC\nnCGGGCACAGCGCTTACCCGCACGCCGCGC  
TTGAACCATTCGATCGAAGCGGTCTTCGTATCCCGATC\nnACCGCGTGTGCTCGCGATATAGACCGACGATTCTCG  
AAGCCGATCAGCCCCGCCATCGATGCGTTGT\nnTGATGATGCTGCCGAACCCTGCCGAGCATGATTTACGGGCGTAC  
TTCATCGAGTTGAATACACCGCG\nnCACGTTTCGGCTCGAACACCATGTGGAACCGCTCGGCATCCTGCTCGAGCAGCGGC  
GCGAACACGCCTTCC\nnGTGCCCCGATTGTTGAACGCGACGTCGAGGCGGCCGTAGGTCGATACCGTGAAATCGAACAGC  
TTGCGCA\nnGATCGTCTTCGTTGCGGACGTCCGCGACGAAGGCCTTCGCTTCGCCGCGCGGTCACGATTTCTGTCGAC\  
nAGTTGATCGAGCTCGGGCTTGCGCCGCGCGGAGACGACGACCTTTCGCGCGCGACGCGGAGTTCGATG\nnGCCGACG  
CGCGCCGATGCCGAGCTGGCGCCGTTACCAGGGCGATGCGGCCGTCGAGTTGGGCCGTT\nnTCGTGATTGCGTTCA  
TGATGAAGTTCTCCAGAGTGATGTTGAAAGCGACGCATGTGGCTGAGTCGAATC\nnGGTTGCGATGCACGCTTTACGAG  
CATTGTGGTCGCTCGCGCTCGTAAAAAATGGAACAATGGCGCATT\nnCAGAATTCGCTTTCGGGAATGAAGCGCGAA

GCCGCATGCGCGCGCCAGGAGGGTTCTCCATGCTCAATC\nGACTCGACATCCTGAAGATCTTTGCCGCGGCCGCGCGG  
 CGCCACGTTCCGCGAGGCGGCCGCGGCCT\nCGGCGTGTGCGCGCAGGTGGTGACACGCGCGGTGCGCGACCTGGAGG  
 AGATGCTCGGCGAAACGCTGTTT\nCACCGGACGACCCGAGCATCCGGATCACCGCGTTCGGGCAAGCCTTCGCGCGCG  
 ACGCGCAGACGGCGC\nTGGCTGCCGTGACGGGCTGTTCCGGCCCCGCGACCGGCCAGGCCGACGAGCCGGTGGGCGTCCG  
 TGCGGAT\nCACGGCGCGGTGCGGCATGGGGCGTCTTACGTCCAGCCGATCCTGGGCGATCTCATGCAGCGGTACCCG\  
 nGGGCTCGTGCCCGACTTGCGGCTGTCGGACGTGCCGTGCGCGGTGGTCGACGAGCAGATCGACATCGGCG\nTGCGCGT  
 CGGGGCCATCGGCGACAACCGCTTCGTGCGCGCATGGTCGGGCGGTGGCCGATGTGGATCGT\nCGGCGCGCGTCTCGT  
 GATCCGGCGGCTGGGTGAGCCCCAAAAACCGCAAGGACCTCGAATCGATGCCGGTG\nACGTGCCTGATCGATCGCGCGAG  
 CGGCCGCGCGTGGCCGTGGATGTTTCGGGGGGAGCAGCAGTTCGTCC\nCGTCTGCGCGCGTACCTGACGGACGATAC  
 CGAGGTGGAGATCGAGGCCATCTGTGCGGGGGTGGGGCT\nGGGGCAGTGCTCGGAGTACCTCGTCCGGCCGTACGTGCA  
 CAACGGGCGCCTCGTGCGGTGCTGCCCGGC\nCTCGAGCCCGAGCCGTGGAAGCTGCACGTCTATCGGCCGCGCGCGG  
 GCCGATGCCCCGGCGCATCCGGG\nTCGTGTACGACGAGCTGGTGGCGAAGCTCGCGGACGCGACGTGGCGGGGGTAGGC  
 GGTGCGGCCCCGGCCG\nCGGGTTCGCGTTGGCGCGAAGCGGTGTACGCCAGGCTCGACGACGTGGCGCGACCTGCGAACC  
 AGTAACG\nTGCGTTCGCGCACGGCTATCCGGCCTGGCGTGGTACGCGGCAAACCGACATATTGCGGCATCATATGGAC\  
 nGCTGACGCACCCGCATTTCGTGCCCGATATTGCACTGGAGACGTTACATGCCGATACCCCGATTGAAAC\nCCGCGAC  
 GCGATGGGCGACCCCGATGTGGTCTATCTCAACCCGGATCACGTGCTGTGGATGTACAGCCG\nACCGCTGCGCAGCGG  
 CGGCCGGGCGGCACGGCGATTTCGCGTGGACGACCGCGTGAAGGGCGGTTTCGTTGC\nTGATGGCCGACCATCCGCCGGCC  
 GCCAACTGCTGCAGGATCTCGGGCCGTTTCGTGACGGTGACGCTGGC\nGAACCCGTCGTCCGACTATCCGGACGGTCTC  
 GTCCATGTGCGTGCGCAGGCGATCGTCAAGATCGCGACG\nGACGACCACGACAAGCCGCTCGCCGAGCGACGCTCGGC  
 TGGGTCCACGTCAAGGACGGCTCCGCGTTCA\nAGGTTCAGCGATTACGCCGGCGTGGCCGCGCAATGGCAGGCGTCCGGT  
 GCGGGTGCCGGCCGCTGAACGGG\nGCTTCACGCCTGCGTAATGGCGCTGCGGCCCGTCTGTCCGCGAGTCCGCCGGCCGT  
 GGTTCCTCCGCGGCAA\nAACGGATGCCGTGCCTCGTGACCCGCGCGAACGGCTCAGGTGCCGCCTGGCCGGGCGAACCA  
 TGCGCCC\nACCCGGTGCCGCATTTCGATACGCATTGCACCGATCCCCCGAGTAACATGGCGCATCGCGACGAACCGGCG\  
 nCCCGTTCCGGGCCTTCGCCGGAACCGCGCCCGTCAAGCGCCCGATGCAGGTGTTCTCCACCGGCTTGCCG\nAGGAACC  
 ATCCGTTCATGAACCTTTCCGCGTCCGATATCACACCGGAGGAACTCCGGCTGTTTCGGCGACAT\nCATCGCCCGCCTGGA  
 AGAACTGGGGCCGGTCTGTCGACGTGAGGACGGTCTGTGCTGCCCGCATCGCGCAA\nCTGCTCCGCGCCGATTTCCGCCG  
 GTCGTTTCGATTGCGACGAACACCCGGCCTGTGGCGCAACGGCTTCT\nCGTACAACATCGATCCGTTGAGATCGCCCG  
 GTACGAAGCGTGGTTCCAGCATATCGATCCGGTCACATC\nGCAACTGCGCGCGCGCCGCGTGGCGACCTGCGTGCAGCA  
 GGTGGTGAGCCGGTGCGAACGGAACCGGACC\nGAGTTCTACAACGACTTCCTGTGCCGCGACGGCATGCATCACGGCAT  
 CAACGTCTATGCGTTTCG\n>gi|142022655|gb|EQ086233.1|45 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCGTGCTC  
 GGCACGACTATCAGCCCGTATCTGTTTTTCTGGCAGGCCTCCCAGGAAGTCGAGGAGTTGCAC\nGCGGCGCGCGGTGAT  
 CATCCATTGCGCCGCGCCCCCTTTCAGGCGCCAGTTCAACTACGCCGATCAGTT\nTCGATACCTGGTTCGGCATGGGC  
 GTATCGAACGTAATCGCGTTCTTCATCGTGCTGACCGTGCCCGGAC\nCCTCCATGCGCATCACATCGTCTCAAGACA  
 TCGGCCGACGCGGCGCGTGCCTCGAACCGTTGGCGGGC\nCACTTCGCTACGTCTGTTGCACTCGGCATCATCGGC  
 ACCGGCTTGCTGGCGCTTCCGGTGTGCGCG\nGCTCGACCGCCTATGCGATGGCGGGGACCTTCAAATGGAGAAACAGT  
 CTCGCGTTACGCGCGACGGTCGC\nGGGAGAGTTCTATGCGGTTCATCGCTGCCGCGGTGCGGATTGGACTGGTTCATGACG  
 TTCGTTTCATTTGAC\nCCGATTTCGAGCGCTCTACTGGAGCGCCGTATCAATGGGATCACGGCAGTGCCCATCATGGTG  
 GTGATGA\nTGCTGATGGCGCAGAGCCGGCGCGTGTGGGCGAGTTCGCAATCAGAGGACCGCTTGGCTGGGGAGGGTG\  
 nGCTCGCGACGCTCGCCATGGCGCTCGCGGCGGCGCGAATGCTGCTGCCGGGATGAGCCGGCAATCCGGAT\nGGGAGAAT  
 GCGCATGCCCCGCGACGCACCGGCGACGCCTCGCCGACGGCGGGCGTGCATTGCCATTTCGC\nCATTCGCCATTTCGC  
 ATTTCGCCATTTCGCCATTTCGCCATTTCGCCATTTCGCCGAGCGC\nTCCATCGACGACGGTGGCGGCCA  
 CGCCCCGGAATTTCGACATGCCTGCATCCTCCGATACGGCGAACCAGG\nGGGCGTCATCAATCGCGCGCATCCAGCGCGG  
 GCTGAAGCGCGGGCTCGGCCGCGCTGCCGGTTCATGGC\nCGCGTGGCGCGCGGCGGTGGAATGGCCGGGCCGGATCC  
 TGAACCAGATCGCATACATCGCGGGCAGGAA\nCACGAGCGTGAGGACCGTCCCGGCGAACGTGCCGCCGATCAGCGTGT  
 ACGCGAGCGTGCCCCAGAACC\nGAATGCGTGAGCGGAATGAACGCGAGCACGGCCGCATCGCGGTAAGAATCACCG  
 GGCGCGCCCGCTGCA\nCGGTTCGCTTCGACGACCGCGTGAACGGATCGAGTCCCGCGTGTTCGTTCTGGTGGATCTGGC

CGATCAG\nGATCAGCGTGTTCGCGATCAGGATCCCCGACAGCGCGATGAGGCCGACCAGCGCATTGATGCCGAACGGC\  
nTGCCCGAACAGGATCAGCGTCGGCACACGCGGATCAAGCCGAGCGGGCTCGTCATGAACACCATCACCA\nTCGCGGA  
GATCGAGCGCACCTGGAAGATGATGATCAGAGCGTGACCGCAGCATGATCGGGAAGAGCGG\nnCAGCATCGCCTTCGT  
GGCCTTGCCGGATTCTCGATGGAGCCGGCTTCCTCGATTGATAGCCGCTCGGC\nnAGCGCGCGACGATCGGCTGAAG  
CTGTTCTGTGATCGCGGCCGACAGTCCGGCGGGCTGCAGGCCGTGG\nnCGATGTGCGCCGCGACCGTGAGCGTGGCGAC  
GCGATCGCGCCAGCGCATGATCGGCTCTTCCATCCGCAC\nnGTGACCTTGCCGACCTGCGACAGCGGGATGCGCTGCC  
GTTGCGCGCGGCGAGCGTGAAGTCGTGAGC\nnCGCGCCGGTTCGAGCCGCGTATCGCCGCCGAGCGCGCATCACCTG  
TACGGTCCGGATATCTCGCGTA\nnCCGCGGTGACGGGACCCCGTTCAGCAGGAAGTGCAGTTGCTGCGCGACCGCGCT  
GGACGTCAACCCGAC\nnTGCTGACAGGCGGTCTGCTGCAACGTAAAGTGCAGCGTCGGCGCGCGCTGCCAGTCGGC  
ATTGACC\nnGTGCGCATCATCGGGTGGCGTGCATCACACGCTCGACGTCGGTTGCGATACCGCGCAGCTTGTCCGGAT\  
nCGGGGCGCGTATCCGGTAGGCCACCGGGAACGGCGAATAAGGGCGAACACGAGCTGCGTCACGCGCAC\nnTTGCGCT  
TCGGGCGCGAGGCCGTGCGCGATGGCGCGCCGACCGCGCCTTCAGCGCTTCGCGCTCGTCC\nnTGCTGTCCGTGCGG  
ATCAGGATCTTCGGAACGACGATCGGGCAGCTCCGGCCCCATCGCGAGATAGA\nnAGCGCGCGCGCCCTGGCCGACG  
TACGCGGTGACGATCTCGCTTCCTTCTGCTTCGCGAGCCACGCTC\nnGACCTTCTGCGTGGCGCGCTCGTTTTCGTA  
ATCGACGTGCCGTACGGCATCTGCACCTCGATCAGTACC\nnTCGGGCGATCGGAGATCGGAAAACTGCTTCTTCAG  
ATCGCCATGCCGAGAATCGCGAGCACGAAGA\nnGGGCGACGACCGAACCGGCGACGAGCCATTTGCGGGCGATCACGCGC  
GTCAGCAGTTGCCGAAGCGTT\nnGTAGCGGGCGTGTCTGATGATCGATCGTGGCCGCTTCGAATTTCTGAGGTT  
GGCAGCATCTTCAG\nnCCGAGGTACGGCGTGAACGCGACGGCGACCCACGATGCGATCAGCGCGATTCGACGATC  
CAGAAAC\nnTGTTGCTCGTGTATTGCGCGCGGTGGAGCGCGGAAGCCGTTGCGCATGAAGCCGACCGCGGTGACAG\  
nCGTGCCGGCCAGCATCGGCGCGCCGTATGGCTCCACGCGTAGGTGAAGCAGCCACGCGGTGATAGCCT\nnTCCTCCA  
TCTTCAGACCATCATCTCGATCGCGATGATCGCGTGTGACAGCAGCCCGAGCGCGAGGA\nnTCAGCGAGCCGAGCG  
TGATGCGGTGCAAGTTCTTGCCGTCGCGCCATCACGACGAACACGATCGCCAG\nnCGTCAGCGGCACGGCCGCGCGA  
CGACGAGCCCCACGCGCCAGCCATGCTGACGAAGCTCACGAGCATC\nnACGACGAGCAGCGCGCGGAAGAACTTCACCA  
TGAAGTCTGTCAGCGCCGAGCTGATGTTACCGACTGGT\nnCGTGACCTTGGTCAGGCTCATGCCAGCGGCAGCGCCG  
CGTTGATCGCGCCGACTTCGCGGTCCAGCGC\nnGCGGCCGAGCTCGAGCCCGTTCCAGCCGTCGCGCATCACGACCCGA  
GCAGCAGCGCGGTTTCGCCGTTG\nnTTGTGGATCAGGAACGTGGACGGATCTTCGTAGCCGCGCTCGACGGTCCGCGATGT  
CGGACAGTTTCAGCG\nnTGCGGCCCTGCGAGACGATCGGCGTGTGCGGAATCTTCTGACGCGTGTGAAACGCGCCGTCCA  
CACGAAT\nnGAAGATCTCCGGCCCGCGGTCTCGACGGAGCCGGCCGCGTCAGCAGGTTCTGGCCGTTGAGCGCGGCG\  
nAACACGTCTGCGGGCTCACGCCGAGCGTCGCGAGCCGGTCTGCGACAGCTGCACGTAGATGCGCTCCG\nnCTGCTC  
GCCG\n>gi|142022655|gb|EQ086233.1|396 marine metagenome JCVI\_SCAF\_1096627390048  
genomic scaffold, whole genome shotgun sequence\nnTGATCGGCGACGGCTACTGGTTCTCGAGA  
AAGCGCGCAACTACGCGAGCGAGCGCAAGGTGTTTCGGCC\nnGGCCGATCGGCCAGAACCAGGGTATCCAGTTCCCGCTCG  
CCGAATCGTTTCATCGAGCTGGAGGCCGCGAA\nnCTGATGCGCTGGCGCGCTGCGAGAAGATCGACGCGCGCCAGAATG  
CCGGCGTCGAGGCCAATGCGC\nnAAGTACCTGGCCGCGAAAGCGAGCTGGGAAGCGCCAAATGCTTGCTGCGAGCGC  
ACGGCGGCTTCGGTT\nnTCGCTGCGAATACGACGTGGAGCGCAAGTTCCGCGAAACGCGCCTGTACCAGGTGCGGCCGA  
TTTCCAC\nnCAACATGATCTCGGGCATGTGGCCGAGCACGTGCTGACGTCCCTCGTTCTACTGATTCCGGAAGGAT\  
nTCCTCTCCATGAAGATGCTCGTACCCGATCGACCAGACCGCGAGATCGTCGCGCCGCAACAGTACATCG\nnCGGCCGG  
CATCTCGGGGACGATCCAGCACTTGCCGGCATGAAGGTTCAACGGCAGTCGTCTCGATCGA\nnCAAGCATTCCGAAGC  
ACTGATCTTCAGCGTGGCCGACTACAGCCTCGAAACGACCTGCTCCGGCCGTAC\nnCCGACCAGGTCAACGCACGCTGA  
CGGCAAGGGCAGCCGGCACCTGAAATACGGCGTGCATTGCCTGCCG\nnGACACTTCCATCACCTAGTACGAAGCACCAT  
GAATGTCGAGCTTTCCGAATCTTCGACGCGTGGGTTCG\nnCCGACGGAAGGCCATACGACCGGATTACTCGGGCGCC  
CGTCCGTCTGTTGACGGCAACTCTGGACTAC\nnGCGACCCCATCCGAATCCCGCGACGCTGCCGCCACTGTGGCACTG  
GCTGTAATCTGCGCGGCGAGC\nnGACAGTCGAACATCGGCGTCGACGGTCACCCGACGCGCGCGGCTTCTGCGGCC  
CGTCAGGCTGCCGCG\nnCCGCATGTGGGCCGGCGGACGGCTGCAGTTCTTCGCGCGCTGGCCGTGATGCACCCGTGCA  
GCGCCG\nnTCCACGATCTCAGCGTGCAGAGTAAATCCGGCCGACGCGCGGCTGGTGTTCGTACCGTGCTGCACG\  
nAAAACGAGCGACGCCGAGGGCGTGGCCATCCGCGAAGAACAGGACATCGTGTATCGCGATGCGCCGCCGCC\nnTGCCACT  
GCCGGCGCGCCGACACCTGAGCCAGCACCGGCGCCGACCGACGAACAGTTTTTCGCGCATCGTC\nnACGCCCGATCCGGTG

CTGCTGATGCGCTTTTCCGCGCTCACCTTCAACGGCCACCGGATCCACTACGACC\nGGCCCTACGCGATGCAGGAAGAA  
 GGCTATCCGGGCTCGTGGTGCACGGCCCGCTGATCGCGATGCTGCT\nnGCTGGAGGAAGTGCGCCGACGCACTCCGGC  
 AAGACCGTTTCGCCGCTTCGAGTTCAAGGCTGTGAGCCCG\nnCTGTTGACACCGCGCCGTTACGGTGAACGGCAAGCTC  
 GAAGGACATGCCGCCCGCTGTGGGCGCGG\nnGCCCGCAAGGGCGGCTGGCCATGCAGGCCAGCGCCGAACTGGAGTAA  
 CGTTCGGACCATGCAAGCTCGCA\nnAAGCTCAACCCGGCCCGCTGAGCGGGATGACCGTCGTCGCGCTCGAACACGCGAT  
 TGCCGCGCCGTTCTG\nnCACACGCCAGCTCGCCGACCTCGGCGCCCGTGTGTCATCAAGGTGGAACGCCCTGGCGTGGGCGA  
 TTTCGCG\nnCGGGCGTACGATCACCGCACCCGCGGCTCGCGTCGCACTTCGTGTGGACCAACCGGTGGAAGGAAAGCC\  
 nTCACGCTGGACCTGAAGCAACCGGCCGCGCAGGCGGTGCTGGGCGAACTCGTGGGCCAGGCCGACGTGCT\nnGGTGCAG  
 AACCTCGCGCCCCGCGCGGCCGCGCGCTGGGCCTGTGCTTCGATGCGCTGCACGCGAAGCAC\nnCCGCGGCTGATCGTC  
 TGCGACATCTCGGGCTATGGCGCCGACGGCCCGTACCGCGACAAGAAAGCCTACG\nnACCTGCTGATCCAGAGCGAGGCG  
 GCATTCCTGTGATCACGGGCACGGCAGACGAGCCCTGCAAGGCCGG\nnCAACTCGATCGCCGACATCGCGCGGGCATG  
 TACGCGTACACCGGCATCCTCAGCGCGCTCCTGCAGCGC\nnGGCGTACGGGCATCGGCTCGCACGTGGAAATTTGATG  
 CTGGAAGCGCTGGCCGAGTGGATGAGCTTCC\nnCGATGTACTACGCGTACGACGGCCAGCAGCAGCCGCCCGCAACGGC  
 GCCGCGCACGCGACCATCTACCC\nnGTACGGCCCGTTCAAGGCCGCGACGGCCGCGTCTGTGATGCTCGGACTGCAGAAC  
 GAGCGCGAATGGA\nnGCGTTCTGCGATGTGCTGCTGCGGACCCGGCGCTCGCGACCGACCCGCGCTTCGACGCCAAC  
 GTGCGC\nnGTTCCGAACATCGTGCGAAGTGAAGGCCGTGATCGAACAGGCATTCTCCTCGCTACCGCGCAGGACGT\  
 nGATCGCGCGCTCGACGCCGCGCAGATCGCGAACGCCCGCGTGAACAGATCGGCGATCTCTGGACGAT\nnCCGCAAC  
 TGCACGCGCGCCAGCGCTTTTCGACGATCGACTCGCCGGCCGGCGAACTGCGGGCGTGTGCTG\nnCGCCCGCAACGATCG  
 ACAGCTTCGACGCGCGCATGGATGCAGTGCCCGCGTGGCGGAGCACAGCGCGGC\nnCATCCTGCACGAACTCGGGCGCA  
 CCGACGCCGATATCGAACAACTGCGCGCCGCCGGCGTGTGATCTGAGCC\nnATGTCCAAACCCCTCTCCCCCGTCCATCACG  
 CGCGCAGCCTGTGTTCTGTGCCGGCCACTCGCCCCGAGC\nnGCTTTGCGAAGGCGTTGACTCGGGCGCGGACTGCATCA  
 TCGTCGACCTCGAGGACGCGGTGAGCCCGA\nnCGCAAGGACGACGACGCGCGCAACTCGCGCAGACCTGCCGTTGC  
 TGACGGCCGAGCAGCGTGCAGC\nnACCGTGTGCGCGTGAACGCGGTGCGCACGCCGTGGCACGCGGACATCGCGC  
 TGCTGCGCGACCGGG\nnCCGCGCAACCGGTAGTCGTGATGCTGCCGAAGTCAGAAGATGCGGGCGCGTGCGCCAGGTGCG  
 CGGAGCA\nnGCTCGGCGCGCACGCGCAACTCGTGGCGCTGATTGAATCGCTGGCGGGGCTCGATGCGGCCGATGCGCTG\  
 nGCGCGGACCCGCGAGTCTGTGCGTGTGCGGTTTCGGGCATCTGGATTTCAGCTCGACCTCGGCATGCGCG\nnCGACGCC  
 CGACGAGCCGAGCTGGCCTTTGCGCGCATGCGATCGTGGCCCGTTCGCGTTCGCGCACAGTT\nnGCCCCGCGCGATCGA  
 CGCGGTGACCACGCGCACCGACGATGCCGAGCGTCTGACGGCCGATGCCCGACGC\nnGCGCGCGCATTCGGTTTCGGCGG  
 CAAGCTGTGCATCCACCCGGCGCAGGTGCGCCGGCTGAACGATGCGA\nnTGGGCTATTTCGAAGACGAGCAGGCGTGGGC  
 GCGCCGCGTGTGACGCGCCCGCGAAACACGGCGGCGC\nnCGCGTTACGCTGGACGGCCGAATGGTCGATCTCCCGGT  
 GATTCGTGCGGCCGAAGCGATCCTCGCCGGT\nnACGCGGAAATGATCCGTGCTACGACGCGGCAAGATCGGCGAATTGCG  
 ACAGCCGGCCTCGCGCGGCTGCG\nnGCCGAACGCGCTGCCGGGAAGCCCGGCTCGCCGTTGAACGGCGACACGACCGTCC  
 CGCCGAAGGCGGAGT\nnCGCCGCGACGGTGCCTAAGAGGCTCGCCCCCGGTCAGACGCCTTCGACCAGCACGGCCCGAA  
 AGCGTCC\nnACCTTTGACGCGATGCTGAACAATGGCCTGGTATGCCGGGCTGTGCTACACGCGCGCGGGCGTCCGCG\  
 nGACGGAACTCGACGATCACGACCCCTTCGGGGCCGTCACCTCGAGCGCTTCCTGCGGGCCATACGCGG\nnCCAGCAC  
 CTTCAACGGGTGGCCGGCGATCGTCGCGCCGACCTTGTCCTGGTAGAGATCGAGTTCCTGCTG\nnATCCTGCGTGTTC  
 A\nn>gi|142022655|gb|EQ086233.1|250 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\nGCCCGTGCCGACGCGCCGATCGTCAGCGCG  
 CGGCGCTCGTGGTCTGCCACGCGGGCACGGTCATCGCG\nnTCGCCGTGCTCGTGAACCAGCCGATCTCGGCCATGCC  
 GGCGCCGCGTTCGCGGTGCCCCGACGGATAGC\nnCCGGCGTGGACATCACCGGATACATGCGCCGACGCGGCAAGCCGC  
 GACACGAAGCGCATCAGCGTTAC\nnCGCTTCGTGCTGCGCGCCAGATCCAGTCGAGCCATGACAGCTCGTTGTCTTG  
 CAGTACGCGTTGTTG\nnTTGCCGTGCTGCGTGCAGCCGAACCTCGTCGCGGCCACAGCATCGGCGTCCGAGCGCGGTG  
 AACAGCG\nnTCGCGAGCATCGAGCGCGGACGCGTGCAGCACGTCACGGATCGCCGCGTCTGCTCGGCCCCTCGAC\  
 nGCCCCAGTTTCGCGTGCAGTTGTGCTGCGGGCCGTGCGGTTGTCTCGCCGTTGGCTTCGTTGTGTTTG\nnGTGCAAT  
 ACGAAGCGAGATCGGCGAGCGTGAAGCCGTGCTGCGCGGTGATGAAATTGACCGACGCCAGG\nnTATGCCGGCGTGGT  
 GGTGAAACAGGTGCGCCGAGCCCGCAAGGCGCGCGGAGCTCCGGCCGCTGGCC\nnCGGGTTCGCCGCGCAGAAGCGCC  
 GCACGGTGTGCGGAAAGCGGTGCTTCCACTCGGCGAAACCGGGCGGG\nnTGGCGGCCGAGCTGATAGCCGCCGGGGCCGA

GATCCACGGCTCGGTGATCAGCTTGCCTGCGCGAGCA\nCCGGGTCTGCGCAACGCGTCGAAAAACCCGCGCCGG  
 GATCGAAACCGTGCTCCTCGCGCCGAGCGT\nCACGCCGAGGTGGAAGCGGAAGCCGTGATGTTGAACGCGGTGCCCC  
 AGTAGCGCAGCGAATCCATCACC\nATCTGCACGACGCGCGGATGCGACAGGTTGAGCGTGTGCCGAGCCCGTTTCGT  
 CGACATGAAAGCGGC\nGGTTGTCCGGCATCAGCCGGTAGTAGCTCGCGTTGTGAGGCCGCGCCACGACAGCGTCGGGC  
 CGAGCTC\nGGTTGCCCTCGCACGTATGGTTGTAGACGACGTCGAGCAGACCTCGATGCCGGCCGATGACAGCTGGCGG\  
 nATCGCGATGCGCATCTCGTCGAGCCGCGCGCTCGCGAGATACGCGGGCTCCGGCGGAAGAACGCGGCCG\nTGTCGTA  
 GCCCCAGTAGTTGCGCAGCCCGGGTTCACGAGCGCGCTGCTGCAGGAACGCGTGGACGGG\nCAGCAGCTCGACCGT  
 CGTCACGCCGATCGACAGCAGGTGATCGATGAACGCCGGATGCGCGAACGACGCG\nAACGTGCCTCGCTCGGGCGGGCG  
 CAACCCGGCAGCCGCATCGATGCGCCGCGCACGTGGGTCTCGTAGA\nTCACCGTATTGCGCCACGACAGCGCGGGCG  
 CCGGTGCGTGCTCCAGTCGAACGCTTCGTGACACCAC\nGCACTTCGGCATCGCGGGCGCCGAGTCGCGTCGGTCCAT  
 CGACAGGTCCGCGCGGTTGAAATGCAGGCGA\nTAGCCGAACAGCGCTCCGACCAGCGGAACGGCCGACAGCTTGCG  
 CGCATACGGGTGAGCAGCAGCT\nTGGTCGGGTTGAAGCGATGGCCGTGCTGCGGCTGGTACGGGCCGTCCGCGCGAAA  
 CCGGTACACGGTGCC\nCGGATGCGCGTTGGGCAGGTAGCCGTGCCAGATCTCGTCGGTGATTGCGGAGGTCCAGGCG  
 CGCAAGC\nTCCTTGCGGCCGGTCGGATCGAACACGCACAGCTGGATGCGCTGCGCGTGCAGGAAAAACCCGCGAAGT\  
 nTGGTCCCAGGCGGTCAGGTGCGCGCGAGCGGATAGCTGCGGCCGATTGAGACGGGTGGGCAGGGC\nGGTCGGC  
 ATGGGGCGGATCTCCGTGTGAAAGCACGATTATGAGGCGGGCGTCAGCCACAGCGTCGCGA\nGCGCGCGCAGGCGCA  
 AGGTGCGCGACACGCTTCGCCGTGGGCGGGCACGGCTTCGGCCACACGGCGCC\nGTGCTTGCCGGCGTTGGTGCCGC  
 CGTACGGCGCGGCATCGGTGTTTCATCAGTTGCGGCCAGTGGCCGGG\nGCGGGCAGCCCGACGCGATAACCCGCGCGCG  
 GCACGGGCGTGAAGTTGACAGCCGCGACGACGAGATGCC\nCCGCGTCGTGCGGCGCGCGAACGCGAACACGCTGTTGT  
 CGCGATCGTCGCGATCAGCCAGGAGAAGCC\nGGCCGCGTGGCAGTCGAGCGCATGCAGCGCCGTTCCGCCGCATAGG  
 TGCGGTTGAGGTGCGGCACGAGC\nCGCTGCACGCCGCGATGCGCAGGTGCGTCGAGCAGGTCCAGTGCGGCGTCGCGT  
 CGTGCGCGAACTCGG\nCCCATTGCGCGAATTGCTGCCCATGAACAGCAGCTTCTTGCCCGGTGGGCCACATGAAGC  
 CGAAGTA\nCGCGCGCAGTGTGCGGAGCCGCTGCCATGCGTCGCCGGGCATCTTCGCGACGAGCGAGCCCTTGCCGTGC\  
 nACGACTTCGTGCTGCGACAGCGGCAGCACGAAACGCTCGGAGAACCGGTACACGAGCCCGAACGTATCC\nGGTCGTG  
 GTGATAGCGGCGGTGGATCGGGTCTTCGTGCAGATACGCGAGCGTGTGCTGCATCCAGCCCAT\nGTTCCACTTGAAGTC  
 GAAGCCGAGCCCGCGTCGCCGTCGGCGCGGTGACGCCCGGCCACGCGGTGCGAT\nTCCTCCGCGACGGTGACGACGCC  
 CGCCGGCGCGGCCGACCGTGCAACGTGTGTTTCAGCATGCGCAGGA\nATGCGACCGATTGAGGTTCTCGCGGCCGCC  
 GTGCACGTTCCGGCACCCATTGCGCTTCGCGACGCGA\n>gi|142022655|gb|EQ086233.1|322 marine  
 metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
 ce\nTGCGGGCGCCTGTGGGCGGTGCGCGCCGGGCTGTGCGAGCCGTGTTCCAGGCGGCGCGCTGTGGCT\nnCACC  
 GCCGCGCAGCGCGATGCGTACGACGCGAGCCGTGCTGCGAGTACAAGGCGACGGTGCTGTCCGCAT\nTCGGCGACGTTG  
 CCGATTGCGTTGCCGCGCTGGAACGACGCCGAGGCGCTGGCCGCGAACGAACTCGC\nGGCGCGCGCGGCCGGAAG  
 GCTTCGACGAAACGGTGCAGCGCCGCGGCTCGGTGCCGTGCCGATGACC\nGCGCAACGGGCGAGCGAGCGGCAGTACC  
 TGAATGCGGAGCTCGATTGCTGCGCGCCGCGAGCCGCCGGA\nTGACCGATACGGCGGCACTCCTGAGGCGATGGGCG  
 AATTGCCGCCCGACCCGCGACGGCAGTCGCGGC\nCCGGAATGACGGCTGCTTGCG\n>gi|142022655|gb|EQ0  
 86233.1|88 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole  
 genome shotgun sequence\nGGACCGGCGAGGAAGTGGTCGGCGTCGAACAGGACGATCGCTGCGTCACGTTGCA  
 GGTGCGCGGTGCCGA\nCGGCCGCAATATGCGCTCGATGCCGACTACGTGGTGGCGGCCGACGGCGGCGCAGCCCGAT  
 GCGCGAC\nTGGCTCGACATCAAGCTGGCCGGCACGACGACCCGATCAAATGGGTGCTGGTGCAGTGAAGAACGCG\  
 nGGCTCGACACCGCTGCACGGCGCTGAACTGCGACCCGCGCCGCCGAACGTGTGCATCTATCTGCCGTT\nCAATTTT  
 CGGCGCTGGGAGTTTCTCGTATTCCCGCATGAAGACGCGGAGGCGATTGCCCGGCCGGAATCG\nATCCGCGCGCTGATC  
 GCGCCGTACGTGACGACGTCGAGCGCATCGAGATCGTGCAGCACGGACCTATA\nCGCACCATTCGCGTGGCCGAG  
 CGTTCTGTGGCGGGCCGCTGGCGCTCGTCGGGACGCGCCCACT\nGAGCCCGCGTGGATCGGCCAGGGCCTGAAT  
 GCGGGCTTGCGCGACGTCGGCAATCTCGCGTGGAACCTG\nGCAGGCGTGTGAACGGCACGCTCGATCCCGCGTGATC  
 CACACCTACGAATCGGAACGCCGCGACACG\nCGAAGGCGATGATCGATCTCGCGACACGTTGCGCGCATGCTGATG  
 CCGACCAACCGGCTCGTCGCGTT\nTCTGCGCGACCGTTCTCGGGCTGCGCGATATGCGCCCGCTGAAGGACTAC  
 GTGCTGCAGATGCGT\nTTCAAGCCGATGCCGAGCTATACGCGGGCGCTGCTGCTGCCGGCGCGCCGGGCGACGCGACG



GGCAGGA\nTGATCGCGCAGCCGACGTCGAAACGGCCGACGGCGTGCGCCGCAAGCTC\n>gi|142022655|gb|EQ086233.1|594 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGCCGCGGGCGCCGACGGAGCACGCGCATGATCCCGCTTTCCCATCCGTCCTTCCGACCGCAACCATGAT\nnCGGCGAAATCGACATCTTCGGCGTATTCTGTGCCGGCCCCGCTCGTGCTGATGCTGATCGCCTACCTGATC\nnAACATCGTCTGCGCGCCGTGCTCGAGCGCGTCGGCTTCTACCGCCTCGTCTGGCACCCTTCGATCTTCG\nnACCTCGGCATCTACGTATTCTGTGTTGGCCGCCGTCTGTCATCGTTTCCACCATCTCTGTGGCTACCTAACG\nnTGAAAAA AACCTGGCTCTCGGCAGGGCAGATCCTGCTCACCCTGATCGTCTGTCGTGGCTGCCATCGT\nnGCTGTGGCGCATCGT CAACTACTACATGTTCTCGCCGTGGACCCGCGACGGTCACGTGCGCGCCGACGTG\nnATCCAGGTGCGCGCCGACGTGGC GGGCCTCATCACGTCCGTGCAGGTGCGCGACAACCAGGAGGTCAAGC\nnCGGGCCAGGTGCTGTTCTGTGATCGACAGGC CCGCTACACGCTCGCGAACGGTTTGCCGAAGCCACGCT\nnCGCGCAGCGCCGCGACGCTCGCCAGGCGAAGCGCGA GTACGCCGCAACCTGCAGCTCGGCAACCTG\nnGTCTCGGAGCGAACAGGTGAGGAAAGCCGCACGCGCTCGAACAGGG CGAAGCCGCGCTCGCCGATGCGC\nnAGGTGTCTGCTCGACACCGCGAAGCTGAACCTGCAGCGCACGACGATCGTCAGTCC CGTAGACGGCTACCT\nnGAACGACCGCGCGCCGCGTGTGCGCGAATACGTGCCGGCCGCGCGCGTGTGTGCGGTCTG CGACCGC\nnAATTCTGTTCCGCGTCGACGGCTACTTCGAGGAAACCAAGCTGCGCGGCATCCACATCGGCCAGCCGGTCTG\nnACATCATCGTGATGGGTGAGCCGCGTGCCTGCGCGGCCACGTACAGAGCATCGTCTG\n>gi|142022655|gb|EQ086233.1|293 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nACTGCCGCGTTACGCGGAGAATGTTCTGTCTGGAACGCAATCCCTTCGATTACCG TAATGATTTGGCAA\nnCCTTGCCGAACTGCTCGAAATGGCGCGCATGGAGTCGACCACCTGGTTCACCACTTCGCCGC CGCGCAT\nnCGCAACCTCGCAGGCGCTGCCTGCGAGCGGTTTCGCTGCTGGCGTTCTCCGTATTTGACGCACCGTC\nnCGGGTCAACTGCTCCATGCTGGCCGCCGTTTCTCGAGTGAAGCCGCTGCTCTTCCGTGCGCTGCGACA\nnGATCCAG GTTGCCCTGCGCGATTTCCGCCGCCGACGTCGTGATCGATTCCGCGGTGCGATTGAATTTCCGA\nnCAGGATGCCGCTCAG TTGCCGCCGATCGATTGAGCGACGCCATCAGGCTGCCGCGATCGTTCCGGCGC\nnACCGGCACGGGTGCTTCGAGATT GCCTTCGGCAATACGCCGCCACGACCGCGCGTCCGGCAGGTTTAC\nnCGCCAGCGACCGCATGATTCCGCGCGTGAT GACGGTTCCAGCACGATCGACAGGATGACGGTGATGAC\nnGCCTGCTGCCGCGAGCGTCAACTTGGTCAACGCAGTCAA GCTCGCCACGTCCCCTGAACGCTGATCGAGC\nnAGCGACTGCTCGGCCTCGTTTCAGCTCGGCCGCCACGGCTCGATAGCG ATCCATGAAGTGTGCTGCGC\nnCCTGCTTGAAGTAGTCGATCAGCACATTGGCAGGCTGCGTGCCTGCATTGACGTC ACGCCGATGGCGAT\nnCAGCTTGTCTGCTGACCTCCGCCACCTTCCGGTGATGTGCGCGAGTGTCTCGAGACGACGCTG TTGCGCT\nnGCATTGTCCGAAGTGAGGCTGCGGACGAGATCGAACGACGTGGCGAACTGCTTGAGGCCGGCCTTGTAGG\nnGCTCCAGGAACCGGTCTATCCCCGAGGCCACGTAGCCGCGCACCCCTGTCTCCAGGTTGACCATGCTGGT\nnCAGCATG TCGTCTGTTGCTCGCAGGACCTGGTAACTATGAATGTTCCAGCCGTTGCGCTGATTGAGCTTC\nnAGAAAATTGACGATC GAGATCGCGCTGCCGATCAACGATATGAGTACGACGATACCGAATGCTACGCCCA\nnGCTTCTGCGGACTGTCTATGGAG CCCATGCTCTTCTCCGTGTATTGTGCTTTGTGCGCACGCAAGCAA\nnACTTACTGTAATGCACGATTCCCCGCGCTGA AGAGACATTTGCGAGGACGGTCGAATTTATCGGTGCGCT\nnTGCCCTGACACGCAGTCGCATCAGGTGACTATTCGTCTT CGCAAACGATCGCGGGCGATTATCTTTT\nnTCGGTTGCCGCAAAGAACATTGCGCATGGTGTGAATCCCTCATCGCC TCACGCAGGTGATGCGATGCGC\nnGGCCGCGAATACGACCGTTCTCGTCACCAGGAAAAGAGGAATCGGCAAGGAGAGC GCGGCCAGGTTCTCC\nnATGTCTGCGGTCTCGTACGGATTTCCATACACCCGTGTACGATCCGTTGACACACGGGTTG CAGCACT\nnGCCCCGAAACCTTGACGGAACGAGCGGACCAGATGGCATGCGTATTGCGTGCATTGCGCGTTCGATGCC\nnGGTCCAGGGCGTTCAATACCGGCTCCCGCAGCGTGATGGGCTCTCGACACAGATCGACAAAACAGTTATC\nnCCGGCGG ATTTGCAACGATGAACGACACCACGCGACAAGCGCCGACATGGCGGCCGATCCAGTTTTTGGC\nnGACCTTCTTTATTT GTTCGACGCCCCAACTGGACGAAGCGGTGGGATACAGGTCAAACCTCATCACGCAC\nnAAAATGGAGCCGCACGCGCTGGA CGGCTCGATGCTCAAACGAATCCGGCATTACTACGCGCGCAACGAG\nnGCACATTACCGATTACGACGAGCAGTTTCAT GCGCTGGCAAGCAGAGGCCACGACCGGAGCAGCGAGA\nnAATGCTTGTCTGCGTCAAGTGTACGCCGATGAACCTCTC CGGGCTGTTCAATTCGATCATTCGCGCAATC\nnGATGCGCTGTCTCGAGATGCGCTCGACAGCAGGAGACTGGATGCGTC AACCTCCGCCCCGACCTCGCATCG\nnCGGCATCGCCATGACGCCGTACATCGGGGCCATGCCGACGGTGCAGAACGCTT CGTCCGTACGTACG\nnGAATGAGACCCGTGCGCGCAATCTTCGACGCGATCGTGCCGCGGCAAGCCGCATCGTCCGGCG CGCAATA\nnGTGCGATCGAGATCACCGGTATTGCGTAGCGGCTTCAATCGCCTTTGCTTGGCGGAGCAGCCAGTCTCGA\nnTCCGCGGCAGGGACGGCGCCGTAATGCCGGTTGCGCTGATCCAGTTGCCGAACAGCGATTGCAATGCAA\nnCCGCGGC

GACCGCGTTCGTGCACCTTCGGCAAGATCTCGAAGCCGCGGTTTCAGGATCAGTTGCACAGTCGG\ nGTAGCGGATTTCAC  
 GGCGCGGATCACCGCGACCAGGCCTGCCTCCTGCTTCGCGCGTGCCGCTTCCGTC\ nTTCGCGACGAGCTGGTACGAATC  
 CAGCGTATCGAAAAAGAAGCCGCGTAACCGTGCGCCACAGCGCG\ nCGATCACCTGTTTCGACGAAAAACGACGGCCA  
 CCCCGGCGCGGACTGATCGACGACCGCGGACGCCATGT\ nCGTATTACGCCCCGAACAACCACTGTTTCGGCATCCTGCG  
 CGCATAGTCACGCTCTTGCGTGACTTCGCGG\ nACGCTCACATAGGCGAACACGCGAGGGCGTCGCACGTCTTGCGCAGC  
 TGGATCGAAGCCGCTGTGCGGTT\ nCGACGACCGCGATATCGAACTTTTCAGCTGGCGAACCGGCACGTCTTGCCGTA  
 GAAAAACGCGACGGA\ nGTGCGGACGCATCGCGTCGGCCGATCGCACGCGTCAGGCGACGATACGGGCTGTGCGGGCTG  
 CGCATAA\ nCCTGCCCCGCCGCCAGCGCGCGGATGATGACGAATGCGATGACAGCCACGGACGATGCGGTGCGCCATCA\  
 nTCGACGGCCTGCGCGACGGACGAATGCGGAGTGAAAACATATCCCCTCAATGACGCAACATATAGTTTC\ nGTATTTCG  
 AGGCGGCCGAACCTTGCGGTTCGAGCAGCAGCACCGCCGCGACGTGCGAATGAGCAGTGCAAGA\ nGCAAAGCCGTAGCCG  
 TATGCATCCGGCGTCATCCATAGCGTGATACCGGTGACACGCCGTTGAGCGCGA\ nCGAAGGCCCGGACAGCGCCATG  
 ACCAGCCGGCGCGCGTCGAGATAGAAGAATACGTTTCAGCAAACCCAA\ nCAGCAGCACCTGCAGACTCGCGGCGATCACG  
 TCGACGAGCAGAAGCGGAAGGTACAGCGTCGAGATATGC\ nAGGGCCGCCAGCAGCCAGTGACCGAATGCGACGATCACG  
 ACCAGCACGGCAGCCTGCACCTTGATGATTT\ nCATACAGCCCCGCGCGCACGCTGCCGACCATCATGTGCGCATGTGCG  
 TTGATGTGCCGACGCTCGCACC\ nGCTGCGCACGGCTTCGTAACACGCATCGTAATACTCGACGAAGTCCGCCTCGATG  
 CGCACAGAAACGTC\ nGCCATCCCGGGCATCACGAAACGTACGCGATGAACACGGGAATGTCTGATGATCACCGATGCA  
 TGCAACG\ nGCCCCGATCACCTGCGGCCCCGTGCCCCGCGCATACCAGAACATGAACTTGTGAGCCATACGCCGAGATT\  
 nGAACAGCAGTCCGACGAGCGCCAGCGTCGGATACGCGAAGCGCCGGTCGAGCATGTGCAACGAGACGAAG\ nCGATCGC  
 TGCGGTAGTTTCGGTAGATCAACCCGAAAGCCCGGTAAGCAGGACGAGATGCCCGGCGACGA\ nAACCGCCGAGCAGGC  
 CGCAGATGCCCTGCCGTTTCAGCATCAGCGGAACGCGACCGTGACGCCGTATCC\ nGACCAGGAATACCGCAAGAATCT  
 GTCGATACTGCTTCACGCTGGACAAAAATATCACCGCGATCCAGATA\ nTTGCTGACGATGACGAATCCCGCGACCATCA  
 GCAGCCGATACAGCAGCGGCTCGCCGCGAAAAACCCGCC\ nCCACCGCGATGAAGCCGACGCATCCCGACGCCACCGTGG  
 CAATCAGCGCCACACCGTTATAGTTTCGACAG\ nCACGAGATCGTCGCGCTTCTCGAACAGCCGGTCGGAGATGAAGCGCG  
 TGAACGACAATTGCAGCGGCCG\ nGTCAAAATCAGGCTGCACGCGATCAGGTAAGTCACCGACACCTGAAACTGGACGA  
 TCGCATATTTTCGGTA\ nTCACGAACGCGACGCTCATCACCGGATGATCAGAATTCCGAATATAGAAAGATCAGCGGAC  
 CGGAA\ n>gi|142022655|gb|EQ086233.1|75 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\ nTATAGGCCTGGACCTGCGTGCGGGCGAGATC  
 GCGTCTTCGCCGAAGGTGCCGTTGATGCGTGCCAGCG\ nCGGTTTCGGTTGCCAGATCGGCTTGCGGTGAAAGCGCCAC  
 GGTGATGCCGACGGCGCGATACTCCTGGCGC\ nGCGATGTGCGCGAATTTTCGCGTGAGCGCCGTATCGCCGATCGCGGC  
 GAAGCCGAGCGTTTTCCGGCCACT\ nGCGAGAAACCGCTGCTGCCCGCGCTTGCGCAAGCGTGTACTGGAAATGATTGCG  
 CGGATCGGAGCTGAT\ nCGACACCGGGATGCCGAGCCGCGACTGCTCCGCCAGCGTCTGAATCTGGTTTCGCTGGTGGC  
 CATGGTG\ nCGCGCATCGGCACTCATGCGCGTGATATAGGTGTTGACGCCCCGCTGATTGATGAGCGTCTTGAGCGCCG\  
 nCCAGGTTCGTACGCGGAGCCGGTGCCCGCACCGGTCTGTGCTTCAGCACCGGCGCGGTGCCGTGCATCAT\ nCAGGCCG  
 GCTTTCTCGGCCAGCGTGAGCGGCCGACCAGGTCCCGCGCACGCACTTCGGCACTCAGCCGC\ nCAGTCTTCGTACGGT  
 TCGATCTGACCGCTCTTGTCCATGTCTTGAACCTTCAGGCCGTCCCGGGTGATCA\ nGGTTCAACGTGGTGTAGCCGAAG  
 TCCTGCTGCGTGACAGGCGGGGAGAGGAGTCGACGCCGCCGAGG\ nAAACAGCCCCAGCGTGAGCGCGGTGCCGGAT  
 GCGGCCCGTGCCGATCGGCGAATGGCTGATCTCGTGATG\ nCGTTGCATTGCTTGTCTCCTGAATCGGATTTTTGGCGCG  
 GCCGTGATGCATGACGAGGGCCGCCGCCA\ nCTTTACCGAGCTGTTTCAGGGGACGTTTTGCCGCCGGGTGCAGTTTT  
 TGCCCCGATGGGGCGCGATGTGT\ nCGGGCCATGCGGTGCGGCGGATTTTCAACGCGGTCCAGCGGCTTCGGGAGCCGGC  
 TTGCCGCTCGTCCGC\ nCCGCGCACGCTACGCATGAACGCCACCCGATACCGGGTGGTCTGTCTATAGACGGCGTCCGG  
 ATAAAGG\ nATCACGTCGCCGCTCAGGCCGGGCATGTTGACCTGGTCCGAAAGCCGCCGCCCTCCAGGCACAGCGCCG\  
 nCGTGTTCGCGTGAACGTGTGCCCGGCTCACGCGCAGGCCATCGAGATAGTTGCCGGTATAGAGTTGCAG\ nGCCGCGC  
 TGATCGGTGACACGACAGCTCGCGGCCGCTTTCCGGATCGTAGACGCGTGCGACGGGCCG\ nGCCCTCGTGCGCGCG  
 TCGCCAGCACGAAGCAATGATCGAAGCCGCGTGACGAGCCAGTTGCGCGTGCG\ nGCCAGTCGAGGCGCGCGCCGAG\  
 n>gi|142022655|gb|EQ086233.1|454 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\ nTCGCGGGCCTCGCGCTGCCGTTTCGCGGCGCT  
 CGCGTTGCTGGAGGTGGTCTGCGGCAGGGCTGGCTGCC\ nGAACCATCTCGTCCCGCGCCGAGCGAAATTCTCGACAC

GCTCGCCCGCCTGGGCATCACGCGCGTGGCG\nCGTCATGTGCGCGCGAGTACGCTGCGGGTCGCGCGGGGTTTGGCGC  
GGGCGCCGGCTTGGCGCTCGTCG\nTCGGCGCCGCGATGGGCCTGAGCCGCGGATCGACGCGCTGTTCAACCGACGTT  
CCAGGCGCTGCGCGC\nGATTCCATCGCTCGCGTGGGTGCCGGTGTGCTGCTGTGGCTCGGCATCGACGAAGCACCGAA  
GATCACG\nnCTGATCGCCATCGGCGCCTTCTTCCCGGTGCACCTCGCGGTGGTCGCGGCATTGCGGACGTGACCGCA\  
nAGTCGTGCAACTCGGCGCGGTGTACCGGCTCGGCCCCGTTGCGGTGTTCCGTGCGATCCTGCTGCCGGC\nnCGCGCTG  
CCGAGATCGTCACCGGCTTGGCGACCGGTCTCAGTCTCGCGTGGATGTTTCATGGTCGCGCGC\nnGAGCTGATCGCGGCC  
ACGCGCGGGCTCGGCTTCCGTGTGAGCGACGGCCGCGAAACCGGCCCGCGGATC\nnTCGTGTTCCGGGCGATCCTGCTG  
CTGGCGTTGCTCGGCAAGCTGACCGACGGCGCGATGGCGCGCATCGA\nnGTCGCACTGGCTCGGCTGGCGCGACGCCTAC  
GACCACACGCCACGCAAGGGTCGGCATGAGCACGCGTT\nnCTTCTGCACCCACGGCGCGAGCACCGCGACACCGCTGC  
TCGACCTGCGCATCACCGCAAGCTTTACGG\nnCGATCGCACGATCCTCGCGACGTCCCGCTGCAGGTTGCGCGCGGCG  
AGATCGTCTGCGTGGTCGGCCCG\nnAGCGGTTGCGGCAAGAGCACGCTGCTGCGGATCGTCGCGGCCCTCGACACCGATT  
TCCGCGGCAGCGTGA\nnAGCTGGGCGGCGTGCAGCTCGACGGCCCCCTCCCGCGTGGGGCGTGATCTTCCAGGAGCCGC  
GACTGCT\nnGCCGTGGCTGTGATCGCGGACAACGTGCGGCTTTGCGGCCGGCCCCGCGCGGCGGGCGGAGCCGGCCGTC\  
nGCCCGGCTGCTCGACGAAGTCGGCTGGCCGGGTGCGCCGGCAGCTGCCGGCCACGTTGTGCGGCGGGA\nnTGCGCA  
GCGCGCGGCGATCGCGCGGGGTGTTGCGCGAGCCGACCTGCTGCTGCTGCTGACGAACCGTT\nnGAGCGGGTCGATGC  
GATCACCCGGATGCGGTTGACGTGCTGCTGCTGCTGATGCTGCTGAGCGGCACCGG\nnACGGCCGCGATCGTCGTCACGCA  
CGATCTCGACGAAGCGCTGTATCTCGCGATCGCGTGTGATGCTCG\nnCGCCCAACCGGGCCGATCGACGACGAGAT  
CCAGCTCGGGGTGCTCCGGCCGCGGATCGACGCGACCC\nnGGCGTGGCCGTGCTGCGCGCACGGCTGCTGGACGCATT  
CCAGCAGCTTCAGGATCGTGCCGACGAGACG\nnAAGCGCCGAACGCACGACCTTCAGGCCTCCGAAGCCTGACTTCGCGC  
CGGCCGATTACGACCGGCGTCG\nnCCCGGCCTCGACTCGGCACGCGTTTCAGCGCCGGCATGACGCTTCCGTGATCGGAT  
ACGGGTACGACAGGA\nnATCGCAGTCGCACCGGAACCACGATCGATCGGTGCGCGGAAAATCGCGCGCCCGCTCACCCCG  
TAGCGGC\nnGTGACAGGCTGCAACGTACCGGTGCACCGGCGCACTTCAACCGTGCGTATGCCGCAACTCGCGTCGATC\  
nGTCTGCGGCAGTCGCAGCAGGTTGAAGATCGGGCATAACGGCTCGACTGCCTGATGCAGTGCACGAATCG\nnATTCCGG  
CGGTTCCGGGCGAGACGATATCGAGGGTGTAGCGGAGGTTATGCGGGAACACCGGCACATGCTC\nnGAAACCGGTTGCTG  
CGCGAGTGGATGCTGGTCGCCGGTCACTTCGACCTCGATCGAGTCGATTTTCAGG\nnCCGCGCTCCGCGCCTGAATCAG  
CGTGATGTGCGTCAGGCAACTGCCAGCACGCCGAGCTGCAGTCCG\nnGCGATGCCGGCCCGAGGTTGTAGCCCGCGAA  
ATCGGGCGGGCTGTGCTGATCACCTGGAAATCGCGAAT\nnGCGGATCTCCCGCACGCCGCTGCGGCCAGCGCGCGCAC  
GGTCGCTTTTCAGCGGGGTGGCCTTGAGTTCC\nnGGATTGCTGCGCGCGGCTTCCTTTTTCGCGAGCCAGGCGTCAACGCTT  
TTGTTGAGATAGTCGTTCACTG\nnTCGTGATGTGCGGCCGATACGGTCGCAATGACTGGCCACGCACACGGCGGCCAAT  
CGGGAGGGAGCTTCA\nnTGCTATTCCCTTCCGCGACCGCCACTAAACGATCAATCCAGCAAAGCAATTCGAGTATGTGA  
TGTCGCA\nnTTGCGGCACGCGCGCGGGAATTTCACTGTCGTTTCATCCACCTCGAGTTATCCACAATTTCCGTGAT\  
nAACCTGTGGAGAACGTGCGGAGAAATTGCCCGTCCGCTTTACCGGCAAGGATCGCGCGACGCGGTGCT\nnGCCTCGC  
GGGCCCTGCACGCAGACAGCTCGCATCGGGTTCGGCGATCGTCAGATTGTGTGCTTACCGG\nnCGCCGGCGCATCGAT  
CGCATCGACGTGCGCAGTCACGAAGTCCGCACCGGAAACGTGCCGTGTCGACACC\nnGATGCACCGTGTGGGACACCCGT  
CCTGCCCGACGATCCCGCTCCGGCGCCGACATCGATGCATCAGGCGA\nnGCGCCTCGTCACGCAGACGGAATTGCGCAAC  
CGCCGCACTCAGCCGCGCGCCCTGTTGCTGAGCGACAG\nnCGCCGCGGCCGCGGCCTCCTGACACGCGGGCATTGCG  
CTGCATCGTCTGCTCCATCTGGATGACCGCG\nnGCGTTACCTGCTGATGCCCGACGCCTGCTCCTCGAACGCCTGGCC  
GGTTTCGCTCATGATCATGTTGA\nnCGCGCTCGACCGCCGCGACGAGTTCGGCCATCGCCGAGCCGGCCAGTGCAGCAG  
GCCGCTGCCGTCTTC\nnGACACGCTGCGCGGAATGGCCGATCAGTTCGCGGATCTCCTTCGCGCGGACGCGCAGCGCTG  
CGCGAGG\nnCCGCGCACTTCCGTGCGGACGACCGCGAAGCCGCGCCCTGCTGCGCGGCGCGCGCGGCTTCGACGGCAG\  
nCGTTCAACGCAAGGATGTTGGTCTGGAATGCGATGCTTTTCGATCACGCCGACGATCCCGGAGATCCGCGT\nnGGAACGTG  
AGCGAAATCGCCGACATCGTCTCGACCACGCGCTGCATCGCCTCGCCGCCGCGCGACGCGATG\nnCCGACGCGCCTTCG  
ACATAGGTGCGTGCATCGCGTGCCTGTCAGCGTTCTGGCGACGGCCCGCTCA\nnGTTGCTGACGCTGCGCCGCCCG  
TGTTGACGCGAGACGGCTGCTGCTCGGTGCGCGCCGACAGGTCGAC\nnGTTGCGGCTCGCAATCGTGCGGACGTGCGCG  
ACGATCGTCTCCGTGCTGCGCGCACCTGGTTACGGTG\nnACGACGAGGCCGTCTGTCATCCGCTTCAGCGATGCAGC  
AGATAGGCCATCTCGTTCTCGCCGGGCACGA\nnCGACCGTTCCGGTCAGGTCGCCTTTTCGAGATCTTCTCGAACTGGCCG  
ACCGCGAGCTTGATCGGCTCGAT\nnGATCGCCTTCGCGAGCACGCGCTGCGCCAAGAAGCCGACGACGAGTGCGAAGGCC

GCGACCGCCGCCATC\nCCCCACACGATCCGGTGGAACTGCACGCCGGCCGCGTCTAGCGCGCCTTTTCGCGCGCGACC  
TGGAACG\nnCCTCGAGCGCATCGATCGCCTCCTGATAGGCGGCGAACAGCGCGGGCGGCGCGGTACGCTGCGTGTCGAG\  
nGAAGCCGAAACCGTCGTCGGCGTCGAGTTGCGAGAGGGCCTTCAGGAACACGCCGTCGAGCAACGCCCGA\nnCGGCGAT  
TTTGAGCGTGTCGAAGAGCGCCTGCTCGTCGGCGTTGCGCGCGTGCAGGCCCGCATAGGCAT\nnCCAGTTCGTGCTTGC  
TCTGCTTGAGAAGCGTGCGGAGTTTCGCGATCTCTTCCTTGCCGAGTGGCCGGC\nnGCTGATGATCTGCGCGACGTGCG  
TCATGCGTTTCGCGCATCACGAGCATGCGCTCGGAGCTGGTCTTCAGG\nnTGCAGCAGCGACGCGGTGTGTCGCGGTACA  
TCGCTCGAGCGAACTGTTGCCCATATAGAGGGCCGCGA\nnTGCTTGCGCCGACCACCAGCACGAGCAGCACGGTATAGC  
CGGCGATGGTCGCGACGAGACCACCACGGAT\nnAGTGAGTTGTTTTCGCATGACGTTCCCTTGCGCCGGCCGGGCGCGGC  
CAGCGGCGGATGTGGATTCCGGA\nnTCACGAAAGACGCGCGCAGTGTGCAACACGGGGCGCATGGGCTGCGAACGGGTT  
AACGGCGGCGGGAAC\nnGGGGAGGTTTACTGATACGCGACGTGAATCTTGACGATTGCGACAGGGGCGTGGGGATGAAT  
GCGCCC\nnATTTTCGCCCCAG\nn>gi|142022655|gb|EQ086233.1|16 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnGTCGATC  
GACACGACGCTCGCGCAGCGCGACGGAAGGCCGCGTGAGCGCACGACGCGCGTCACACACCA\nnCGAGCACAACGAACA  
CGACCCCACTCTCACGAGCCGACCATGGCCGACCTTCGCTGCACCATCGCGGGC\nnATCACTTCGCCGAACCTTTCTG  
GCTGGCGTCCGCGCCGCGACCGACAAGGCCTACAACGTGAACCGCG\nnCGTTTCGAGGCGGGCTGGGCGGGGTGCTCTG  
GAAGACGCTCGGGCTCGATCCGCATGTCGTCAACGTCAG\nnTTCGCGCTATGGCGCGGTGCAGTGAACGGCCAGCGCAT  
CGCGGGGCTGAACAACATCGAGCTGATCACC\nnGACCGTCCGCTCGACGTGAACCTGAGAGAGATCGCGCAGGTGAAGCG  
CGACTGGCCGACCGCGCGCTGA\nnTCGTGTGCTGATGGTGCCGTGCAACGAGCGCGACTGGAAATGGATCCTGCCGCT  
CGTCGAGGATACGGG\nnCGCCGACGCGGTGAGCTGAACCTCGGTTGTCCGACGGGATGAGCGAGCGCGGATGGGCGC  
GGCGGTC\nnGGGCAGGTGCCGAATATGTGGAGATGGTCACGCGCTGGGTGAAGGAAGGCACGAAGCTGCCGTGCCTCG\  
nTGAAGCTCACGCCGAACATCAGCGACATCCGGATGGGGTCGCGCGCCGCGTACAAGGGCGGCGCGGACGG\nnCGTGTCG  
CTGATCAACACGATCAACTCGATCGTCGGGTGATCTCGACCATATGGCGCCGATGCCGACG\nnGTGACGGCAAGGGC  
ACGCACGGCGGCTATTGCGGCCCGCGGTCAAGCCGATCGCATTGAACATGGTCG\nnCGGAGATCGCACGTGACCCGAA  
ACGCCGAACCTGCCGATCTCGGGCATCGGCGGCATCTCGTCATGGCG\nnCGACGCGGCGGAGTTCATGGTGCTCGGCGCC  
GGCAGCGTGCAGGTGTGCACCGCCGCGATGCATTACGGA\nnTTCCGGATCGTGTCGGACCTGGCCGACGGATTGTGCAAC  
TGGATGGACGAGAAGGGCTACGCGACGCTCG\nnACGACATTCGCGGCCGCGCGGTGCCGAACGTGACCGACTGGAAATAC  
CTGAACCTGAAATACGACATCAA\nnGGCGCGTATCGACCAGGACCGCTGCATCCAGTGCGGGTTGTGCCATATCGCGTGC  
GAGGACACGTGCGAC\nnCAGGCGATCACCGCGACGAAGGACGGCGTGCGGCATTTCAAGTGGTGATTTCGGCGTGCGTC  
GGGTGCA\nnATCTTTGCATGCATGTGTGTCCGGTCGAGCAATGCATCACGATGGAGCGTGTGATTTCGGGCGACTACGC\  
nGAACTGGACCACGCATCCGAACAATCCGGCGAGCGCGGAGGCGGGGCGAGTGCAGGCGCGGCGGCACCC\nnGAGAAGC  
ACGCGAAGAAGGCTGCTTGACGGCGTCCGGCGATGCGGGCCATCTGCATCGCCGCTTTTCGT\nnTCCACCCGGGCGCGC  
ATCGAGTGATGCCGGCGTTGACGTTTTCTGTGGAGTGAGTCAGATGAATCACGCAG\nnCGAATCCCGCGATCCCGATCGC  
GCCGCGGCGCAGGGCGGCAGCCTGTACAACGACGATCTCGCGCCGAC\nnGACCGCGGCGCAGCGCACGTGGAAGTGGTAT  
CACTTCGCGGCGCTGTGGTTCGGGATGGTGATGAACATC\nnGCGTCGTACATGCTCGCGGCCGGGCTGATCCAGGAAGGC  
ATGTCGCGGTGGCAGGCGGTGACGACGGTGC\nnTGCTCGGCAACCTGATCGTGCTCGTGCCGATGCTGCTGATCGGCCAT  
GCGGGCGCGAAGCACGGGATTCC\nnGTACGCGGTGCTCGTGCGCGCTCGTTTCGGCACGCAGGGGGCGAAGCTGCCGGCG  
CTGTGCGCGCATC\nnGTGCGGTGCGGCTGGTACGGGATCCAGACCTGGCTCGGCGGCGAGCGCATCTATACGCTGCTG  
AACATCC\nnTGACCGGCAACGCGTGCATGGCGCCGCGTGCCTGTCATCGGCATCGGGTTCGGGCGAGTCGCATGCTT\  
nCCTCGTGTCTGGGCGCTGCAGCTCTACTTCATCTGGCATGGCACCGATTTCGATCCGCTGGCTCGAAAGC\nnTGGTTCGG  
CGCCGATCAAGGTGCTGATGTGCGTGGCGCTGGTGTGGTGGCAACGTGCAAGGCGGGCGGCT\nnTCGGCACGATGCTGT  
CGGCGCGCTCGCAGTTTGCCGAGGCGGCAAGAAAGCCGGGCTGTTCTGGGCGAC\nnCTTCTGGCCGGGCTGACCGCGA  
TGGTCGGCTTCTGGGCGACGCTCGCGCTGAACATCCCCGACTTCACG\nnCGCTTCGCGCATTCGACGCGACACGGTGA  
TCGGCCAGTCGATCGGGCTGCCGTTGCCGATGGCGCTGC\nnTGTCGGTGGTGTGGTGTGTCGTGACGTCGGCGACCGTGC  
TGATCTACGGCAACGCGATCTGGGATCCGAT\nnCGACCTGACGAGCCGGATGACGGGCATCGGCGTGGGCATCGCGCTCG  
TGATCCTCACGCTCGACACGATG\nnTGCTGCAACCTCGCCGCGAATCTCGTCGGCCCGGCGTACGACTTCTCGAGCCTGT  
GGCCGAAGGCGATCT\nnCGTACCGCACCGGCGGGATGATACCGCGACGCTCGCGATCGTGATGATGCCGTGGAAGATCC  
TCGCGAC\nnGACGGACGGCTACATCTTCACCTGGCTCGTCGGCTACTCGGCGCTGCTCGGGCCCGTGGCGGGGATCCTG\

nATGGTCGACTACTTCCTGATTTCGCGGCACGCGGCTCGACACGCGCGCTGTTTCGACGAGCGCGGCGGCT\nTCAGCTA  
CGCGCGCGGCTGGAACCCGCGCGCTGGCCGCGCTCGCGGTTCGCGGTGCTGCCGAACCTGCC\nCGGCTTCCTGCACAC  
GGCGTTTCCGGCGTCGTTTCCGAACGTGCCGGCGTTCTTCAACACGCTTTACACG\nTACGCGTGGTTCGTCCGCCTCGT  
GCTGGCGTCATGCGTGTACGGCACCTGGATGAAGTGGCGCGCCGGAC\nnAGCACGCGCAGATCGCGAGCGCCTGATTCCG  
CACCCGACAGTCAACGAGGAGGCAACCCCATGGCAATCC\nnTGATTTCGTGGCGGCACCGTGGTTCGATGCGGACCGTTCTT  
ACCGCGCGGACGTGCTCTGCGCAGCCCCGGA\nnGGACGGCGGCACGATCCTGCAGATCGCCGGGCGAGATCGATGCGCCGG  
CCGGCGCGACCGTCGTTCGATGCG\nnCACGACCAGTACGTGATGCCGGGCGGCATCGATCCGCATACGCACATGGAAGTGC  
CGTTCATGGGCACGA\nnCCGCGAGCGACGATTTCTACTCGGGTACGGCCGCGGGCTCGCGGGCGGCACGACGAGCATCA  
TCGACTT\nnCGTGATCCCGAGCCCGAAGCAGCCGCTGATGGACGCGTTCCATGCCTGGCGCGGCTGGGCGGAGAAGGCG\  
nGCGGCCGACTACGGCTTCCACGTGGCCGTGACGTGGTGGGACGAGAGTGTGCACCGCGACATGGGCACGC\nnTCGTGCG  
CGAACACGGCGTGTTCGAGCTTCAAGCACTTCATGGCGTACAAGAACCGGATCATGGCCGACGA\nnCGAGGTGCTCGTGAA  
CAGCTTCTCGCGTTCGCTCGAACTCGGCGCGTTGCCGACCGTGCATGCGGAGAAC\nnGGCGAGCTCGTGTTCCAGTTGCA  
GAAGGCGCTGCTCGCGCGCGGATGACGGGGCCGAGGCGCATCCGC\nnTGTCGCGGCCCGCGGAGTTCGAGGGTGAGGC  
GGCGAATCGTGCGATCCGCATTGCGCAGGTGCTCGGCGT\nnGCCGTGTATATCGTGATGTGTCCGCGAAGGACGCGGT  
CGATGCGATCACGAAGGCGCGCAGCGAAGGG\nnCTGCGCGTGTTCGGCGAGGTGCTGCCGGGCCATCTGGTGATCGACGA  
GGCGTCTATCGCGATCCGACT\nnGGACACGTGCGGCCGCGCACGTGATGAGCCCGCCGTTCCGCTCGGCCGAGCACCG  
CGAGGCGCTGTGGCG\nnCGGGCTGCAGGCAGGGCAGCTGCATACGACGGCAACCGACCACTGCGTGTCTGCGCGTCGCA  
GAAGGCG\nnATGGGCGCGGAGGATTCACGAAGATCCCGAACGGCTGCGGCGGTGTTCGAGGATCGCATGTCCGTGCTGT\  
nGGCATCACGGCGTGAATCATGGCCGCATCACGCCGAACGAGTTCGTGCGGATCACGTTCGACGAACGCCGC\nnGCAGATC  
TTCAACCTGTATCCGCGCAAGGGCGCCGTGCAGGTGGGCGCCGATGCCGACCTCGTTCGTGTGG\nnGACCCGGCCGCGACC  
AGGACGATCTCGGTGAAGACGCATCACAGCAGGTTCGATTTCACGTGTTTCGAGG\nnGGATGACCGTACAAGGCGTCGCA  
ACCCACACGCTCACGCGCGCGCGCTCGCGTGGGCCGACGGCGATCT\nnGCGTGCCGTGCGCGGCGCGGGCCGCTATCTG  
AAGCGCCCGCGGCAGCCAGCTACTACGAGGCCGCGCGG\nnATCGCGAACCGGTGCGCGAACCGCATCCGGTCGAGCGC  
GCCGGTTGAGCGTTGCGTATCGCGCGGGCGG\nnTGTCGGTTCGAACGACACGCCCCGCGCATGTTTGAGCGTGCGTTTAC  
GTTTCGTGCCGGCACCGCTCGTGC\nnCGCCTCTCCGCATCACGCCATCCTCTCAATATTTGGGATGAATTGAGCGCGATC  
GCGCCTTGCCGATCT\nnCCGGATACATAGAACAAGTACGCAAGTCGATGAAACACGCGATGTCGCGCAAATGCGACCAATT  
TTGTTTG\nnCGTTGTGACGTGCATGCCGGCGAGTAATATCCACCGACGGCGT\n>gi|142022655|gb|EQ086233  
.1|584 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
shotgun sequence\nCGCAGGTGGCAGCATTTAAGATCTCGCAAGGTGGTGTCTCCCCATCATTTGGTGGGTCAATTT  
ACGTGTGA\nnTTAGCTTCTCGTCCGCGCTACTGTGGAGCACAGCTGAGCGCCCTCCAGACAGCCTTCGTGCGAGCGCGA  
\nnAGTGATTATCGTCTGACGATGCGCCATGTGTGCGCCGACTAACACGATTAACCGGCCAAATGGGAGT\nnACAGCA  
GACGACCGCGTAACGCACACTCCCGAGGTGTAAGCGATCGATTGATCTTGCCAGCAATCGGGC\nnGAACGTCTCCTGGA  
ACCGACGGTTGAAACCTGCCACCAAGTCTCGGGCAAATCTTCGCGCTCGGCCATG\nnCTCCAGACTTCAGGCAAAATG  
GTGCAATGTGCGTACCAGCTTGACCAACTCGTCAAACGCGATATCTG\nnCGATGACATCTAGTCGCGCCTCCGGCGATA  
GATCATCGTTCGACACATTTTCATAGAGCGTCAGATA\nnTCGTGTTGCCACCGCCTCCAACGCCGCTCGCAAAAGGT  
CTTCTCTGGTCGGAAGTAGTATTGCAAGGTC\nnGCCAGGTGATTCCAACATCGCTCGCCACGCGCGCAAAGCGAAGC  
CGACGTTACCGTACTGCGCGAGGA\nnCATTGATCGCGGCTCCACGATTCGGGGACTCGTGTGTTCTGTTGCCGCGGC  
GCTTCGGCTCGTCGGC\nnCAACTCTTAGCCCTGCCAGTTCAATCGCCAGCTCCCTCATCTTTGACCTCTTGCCGGCAG  
CCGCCGGT\nnGTGCGCTTCGACCAGTCCAACCGCTCAACCATCACCATTCCCTCGAGCGATCGCAGTCAGTTTCATCGGTG  
\nnCTCGCCCTAGCCACCATAGCGAGCAACCACCTTGTCAAAATAATATGTCACGCGACTGATAATTACCAG\nnTCAATC  
GAACATAAAGCGATATGACCCTTTAAATATCATCCATATTTATCAGGC\n>gi|142022655|gb|EQ086233.1|4  
marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequence\nATGTGCCTGAAGACGTTTCAGGCGCGCAATATTGTGGCGAGAGAAAAATCTGTGCCGAATGCCTCCT  
\nCCTCTACGTATCTCGGGCAGTCGCACGGTATCCGACTTGTGCGACGGGCCCGTCCGTCCACCAATCGGGC\nnACTTCG  
ACGATCCCTCGCTCACACTGGTCTGGCCAGTCAGGCCTTAAGTTCCCCAGCGCTGTCTGTTAA\nnACAGGACGTTGCCT  
GTTTTTGGCGACCCCATTCCTATAACGCGGTAACGGAGATCGAAGTCATCATGTTG\nnTCCTCGATTCTGTACGCGATAC  
TCGCGACCTGCGTCGCCATCGTGTAACGGCGCTGGCCGTGACAGGCG\nnGCCTGGTCTATTACGTCGTCAAGCAACACA

ACGACGAGACAATCGACCAGAATCACAAGTCGATCCTCGC\nCGGGCACGCGCTGGCGATCAACGAATGGGTGGCCAGCA  
 GAGGCCGGGAAACCCAGGCACTCGCAGACACG\nnATCGCCATCGGCGAAGGGGATCCCCTGCCGGCGCTGAAGCTGCTCG  
 GGAAGTCGGGCGGCTTCGAGGTAT\nnTGACGCTCGGCCTGCCGACAAGACCGCGTTCTCGAACGTGCCGCTCGCCCCCG  
 GCTACGACCCGACGGC\nnACGCCCGTGGTACAAGCAGGCCGTCAGCGCCGGCAGCCTGGTCTGACGGCACTCTATCGTG  
 ATGCAACG\nnACCGGAAGCCGGCTGTTGCGTTCGCGGTGCCGTGGTGCGCGACGGCACGGTGAAGGGTGTGCTGGCCG  
 \nCCAGCCTGTTTCATGGAGAGCGTGAGCTCGATCGTGACCTCGGTACACCCGACACCCGCGAGCTTCGCATT\nnTCTGGT  
 GGACAAGAGCGGGCGCGTGATTGCGTCGGCCAAGACTGACCTGATCATGAAGCCGGCAACCGGC\nnCTTCCCCTGCACT  
 CGACGCCGACCACCTCGGCGCGCTGCAGTCAGCGTCCAAGCCCGTCGCTGTCGACA\nnTCGACGGCGCCACCAAGCTGGT  
 GCGGGCCAGCCGATCGCCGGTACCGATTGGTCCCTCGTGCTCGCGCT\nnCGACAAGGCGGACGTGACGGCCGGCATGCG  
 CGCGTTCGCGACGACGACGCTGGCGCGCATCCTGATCGTG\nnGCCGTGATCGCGCCGCGCTCGTCGGTTCGATGACGAA  
 CGCAGCCTTCAAGCGCATCCTGCTGGTCCGCG\nnACGCGCTGACGGATGTGTGACCGGCAGCGCGACCTGACGAAACG  
 CCTGCTGCCGACGGCGCGGACGA\nnAGCCGCGCAGATCGCGCATGCGTTCAATCTGTTCCCGAAAAAGATCAGTACGAT  
 CCTGCTGCAGATTTCG\nnGAGTTCAGCGAATCCGTCAATCTCGCGAGCGCCGAGATCGCGCAGGGCAACCAGGATCTGTC  
 CGGCCGGA\nnCCGAGCATGCCGATCGAGCTTGAGGAAACGGCCCGCGCTCGAGGAAATCGCCGACGGCCCGCAA  
 \nCTCGGCCGATGCCGCGAGCCAGGTGAGCCGTCTGGCGAATCGCGCTCGGACGTGCGGTGCGCGCGCGC\nnACGGTC  
 GTCAGCGAAGTCGTGCGACGATGAACGAGATCACGCAGGCATCGAGCGAGATCGCCAACATCG\nnTCGGCGTGATCGAC  
 GGCATCGCGTTCCAGACCAACATTCTCGCGTTGAACGCTGCCGTGAGGCCGCACG\nnTGCGAACGAGACGGCCGGGGC  
 TTTGCGGTCTGTCGGGGCGAGGTGCGTGCGCTCGCGCAGCGCAGCGCG\nnCAGGCCGCCAAGGAAATCAAGCAGTTGATT  
 CACTCGTCGGTCGAGAAGATCGAGGGGGGCTCGGGACTCG\nnTCCAGACGGCCGGTACGACGATGAACGAAATCGTCGAG  
 GGCGTGCGCGCATCACGAGCGTGATTGCGGA\nnAATTACCGTCGCGCCCAAGGAGCAGAGCACGGGGATCGAGCAGGTG  
 AATCAGGCTGTGTGCGAACTCGAC\nnAATGCGACGCAGCAGAACGCGCGCTCGTCGAGCAGTCGGCGCGGCCGCGACA  
 TTGTGCGCGAGCAGG\nnCCGCCAGGCTCGCGCAGACGGTCGGCGAGTTCAAGCTCGAGGACCGCCGCGATGACGTTG  
 CAG\nn>gi|142022655|gb|EQ086233.1|277 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\nnCCGCGAAGGCCGCGTTCCGCCACGCCCCGCTGC  
 CAACAGCGATCTCGCCGGCACCACGTTGCGTGTCGCAAC\nnCTACAAGGGTGGCTGGCGCGCGCTGCTGCAGGCGGCCGG  
 GCTGGCCGACACACCGTACCGGATCGACTGG\nnCGCGAGCTGAACAACGGCGTGCTGCATATCGAGGCGCTCAACGCGGA  
 TGCGCTCGACATCGGTTCCGGAA\nnGCGAAATTCGGGCCGTATTCGCCGCGCGGCAGAAAGGCCAACGTGCGGCTCATCAC  
 CCGCGTGCGCGAAGA\nnCCGTGAACAACAGGTACGCTCGCGCGCAAGGACACGCCGATCCGCAGCATTGCAGACCTGAA  
 GGGCAAG\nnCGTGTCGGGTACGTGCGTGCGACGACGTGCGACTACTTCTGTATCGCCAGCTCACGGAGCCGGCCCTGA  
 \nGCTTCGACGACATCCAGCCGATCAACCTGTGCGCAAGCGACGGCTTGTCCGCTACGATCGCGCGACAT\nnCGACGCG  
 TGGGCGATCTACGGCTACAACGGCCAGCTCGCGCGCAATCGCTACGGTGACGTGTGTTGAAG\nnACGGGGAAGGGCTAT  
 CTGTGCGGGCAACTTCCCGGTCTATGCGAACCCGCGCACGCTCGACGATGCGCGCC\nnGTGCGCGGCAACGGGCGATCTG  
 CTGTGCGATTCCGGCGCGCGTATGCATGGGCGAACGGGCATTTCCG\nnCGACTATGCGCTCGCGCAGAACAGCGAGACG  
 CGCGTGCCGGTCGCCGATCTCGTCGCGATGTTGATCAG\nnCGCAGCGAAGACTACGCGCTGCTGCCGTGACGCCGAC  
 GTGGTGGCGCAGCATCAGCAGGTGCGCGATG\nnTGTTGCGCGGATCGGTGTGCTGGACGGGCCGGCCAATGTTGCGCCG  
 CTGTGGGATACGTGTTCAATGC\nnGCTGGTTCGCGCGACGTGATCGAGCGTAACGAGAAGGCGCCCGAGTGAGTGGGGG  
 ATGCGGATGAACACG\nnTGCTCTCGAAGGCGCTGTTACCGCGCGCGAACCCGGCATCCCGCCTGCAGTGAGTCTCGTCG  
 CGAGATC\nnCGGTGACCGGATCAGGCGAACGTGTCGCGACCATGATGCCGAACACGACGCCGATCGCAATTCGCGAG  
 \nAGCCGGCCGTAAATCGATGCGCGCACATTGTGCTGAGCACGAACGGCGAGCGTTTGTGCTTCCATACCG\nnTGCTGAA  
 TGCGACGGGGCGCGTCAGCAGGGTGACGCACGCGATCAGCACCGGCGTGCGGATGAGCATCGC\nnGATGCGCGGCCCGTT  
 CCAGTCGACGCCGTGCATCGCCAGCGAAACGATCGCGGCAAGCGGAACCTCGACG\nnCACAAGCCAGCAACGAACCGTT  
 GACGAATGAATCTCCGGGGGACTGGAACATGATCGGTCCTGGCATT\nnAAATGAAATCGGCGCCCGGCTCGAGCTCCCG  
 A\nn>gi|142022655|gb|EQ086233.1|346 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\nnCGGGTCCCGGCGCCGAGGCCGTCCGGCTCC  
 TGCAGCGCGCGAACC GGCTCTCGGGTGATTGCCAGC\nnGTACCGAGATGCGCCCGGCTGGGCCGTGATGGCGCAGT  
 GCGGCC\nn>gi|142022655|gb|EQ086233.1|527 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnGAGAACC

GGGAACCGGAACCATGACAGCCCCGCGCGGTTTTACGCGAGATAGCCGGAACGCCGTCCCA\nGAGCAGTTTCAATGC  
GGTCACCGCCAGCAATCCGTAGCAGCTCCGGTAGATCAGGCGCTGGTCCAGCCTG\nCCGTGAAGCCGCCAGCCGAACAC  
CACGCCGGCCGGAATGGCAAGCAGGCACACCGCCATCAACGCCCAGA\nCGTTTCGCGGTGGCTGCACGATCAGCAGCCA  
CGGCACTGCCTTGATCGCATTGCCCACGGTGAAGAACAG\nGCTCGTCTGTTCCCGCGTACATCTCCTTGCTGAGGCCAAG  
CGGCAGCAGATACATCGCGAGCGGCGGCCCG\nCCCCGAGTGCGCGACCATCGTCTGTGACGCCCGATGCAAGGCCGGCCGA  
GACTGCCTTCGGCGACGAACGCG\nGACGAACCGTTCGGCTCCGCCCCGCCCTCACCCACAGCCCGACGAAGACCAGCGT  
GACCACCGCCATCAA\nAAGCTCGATGGCGCGATGGTTCGAGGAAGCGGAAAGCCAGGTAACCGAACCCGATACCGACCAC  
CAGCCCC\nGGCAGGAGCAGCAGAGGTGCGGCTTCGACCATGTGACGGCTTCCAGTACCGCAGCGCGAACAGGTCCA\  
nTCGCGATGAACAGCGGGCGAGCAAGCCGCCGGCGTACGCGGTCCATCAGAGGGACAGCAGCGGAAT\nGCCGATG  
ATCGCGAATCCACCACCGAACGCGCCGCGCATGAACGCGATCACGAACAGCCGGCAAACGCG\nATCAGGATCGTGGCC  
AGCGTCAATTGCAGGCCATCGCAGCAGGGTTCGCCATCACGACCTCCATGCCGG\nTTTGAATCGCGCGTGGCGGACA  
GCCACGGAGCGGGTTCGACGCGCGGCATCGCCGCACGATGGATCCGG\nGTTGAACGCGTTGCACCCATGCTGCTTCTCC  
AATGAGGTACCGGGCGATGCGGTACACCAACGCACCGC\nAGGCCGATGGGCCGACAAGCATTTCAGCCCCGGTACA  
ATCGACTTGACGAAAGCAGAATGCACGCGC\nTCTATCTCAGTGAATTAACATTGACCTCGGTGCAATATTCATTG  
TTATCGGTGCAATCCATGTCAA\nTTCCGAATACCTGCAGTTGGCCGACGCGATCGCCGCCCAAATTGCCGACGGCAG  
CTCAGGCCGGGCGAC\nCGCCTCGCTCCGCGAGCGTCATTTCCGCCACGAGCATGCGATCGCCGCATCGACGGCGGGACGG  
GTTTACG\nCGGAACTGTTACGGCGCGCCTTGTGGTTCGGCGAAGTCCGCCGAGGCACTTTCTGTGCGGTGAGACGCG\  
nACGCGGGGCGCTGCGCCGGGCGAGCCGCGCGGCTTCGGATCGATTTTCAGTTTAACTACCCGACCGTC\nCCGGCCC  
AGACCGCGTTGATCACCAGAAGCCTGCGCGGATTGCACCGACCTGCGGAGCTCGACGCGCGT\nTACGCGAGGCGACGA  
GTACCGGGACCCCGTCATCCGAAGCGTTGCCGCCGCGTATCTGGCGCAGCATGA\nATGGGCCCATCGCCCCACCAGC  
TCGTGTTTACCGGAACGGGCGGCAAAGCATCGCCGCGCGCGTTGCC\nGCGGTCTGTGCCGACCGGCGGACGCTGTGGCG  
TGGAGGCATTGACCTACCCGTTTCATCAAGGGCATCGCG\nCCAAGCTGGGGATCTCGCTGGTGGCACTGGCGATGGACG  
ACGACGGCGTGGCGCCGACGCCGTGCAAAA\nGGCCACCGCGAGGCCAGGCTGTGCGGATCTACGTTACGCCAGCCA  
TCCAGAATCCGCTCGGCACGACG\nATGAGCGCCGCTCGCCGCGCGATCTGCTGCGTGTCTGACAAACTCGACATAC  
CGGTCATCGAAGACA\nACGTGTACGGCTTTCTCGGCGACGAGCCGCCGCTGGCCGCACTCGCCCCGACGCTGCATCG  
TGATCGA\nCAGCCTGTCCAAGCGGGTCACGCCAGGCCTGACCTCGGCTTCATCGTCCGCCCGCCGCGCTGCGCGAA\  
nAGCGTGATGGCGTTCGCTACGCTCGGGGGGATGGACCGCTTCGGGTTTCGCGTTTGCAGCCGCTCAGCGCT\nTGATGCG  
TGACGGCACCGTGGCCGAACCTCGCGAGATTGAAACGCATCGACGCAATTGCACGTCAGGCGCT\nGGCGATCGAGCGTCT  
GGCGGGTTTCGATGTCCGACCAACGGCAAGTGCTATCACCTGTGGCTGACGTTG\nCCGGCGCACTGGCGTTTCGAGGC  
CCTCGTTCGAGCCGCGGCCCGCGCGACATCGGGCTGACGCCGTGCA\nCCACCTTCGCCGTTTCGTCGGGCCATGCGCC  
GAACCGCATCAGGCTCGCGCTGGCCGCGCGAGCATGGA\nTCAGCTCGACCGCGCCTGCGCACGCTGACCGCGATGCT  
GAACGGCCGCGAGGGCGATTTGACGCGACC\nGAGTGAGCATCGCGCGGATGTGTCTGTACATACCCGCTGATTC  
GGTATCCGGATCAATATCAAGCC\nGTCACCTTCGTGAGCAACTCGGCCTCCAGCTCGCGCAACCGTTCCGGAAACCGGAC  
GCGCAGCACAGGAA\nCCACCCGAGCCCGATCGCCAGCAATGCGACGAACACATACGGCAGCCACGCATAGA\n'

```
[22]: c2 = a2.count(">")
      c2
```

```
[22]: 18
```

```
[26]: length_a2 = len(a2)
      length_a2
```

```
[26]: 48593
```

```
[28]: splita2 = a2.read().split("\n")
      splita2
```

-----

AttributeError Traceback (most recent call  
last)

```
<ipython-input-28-5f9b8505c0f1> in <module>
----> 1 splita2 = a2.read().split("\n")
      2 splita2
```

AttributeError: 'str' object has no attribute 'read'

[10]: a

[10]: '>gi|142022655|gb|EQ086233.1|43 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nTCGGGCGAAGGCGGCAGCAAGTCGTCCACGCGCAGCGCGGCACCGCGGGCCTCTGCCGTGCGCTGCTTGG\nCCATGGCCTCCAGCGCACCGATCGGATCAAAGCCGCTGAGCCTTCGCGCATCAGGCGGCCATAGTTGGC\nnGCCAGTGACCGTACCAACCGCCTTGATGCGGCGCTCGGTCATCGCTGCATTGATCGAGTAGCCACCGCG\nnCCGCAAATGCCAGCAGCAGCCAATGCGTTCTTCATCCACATAGGGGAGCGTTACGAGGTAGTCGCAGACCA\nnCGCGGAAAATCCTCGACGCGCAGTGTGCGGTCTTCGGTAAAACGTGGTTCGCCGCCGCTGGCACCTTGAA\nnGCTGGCGTCAAGGCGATGACGACGAAACCTTCCTTGCCAGCGCCTCGCCATACACGTTCCCCGATGTT\nnTGCTCCTTGAGCTGCCGATCGGATGCGCGCTGATGATGGCGGGATATTTCTTGCTTCGTCGAAGTTCG\nnGCGGGAA GTGGATGTCGGCTGCGATATCCCAATACACATTCTTGATCTTGACGCTTTTCATGACAGCTCC\nnGTTCAGGGGGAGGGG GTAAGTTCGCCAGGCCGAATCGTTGGTAGCCAAGCGGCAACGACTCGAATATAGA\nnGAGCCGATTGGAATTCCGTAAGATCGCAATCTGGACTACAGTGGTATCTTCAAATTGACAATGGCACCTA\nnCATGGATCCCTCACTGCTTCCGTCTCTCGCGTGTTTCGCCACGTGCGACATCATCGTAGCTTCACGAAA\nnGCGGCTGCGGAAATGGGCGTTTCTCGAGCAAACCTGTGCGAGAACGTGAAGGCGCTCGAACGCCGTTGA\nnACGTCAAGCTGCTGTATCGAACGACTCGCGACATGTCGCTGACCGAGGAGGGGGCAGCGGCTCTACGAGGT\nnGTGGTATCCCGCGCTGGTCGCGTTCGAGCGGACGGTCGACGCGCTGCACGAGGAGCGCGACGAGCCGTG\nnGGGTTGATTC\n>gi|142022655|gb|EQ086233.1|160 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nATTGGGGAGGAGGCGAGTTGAGCGGCGGCAGTTTCGCTGCGTGCCTGCGCGCGTCGACATCTGATCCG\nnACCATGGAATCCCGCTCAATCTTTGAGCAGGGATGCGGGGCGATCAAGATGGGGATGCGGGATGGG\nnGGCGACGGTGTATTTCCGCCAGAGATTTTCGCCGCGGGAGCTCGCGGTGCGTACGTGCATGTTCAAACGC\nnACGGTGCGCGCATGGCAGTGGCAGACTGATCAACGCAGCTGGAAGCATCCGAAGCGCGGGGACGCGTG\nnTCCTCGACGCGTGGCCTCACATGCTGTCGGGTGCGTTCAAGACCGAAAAGCCACCGACCGACGCGGAGCA\nnATGCGCTACGCGGATCGCGTTCGACACGAGCCGCGCGGAGGCAAGGCCGACGTATTCGATCTTCCAGAG\nnGAAGCCTATTGGCTCGAGTCGTAGTCTCGATATGGTAGAGCAACATGAATCCCGGGCTAAGTACAAGAA\nnGTAACCCGGAACGAGTGAGATTGCGACGAATAAACGCTTCACCATGATCGCGCTCCTGAGTTGGTTGAG\nnGTGAATTGGAAGTCGATTCTTGGGGGATCATTTCCCGCAAGGCGCGCAATCCCGCATTGTTCTCAAGA\nnTCGCAACGCGATTCTGTCAGGCCGATCTTCATGGGGTGTCTCGCTGGTAGTGATTCCGTGCTGGCCCGCGC\nnATGTGCA TGACGGCATCCGGGGAG\n>gi|142022655|gb|EQ086233.1|41 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGACCTTGATCGGCTGATCGCCGAGCGTGCCGACGTATTCAATCACAACCTGCGGCCGGGGGCGGCTGCG\nnGATTATGGGCTCGACGCAAACAGCCTGCGAATTACGAAGCCGAGTTGATCTGCTGCGAGATCGGTGCAT\nnTCGGCCATGTGGACCGGATGAATACGCTTCCCGGTTACGACCCGCTGATGCAGGCATTTTCCGGGATCAT\nnGAGCATCACGGGCGAAGAAGGTCAGGCACCG



GTGCGAGCAGGGGTGTCGATCGTCGATTTGCGGCGCGGGC\nATGTGGGCTGCCATGGGCATTCTGGCAGCGCTGTACCGG  
CGCCAGATGAAGCACTGCGGGGCGACCGTCA\nATGCCTCCCTGCTGGAGACCGCACTGGCGTGATGTCAGTCGGCATC  
GCGAACTACAACGCCGACGGCGA\nACCGGTGGCCGGCATGGGTCGGGCGTCGCGTTTATCGTGCCGACCGTGCGTTT  
GCGACCGCGGATGGC\nTTTCCTCATCGTGAGTTGCGCGAACGATCCGCTCTTCGCGCGGTTGTGTGCCGCGCTCGACCAT  
CCGGAAT\nNGGGCGACGGACGAGCGGTTTCGCGACGAACGCGGGGCGCTGAAGAACCGCGCCGAGATCGACCGCCTGAT\  
nCGGCGACCGGTTGTGACGGGAACGCGGGAGTTCTGGACGGAACGGTTGCAGGCCGCGGGAATTCCGGTT\nNGCGCCGA  
TTCAGACGACCGCGGAGGTGCTGGCGCACGAGCAGACGCGAGGCGCTCGGCATCATCGAGAAGC\nCGTCGGACGACGAGA  
TCGGTATCGTCGGCCTGCCGCTTTCGTTTCGACGGCAAGCGTTCCCCGCGCTGCA\nTCTGGCGAAGGATACCGGCGCTG  
ACAATTCTCTCCTCGACAGTCTTCTGAAGGTCAGTCATTAGCCGCTT\nTTTTCCGAAGCAGACGACGATGCGGCATGGCG  
GATACAGTGCCGTTTCGAGTCGTCAAGCGCCACTTGGTGA\nTCCCGCCAGCCGCGCAATTTGGGTCGGTTGGCGGGTCAT  
TTATTGGCGTACCGGACAGATGACGTTGGTG\nTACGATCAGGGTCGACGTCAGTCGACGAGGAAGCACATTTCTGCTGCC  
ATCGATGCGTTATCGTGGTGAGG\nCGTTTGAGAGTGGAAGTGCACGCCGTGTGACATGACGCCTGACAACGTCATGAAG  
ACGTGCACGTGAGGA\nGACATGAGTACATTCAACCGATTGCACCTGATTTCGTCAGGTGGATTTGTTTACGCTGAAGCTG  
TTCGTGT\nCGGCGGTTGAGGAGCAGCAGATCGGCCTGGCCGCGATCCGGGAGAACATCGCCGCATCCACTGCAACCAA\  
nCGGTATTTCAGGATCTCGAAGATATTGCCGGCGTAAACTTCTCGAACGAGGCCCGAAAGGCGTCTCGCCC\nAGCCCGG  
CCGGTGCACTACTCCTGCGCTATATCCGCAAGATTTTCGACGATCTGGACGATATGCGATCCG\nAGATCGCGTCGTTCA  
CCGACGGCATGCAGGGCGAGCTGACGATCGCATCCGCACGGTCGATCATCGTTCC\nNGTTTCTATCGCGCGAAGTAGGCG  
AGTTCAGTCGGGAATACCCGCTTGTGCAACTTGTGATGCAGGAGCTC\nGAAAACGCTGAAATCGTCAGGCCGTTGTTT  
GGGGCGAAGCCGATGTCGGGGTGTTCCGCCGACACACG\nAGCTGGATCTCGGCGGTGTCGATGTCACGCCGTACCGGG  
AAGACAGGCTGATTGCCGTGTCGCCGAAGGG\nNGCATCGCCTGAGCGAGCAGGCCAGCGTGACCTTCAAGGATCTCCTGC  
CCGAGAATGTTCATCGCCGTACG\nGCGATGCTGGGGGCGGTACGTGCGGCCACGCGGCGGCTGGGCGAAGAGTACAAGC  
CGCGCTACAGCGTCC\nGGAGCGCGGTGTGCGAGTCAGCCTGGTGACGGCCGGGTTAGGCGTCACCATTCAACCGGAAT  
GCCTCGT\nnCAGTCATGAACTGTTTCAGCCGGGTTGCCGTGCTGGAATTGGCCGAACCATGGGCGTTGCGACGCATTTCAT\  
nATCGCCACGGCGCGCGGCCGAGACCAGAATCCGATTGTGCGTGCAGTTCATCAAGCAACTGCTGGACCGGC\nCTCGCGA  
CTGAGTGGCCGCGTCGCGCGCTGGATGGCCACGTTCCCGCACTTTTCGACATCGAAATCACAACG\nACAGACAAGGAGAGG  
CCAGATGCAGCCCATTTGTGCTTGGCAGGTTGGAGATTGAGAAGTTGTCGAGATG\nGAAACGAGCCTTCCGATCGAAGT  
GCTGTTTCCGGGCGTCACTGCAGGCGACCTTTCCAGCTGAAGGGCT\nGGTACTGGGACGATGTGCTTTTCGCTCGATCC  
GGCTCAAGCGGGGTGCGGCCTGAGCATGCACAGTTACCT\nCGTACGTTTCAGCGGAAAGACGATCCTGATCGACAGCTG  
CAACGGGAACGACAAGGATCGCTCGACCGAA\nNGCCTTTTCATCGGCTGAATACGCCATGGCTCGACAGGCTGCGGAAACT  
CGGCATTGCGCCGAAGATGTCG\nACATGGTCATGTGCACGCACCTGCACGTGGATCATGTGCGGTGGAACACGCGCCT  
CGACAACGGGCGCTG\nGGTACCGACTTTCCCGAATGCGCGCTATGTCATGAACAGGCGCGATGTCGACCATATCAGTTC  
GCCGAGC\nnACCCATGCGATGATTTCGCGAGCCGTGGCTCGATTCCATCTCGCCGGTGATCGATGCCGGGACGGCCGATC\  
nTCGTCGACGGAGACGAGGTCATTCTCGGGAAGGTGGGTGACGGCATGTGGCTGCGGCCGGCGTATGGCCA\nCTCGCCC  
GGCAGCTGCATGATCCACGCTTGTGCGGTGGTTCCCGCGCGTATTTTCTGGGACGTGGTG\nnACCATCCAGTTCAG  
CTGATCCGCCCAGATCTTCATTTCTGTGACGACTGGAACGGCGATCTGGCGGAAC\nnAGGTGCGCATCGCGCTGCTCGAG  
GAAATTGCCGGAACCGATACGATGCTGTTTCCGGCCCATTTCAGATA\nnTGGATCGGGCGGGCATGTGAGCGGGACGGC  
GACAAGTATCGCTTTTCGTTTACCCAGGTCTGACGAGTC\nn>gi|142022655|gb|EQ086233.1|221 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ce\nnGCCCCGCGATTTGACGTCTGCAGCCTCACGTCCAACTCAATTTGAGATTCTTGACAGACAGATGGGCTA\nnATTA  
AAGTAGGTGATCGCATTATGGTGAACATACGATCACCTGCCGATGCCAGTTCAGTCAAGTGTATA\nNGGGGTAAAAATA  
GTCAAAAATATGTAATTAAGGGGGCGTGATATGTTTCAAGCCGCCAACACACCGCCA\nnCCCAGCATCCACCTAGCCC  
AGCGGCCGCCAATTCCCACCTGTGGCCAAGCTTGTGAGGGCCGATCCAA\nnTACCACCATAAACAGCACCAGATGCAAT  
TGCATAAACCAATCCGCTTACGCCCCCGCATTCTGGTTTG\nnGTTGGCCTTCTTGCTCCTCTCCCCCGATACTAGCA  
ACAGTTCATAAGCGGTTAATTCTTGAATGCAGCC\nnAATCTCCAATGAATTTTCAATTGCCGAAGAGTGATTTCATCATTT  
TTTCTTCCGTGTATTTGCATAGTTAA\nnAAAGAACTTCCGCCAAAACAACGGATATAGCACCTGCAAAAACATCTCTCAA  
AATCGATCCTTCTAGAAT\nnTTCAATTTTCCCCAAAATAAACCATATCGAACGCCAAGCAAAAAATCAGAAGGAGTTT  
TGCGGCAACC\nnGCGTTAATGTCCATATCGATAATTTTTTATTTCTCCAGATCACGCCAGAAGGATGAGCCGACACTGTC

AT\nACGATATTCATGTGGAAAATACAAAAATGTAAGCTCAGATCGTCGATTTCATAATCGCGCCACGAACTGCA\n\nATAC  
ACGTTTCTACCGTAGCACTCCGTATTAATGCTCGCAACCTCCTTTTGGAGGCAATCCTATCAAGTT\n\nAAGCCGGAACCC  
TTGTAAACAAAGGCTCGCTCATTGACGATGCGCCTCAACGCCATCTTCAATATTTTCGC\n\nGATATGAAATTGGCTTGACC  
CCATCTTGCCAAGCTCCTCGAAGGCCTTCATACGGGCGATCTGTGCACT\n\nCGCCATGTGCGACGATCCGCGTCGCTTCA  
GCCGACGCCGACGCTCGATTTCTGTC AACCTGAGCTGTTGGT\n\nTTCTAGCGCCCCGCGCGGGGCCATACAGACCGAACT  
CGACTGTTTCTTCGCCATGCTGATCGATCAGGCT\n\nNGAACTGATCCGCCGCGTCAGTGCGCAAGCCGTGCCACGGCGCG  
ACAGCACTTGCCGTTCTCCGTATTCA\n\nCCGACCTCAATGATCAACTGCTCACCAGCATCGAACAGTACTTGTTACGCT  
CCGATGGCCCCGCGCTGCG\n\nGGTCGTGCGGCCAAAGGCTCCGACCTGCGCCTGACCTCGCGGGATGCCACGCGACGCAT  
GATCACCGCC\n\nTGATCCGCGTCCGCGCGTGCTCACGACGAATCTGCTCGACGCGCACGCCTATTGGCCGCCGACATTG  
CT\n\nGCGCTTTAACGCCGCTCCAAGGTACGAAGTCGCTGGCATCGTTTCTCGATCTGTTTAACCAGCGAGTTG\n\nCACG  
TCCGTTGCCCTTTTCGACAAGGCTGACTTGCCGTGATTGTGCGCAATATTGACGGACCGCGATGTCT\n\nTGGTGCTGGGAG  
GTTTTCCGCAGAAAGGCGAAGCTAACGATGCTCCGTATTTTTGCTCGATGGGCCCGGCC\n\nTGTCGATCGATTTTCTGGTC  
GGAAAAAGAGAAGTTTTCGCTGCCTCTATCGCTAACCTTCATTAGCGAT\n\nAGGGTTACGTTTTGGACAAAACTGCCA  
TTTTGCAAAACGCGTCACATCCACCTGTACACCGCTGTCTAT\n\nCCCGAAGCCTATAACTTTTGGATGCCTGTTCTTCCG  
GACTCGCTCTCCTTTCCCGACGCGCTTCCCTGG\n\nCCGTCTTGCGTACCTGGTACGCGAGCGTCGGCTCGCGCGAAGCC  
GTCGAGCGTTATTGTCCCGACGAGCT\n\nCGGAGAAGGAAAGTCGGCCCGTGGTGTGACTGGTCGTATCCGACGGAAACTC  
GTCGAGTTTGCCATCAGT\n\nCGTCATCGTGACGCTCTGGCGAAACTGTTTCAGTGCGAAGCAGGCGAACGGATTTCGTAC  
CGAAAGCGG\n\nTGGCCGGGGCGTCGCCGCGTTGCATACGAACGGCATCAGGACGCTGGCCGAGCTGATTGTGCGAATT  
TC\n\nCCGGCGCGCTCGCTGGTGGCTCGGCATTCCCAGTCTCGGGGAGCGTAGCGCGACCGAGTCGAAGCCTTC\n\nTTTTG  
CGCTTGTTCCGCGCTGACGGAGCGGGCTCGTTGCTGGTGGTGGCGACACCCCGGAATCCGTGA\n\nCGCCGTGGGAAA  
ACATCCGCGTCCCCCATGATGTGACGGTTTCGCGCGGCATCTTCCGTGCGCCGAAGGC\n\nGATGTGCGCACGATCCCAAT  
CAGGCGAAGCGTCCGGCGTACGCCAGGGTCATAACTGTTTTCCCGGGA\n\nGTAGCGAATGGCATCGAGCCAAGGCACA  
GTCGATTTTCATCGTCGAGCAGATGGCGGCGCGCGGTACGGTA\n\nTCGGCGCGCAAGATGTTTCGGCGAATACGGCATCTAT  
TGCGACGGCAAGATGGTCGCGCTCGTCTGCGACG\n\nATCGGTTGTTTCGTCAAGCCGACGGCCGAAGCCGCGCTTTCTC  
GGCAGTGCGAAGAAGGCCCGCCGTA\n\nTCCGGCTGCGAAGCCGCACCTCGTCATCGCCGGCGAGCGCTGGGACGATCGC  
GAATGGCTGTGCGCGCTG\n\nATCCGGATCACCGCCGCGCAGTTGCCGTGCCGATGAAGCGGCGCCGATAACGGCTCATT  
CGCGTGCCAT\n\nGTTGCGCTGCGGCCAGGCGTATTCTGTGACACATGTGCGCGACCATCGCGCATCCGGCGGGAGCACA  
GA\n\nATGAAGACATCCGGACATTTCGCGCGGCACGTGGCTTGCTCGCCGCGCCCGGCGCGCTCGTC\n\n>gi|1420226  
55|gb|EQ086233.1|294 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold,  
whole genome shotgun sequence\n\nGATCAGCCCCGCATACGCTACGCGCTGCGACGCCGAAGAGCGTCTGC  
TGCCGAGGATCTTGCGCAAC\n\nTCGAAGCCAATGCCTGCCATCAGCGGCCCTCGTTGCGCGAGCCGGATGCAGCCACGC  
CCGCACCCGCTTG\n\nATGCACCGGGCAACCGGCTTCGAACGACGCCCCGCGACGCATCCCGTCGGGGGCGATCGCGAGCG  
CATCG\n\nTACAACGCCCGATACCGGTGATCATCTGTGCGCGCTGTAGAAGCGTCGGACCCGGGAGAGCCGGCTT\n\nG  
CTGCGCGGCTGCCAGCGCTGCGCGTCGCCGAGCAGTTCGAGCGACGCTGCGCGAACGCGGCAGGATT\n\nCGCGATCGG  
CACGACCGCACCGGCGGCCCGAACGCGCGGTGCTCCTCTCCGAGCCCTTCGATCAGTTGC\n\nCGGCACGACCCGACGTC  
CGTCGTAATCGACGGCACGCCGGCCGGAAGCCTTCGAGCACGACGAGCGGGA\n\nGCGCCTCGCTGATCGACGTCAGCGC  
GAGCAGTCGACTTTTCGGCAAAATGTCGTGAATCCGCTGAAAGCC\n\nGAGAAAGCGCACCGTCGTGCGAGCCCGAGGCT  
CTGCGGAGCGCCCGGCATTCCAGCGCATACGCGGGA\n\nTCCTCTTCTCCGGCCCGACGATCCAGCCTTCGATTTCCGG  
CATCGTCCGCGATGCGATGAACATCGCGC\n\nGGACGAACGTCTTGATGTCTTGATCGGCACCACACGGCCGATCAGCGC  
GACGATGCGGCGCGGGCCGCC\n\nGTCCCGCTTGTCAGCAGCGGTGCGAAGCCGTCCACGTCGACGCGCTTCGGAATGTT  
GCGGGTCCGGGCG\n\nGCATCGGCCCGTCAGCGATCTGACGGTGTGATTGGCCTCATAGAGCGCGACGATCTCGTCGGC  
CGCT\n\nCGTATGCGAGCTTGCCGAGCGCTTCAAAGAAACGCACCCACAGCGCGCGAAATAGCTGATCGTCGATAC\n\nG  
TCCCGCTCGAACGCGCCCCGTTGTCGCGGATCCACTCGCTCTGACGAGATCGATCTTGCGTTCCCTTG  
nGTATAGATG  
CCGTGCTCGGAAACGAGCAGCGGTGCGCCCGTTCGGTAATGCAACAGCGCACCGAGAAAGC\n\nCCGCATAACCCGTGAC  
ACGGTGTGATAGACCCGCGCCGGCGGACGCCGCGTGCGATCCGCGCAAGCTG\n\nCCACAGCGGCTTGTCATGATCCGA  
ACCGTCCAGAAGTAGTCGGTGAAGGAAGGGTCGGTGCAGTAGCGG\n\nTGATAACGCTCGACGATGAAGTCCACGCGCG  
CGCTGGTGAAAACTGCTGCTCGTTCAGCGGTCCGT\n\nCATCGCCGATCAGTTGGACCAGGTGCGCCATCAGTGCAGCC

GGGTTCCGCGACGCCGTCCGCGCTCAG\nCCCTCGCCACGCATCGTGCAGCGCAGCCGAACGCTTGAACGCTTCGGGG  
TCGCCGGGAATCTGCCGTGCG\nCCGCGCGTATCGGCAGGCACGTGCCATACAGATAATGGGATTCTGAAGTGAACGACG  
TTGTCGGGCAACG\nCGTATGCGGCGCCATGATAGTCTCTTCCCGGTGCCGATGAACACGATCGCGAAACGCTTGTGC  
GGATA\nTGACGGATCATCTCGTTGACCCAACTCGATACGCCGCTCGCACATACGGGAACGTGCCCTTCGAGCAGC\nA  
GGCAGACGTGCATCGCGGTTCGACGGACCGGCACGGCGGTGAATGGACGGATTTCATGACGACCACTAGCG\nA  
ATTGAGAACGGGCAGCGCGGCCGCTTGGCGAGCGACGCGAGCAGCTCGGCCGACGCGAA\nTACTCGCCGCGCAGGAA  
CGCGGCTCCGCAAGCCACGGCACGAGACGCTCGCGCGAAAGCCCCATTCTGA\nGCGCGCGGCTCATGTATTGCACTGC  
TGCATCCGCGTCCGCGTTCGCGAGTGCAGGGCGGCCGGAACGAG\nCCACATCGCCGCGTTCGCTTCGTCGATCGCGAG  
CGCGGCTTCGCGTGGCGGTTCGCGCTTTCGCGCGTG\nTGCCCGTGACCGCGCGAGCAGAGATTCTGATAGACCAG  
CTCGAAGTACAGCTCGGCAAGCCGCGAT\nGGGTTCGCTGGCGGTTGCGTCTGTCGTTGGCGGCGTCGAGCGTCTGCTG  
CGCGCGAAGATGTTCTGCGC\nGATCTCGTTCTCGCGCGGTTCGAGCGTGCCGATGCAATGAGCCGACGCTCTTCGAG  
AGGATCGGCGAGC\nAGATCGCGCAGCAGCGTGCCGTCGTGCGCGTCGGCATGCTCTGGATCGCCACCAGCGCGGAAAG  
ACGAT\nCCGACGCGGCCACCCGCTATTTCAGGACGCGCCTGCAGACGCGCACCGCCGCGCTGCGAGACGCGCGA\nC  
ACGAGATAGGTGACGAATTCGGGCAAAGGCACGTGCGGCTATGCGTCTGCTCTTCGCGGGCCAG\nCAAGCCGCC  
CACACACCACTCGCCAGCAGACAAGGCCGCCGCGACGGGCACGAACGTGCATGCGAGCC\nAGAGCGCCGCGAGCGTC  
CCGTTGCGCGGCTTCGTAACGTTGCGGCAGCAACTGCCGTTACAGCACCGC\nCTGCGCGCGCGCCGCTGCCACCATC  
GCGCGAACACGGGCCACAGGGGTCAACCGGCGGTGCGGGAACG\nGGCATCGCCGCCCCCGCTGGGCGAACCCGCGG  
AACCGCGCGAGCACGGCATACTGCAGCGCCGCGACGA\nTCGCAATGCCGACGGACGACGCGAGTGGCGCCAGCATCGAC  
GCCGCGGCACCCGTAACGCGTGGCGCCC\nCGGACTGGGCGCCACGAACCGGATCGCAGCCTACACATGAACGCGCT  
TCCCTTCAGCAGGCGCTGCAA\nTGCCGCGCCCGGTTTCGAGGCTGCCCAGATGCAGCGTATGCACGCCGATCCGCGCGCC  
TTCGAGATCGAGA\nCCGAACCTGCGCATCGAGGCTCGCCTCGATGCGTTCGAGATAGCCGTCGACGCCGCGCTATCGGT  
TGCCG\nGCGATCAGGTTACGAGCACCGATTGCGTTCGCGCTCTCAGCGCCATATGAGATCGAGCGAACGGCGCCC\nG  
CGCACCATGCTCGAACAGCGAATCGTCCGCTCGTCTGCTGGAACACGAGCGCAACCAGCGACGAC\nGCGGATCCCC  
GACTCGCGTGCAGGCGGTTCAGACGCGCGATGTCGAGCGCAATTCTGACGGGCATGCCG\nGTACTGCCGCCACCACG  
TCGCGCACGCACGCCGATGCTCGACGCCGTCCGCGTAATAGCCGAGCAGCAC\nCAGCAGCAGTTGCAGATTGTGCAAG  
TTCAGCGACAGAAACGGCATGCGCTTGATCGCGAGCAACGCGATC\nAAGCGGCGTCGGCGCACACGACCGGCGCGGCG  
ACCAGCAGTGCCGTTTCGGCGCCACGTGCGTGTGTC\nTTTTTCGATGGGCGACGATTTCTGCTCTCGAGCGCTCGGGCC  
ACGAGGCCGTTCGCGCGGATCGAAATCGAA\nTGATCGCCGAGCCGCGCAACCGGTTTCGCACACGATTCTGGGCGAGTTC  
GTGCGACCTTCGACCGGATAC\nAGCGCGCCACCTCGATCTGACATGCTGTGCGACGAATTCGAGCAGCGGTGCGCA  
CCGGGTAGCCGCA\nCCTTGCTCGCGCGCCCTGCGGCGGCGCATCGAATCCGTGCGCGAGCGACAGGCGGCGCAATTCTG  
GTGAT\nCGAATCCCGCAGCGTGGTTCGCTTCGACAGCAGATCCTTTTCAGCGCTCGTTCGACAGGCGCATCAGG\nT  
AGTGCTCTTCGTGATCGCGACGAGCCGTCGTTTCAGATAGTCTGTGAGCG\n>gi|142022655|gb|EQ086233.1  
|323 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
shotgun sequence\nACGCCGCGCACCGCGAGTACCGCGCGCCGGGCACTCCTTGACCCCGCATGATCGATTCC  
CGATGAAA\nCCCGAAAACCTCGTCGCTGCCACGAATGCGACCTGCTGTTTTGGCGGCCCGCGCTTCGCGCGCTGG  
\nCTGCGCACTGCCGAGGTGCCGTGCCGCGTGGGCGGCAGCGCGCACGGCCGTCCGGCGCTCGACCGGCG\nG  
GATCGCGCTCGCCGCGCTGTTACGCTCTTCATCGCGCAGGCCTTTCCCATCGTCGCGCTCGAC\nGCCCGCGCATCGC  
ATCGCACGCGACGCTGGCCGACGCGGTGGCCGCGTTGCGCTTGAACGGGCAACCGG\nCGGTGGCGCGCATCGTGTTCTG  
CACGACGATGTTGTTCCCGCTGTGGAATCGCCGCGTGGCTGTACGT\nGCTCGTACCGTTGCGCGCGGGCCGCGTACC  
GCCCCGCTTCGAGCCGCTCTGCGCAACATGCAGCGGCTG\nCGCCGTGGAGCATGGTCGAGGTGTTCTGCTCGGCAT  
CCTGGTCACGATCGTCAAGATGACGAGCCTCG\nCGCACGTGATACCGGGCCCGCGCTGTTTTCGTTTCGGCGCCCTCAC  
CGTGTGCTCGGCTTTCTCGCGTC\nATTTCGACCCGGGCGGCTGTGGGAAGCGCGCGACGAATCATCGCGCTGCGCGG  
CGGCGGTACGTCCGCC\nGCGGTATCGCGCCGGCGGCACACGCCGCGACGCGCTGCACCGGTGACGCCCCGACACAGCGGA  
CGAACGA\nACGCGACCGGCGCGACCGGGCGCCACCGCTCGGCAAGCGTCACGGCCCGCGCCGCGGGGCTGGTTCGATG  
\nTCATACCTGCGGACGCGTGCAACCGCATACCGAAGCCGCGACGCGCGTGCACGCGCTGCGGTTTCGACG\nCTGCAC  
GAACCGCGCCCGCGCAGCGCCGACGACGGGCGCGCTCGTCATCGCGGCCGCGCTTCTCTACA\nTCCCGCGCAACCTG  
CTGCCCATCATGCATGCGACCTCGCTCGGCCGCGCCGAGGACGACACCATTCTCGC\nCGGCGTCGCGTACTTCTGGACG

TCCGGCGACTGGCCGCTCGCCGTCGTCGTGTTTCGTCGCGAGCGTGCTC\nnGTGCCGATGCTCAAGCTCGCGATTCTCGCG  
TTGCAGGCCATTGCCGCGCATCGCGGCACGCCATGGCGCC\nnCGCTGGAACGCGCGCGGCTCCATCGACTGGTGGAGCGC  
GTGGCCGCTGGTCGATGCTCGACGTGTTTCGT\nnCGTCGCACTGACGATCGCGCTCGTGCAATTCGGTTCGTTCCGCGAG  
ATCACGCCCCGGCCCCGGCGCACTG\nnGCATTCCGGCGCAGTGGTCGTGCTCACCATGTGTGCGTCGATGCAGTTCGATCCG  
CGACTCATCTGGGACG\nnGCGCGCACCGCTGCGCCGCCAGCCCTCGATCCTGAACGGGCGCAAGCCCGTCACAACGCTAC  
GTCCATTTC\nnACAAGATGCTTCCCCCTCGATTTGCCTGAACCCGAGATCCGGCCGCGCAGCCGCTGGATCCCGTCCCTCGT  
\nCTGGATCGTGCCGCTCGTGTGCGCGCTGATCGGGCTGGCGCTCGTGTACCGCGGCATTGCCGCGACCGGC\nnCCGACG  
ATCACCGTCACGTTCCGGAACGCAGAAGGCCTCGAAGCCGGAAGACCAAGGTCCGTTACAAGG\nnACGTCGATATCGGC  
TCGGTGACAGGCCATCACGCTCACGCCCCGACTTCAAGCGCGTGGTCGTGCGCATCCA\nnGCTCACCAAGGACGCCGCGCAG  
TTCGGAACCGCGATACGCGATTCTGGGTCGTTCCGGCCGCTATCGGC\nnGCCACCGGTATCTCCGGGCTCGGCACGCTG  
CTCTCCGGCGCGTACATCGGTGCGGACCTCGGCCGCTCGC\nnCGCACGAACAGACCGACTTCGCCGACTCGAGACACCG  
CCCGCCGTACCGCGGACCAGCAGGGCCGCGC\nnCTACCTGCTGCACGGCGATTTCGCTCGGCTCGATCGACATCGGCTCG  
CCGGTGTTCATCGGCACCTGCAG\nnGTCCGGCCAGGTGGTCGGCTTCTCGCTCGACAAGGATGGCACGGGCGTCGACGTG  
CAAGCGTTTCGTGAAGG\nnCGCCGTACGACAAATACGTCGGCACCCAAACGCGCTGGTGGCACGCAAGCGGCTTGGACCTG  
CGGCTCGA\nnTTCGAACGGCCTGAAGCTGAACACGCAATCGCTCGCGACGGTGGTCGTGCGCGGCCCTCGCGTTTCAGTCG  
\nCCGCCCCGGGAAGCCGATGCACCGCCGGCAGCCGACAACGCGTCGTTCCGGCTCGCGCAGGACGAAGCCG\nnATGCGA  
TGCGCGACCCCGACGGCACGCCGCTGATGGTCGTTCATGCGCTTCGACACAGTCGCTGCGCGGACT\nnGTCCGTCGGCGCGC  
CGGTCGATTTGCGCGGCATCGCGCTCGGTCAGGTCACGAACATCGGGATCGAGTAC\nnAACGAGCACTCGCGCACGTTCA  
GCATGAAGGTCACGATGGCGCTGTATCCGTCGAGACTGAGCCGGCACA\nnGCGACACCGCGCTGCCGGCCCCGATACGG  
CCGGCGGCCACGACCTCCTCGAGCACCTGGTGCTGCAGGG\nnGTGCGCGGGCAGTTGCGCACCGGCAGCCTGCTCACGG  
GCCAGTTGTACGTGGCGCTCGACATGTTCCCG\nnAAAGCGCCGCGCGGAGCGTGACGTGCACCGCACGCCGGTCGAGC  
TGCCGACCGTTCCGAACACGCTCG\nnACGAACCTGCAGGTTTCAGCTCGCCGACATCGCTCGGAAGCTCGACAAGGTTCCAT  
TCGACCAGATCGGCAA\nnCAGCCTGAACGGTGCGCTCGAGCATGCGAACCAAGTTGTTTCGGACACCTCGACGAGCAAGTCG  
TGCCGCGA\nnGCGCGCGACACGCTGGCGGCGGCGCAACGCACGTTTCGACGCCGCGCAGGCGACGCTGCGGCAGGATTTCG  
\nCGATGCAATCGGACGTTTCATGACGCGATGCAATCGCTCACGCAGACGCTCCAGTCGCTCAATACGCTGGC\nnCGACTA  
TCTCGAGCGGCATCCGAAGCGCTGCTCTTCGGCAAGAAAGGAGAACCGAAATGACGCCGCATT\nnCGTTTCGCACGTTTC  
CGGATGCCGATCCGCATCGCGACGTGCATCGCGCTGGCCGTGCTCGGCGCCTGTAC\nnGTCCCGCCCGTACGGTTCCAT  
ACGCTCGGGATGGCGGATGGCGCGGGCGGCGACACCGACGCCTCGCGT\nnCCCCGATGGCTGATCGACATGCAGCGCGTG  
CACGTGGCGGCGCCGGCGGACGGCAACCGGCTCGCGGTGC\nnAGCGCGGCCCCGAACGGGTCGACATCCTGGAACAGGAG  
CGCTGGGTTGCGCCGCTCGGCGACGAGATGCG\nnCGACGGACTGTCGACGCGCGTCACGTCCCGGCTGAACAGTTTCGAC  
GTTACACCGCTCGCTCATCCGGAT\nnGGCACGCCGGTCTATCGGGTCGCCGTGGACGTCCAGCGTGTGCAATCGTGGCCG  
GCGTCTCACGTGCTGC\nnTCGATGCGACGTGGACGGTGGACGCCGGCTCAGGACAGCCGGCACTGACTTGCCGCGAGCATC  
GTTCCGGC\nnCGGTGCGTCGGCGGGCTACGACGCGCTCGTCGACGCGCATCGCCATGCGCTCGACACGCTCGCGCTCGGC  
\nATCGCCGCCGGCATTTCGCGCGGGCCACGCACGCGCCGCCGCCCGGATGCCGCGCGGAACACCCG\nnACCGAC  
GGATCGCGCCAGCCACCGCCGAACGCGGCGCGCCGGATCCGCTCATCGATTGAATCGGGCAGGG\nnAAAAACGCGCGGCCG  
GGCCGGGTTCGGGTTCGGGTTGGGCGATCGGCCCTCGCGCCGGTGCCCTGCCCGACCCA\nnGCGACATCGTCAGGCGAAACCG  
TCCCGACGAACCACGAATCGGCGCGGCGGCTTCCCGCTCGATCCGG\nnTATTTTCGCGGGCGCGGTGTTGCGCCGCTT  
CGATCATTTCCACCTGCATGGACCACCCTGGATGAAACCC\nnGCCCCGACTCACGCGCGAGCAACGCAGGCTCGACACCCG  
CGAACGGCTCATCGCTACCGCTCGCGCACTAT\nnTCATCGCGAAAGGCTTTGCCGATACGAGCGTCGAACATATCGCCGA  
GGGCGCCGGGTACACGCGCGGCGC\nnGTTCTATGCGCACTTCGAACACAAGCGCGAGCTGTTGATCGAGATCCTTCGCCG  
CGACCATGATCGAATG\nnCTCACGAAAGGTTCGGCTAGTGGCACCGGTGAGCGGCAGCGCTGCACCACGGCCGACTACGC  
GGAGATTG\nnCCGCCGGCTGGGAATGTTTCCCGCTGTGGTTCGAGGTTTCATCTGTACGCGCTGCGCGATGCCGGGTTTCG  
\nGCACATCGTCGACTGCCTGCACACCGAACCCGCCGCGCTACCGGAGACCACAAGCGGCACGGCGGGCGCC\nnACCGAT  
CGGATGCCGTCCGCGGCCGGCTGGGCGGCCGCGCTCGGCGTCGCACTGCTGAGCACCGGCGCGC\nnGGGATGCCCGGCGC  
GACGATCCGGCGTAGCAAGCAGCCACGCGAGCGCCGTTCCGCCCATCGCCCCGGC\nnTATAATTTCAATCGTTCCCGA  
CGAGTATCCGCACAATGACGGACGCCAACCTTCCTCACGTGCTGGTAG\nnTCGACGACGATCCGGCAATCCGGAACCTGC  
TGTCGGCTATCTGTCCGACAACGACATCCGCGTGACGGT\nnAGCCGATTCCGGCCAGGAAATGGCGGGCGCGCTGGACA

CCTTTGCAATCGATCTCGTGATTCTCGACCTG\nCGCATGCCGGGAGAAGACGGCGTGCAGATCGCCCGGAGCTGCGC\  
 n>gi|142022655|gb|EQ086233.1|564 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\nCTGGTAGCCATGCAGCAAGGCTGGCGCTAGA  
 TGTACGGCCAGATTGGCGAGCTGGCCAATAACGCGGGAA\nCGCTGTGCAGATCAATCGACACTCTGAGCGAAAGCGCTC  
 AACAAGCCGACAGCGCCGCCATCCAGATGGC\nnGAGCGGGATGCAACGGGTTCGGCGCCAGATCCGGCATATTCCTCGT  
 AAGCCGCCTAAGTGGAGGACGCG\nnGTGCGGCGGAAGAACGGCCGGGGCCGGTGGTGGCGTGATCGAGCAGACCATCGCC  
 AGCATCTACGCCATC\nnGTCAACGTGATCAAGGACATCGCCGATCAGACCAGTCTGCTGGCGTTCAACGCCTGTATCGAG  
 GCGGCGC\nnGCGCTGGCGATCAGGGGAGCGGCCTCGCCGTGGTGGCCGACGTAGTGGCGAAATTGGCGGAGCGCACCGG\  
 nCCAATCCGCCGACGACATCGCTCGCACCGTCATATGCGCCAGTTAATGGCTACGGCGTCCGACGGCATGC\nnAGTAAGA  
 CGGCACCGCGCCCGCCAGATCGGCGCGACGCGATCGCTTCGATCCGCGCCAGTGCTCAACA\nnGGTACTGCCCTGGT  
 GCACAGATCAGCAGGCCTTGCGCGAGCAGAGCGCGCGCGCGTGCATCGCCG\nnGCCAGGTGAACCGAATCGCCGGT  
 GGCAGCCGCGAAAGCAGCAAGGTGCCCGGACACTGCCGAGGCGG\nnCGCGCTTTCTGGATCGGCTAGTGATGGAAAT  
 GAAGCAGATATGGCGCGCAAACCCCTCTGCATAAACA\nnCAAGGCCGCTTTCCGGCAAGGAAAGCGGGCCTTATGCC  
 GATCTGGATAGGGTCATGGGCCAGCGCCTGC\nnAGCGACTCGCGCATGATCTGGCGGTGCTCGTCCAATCTCGATCAAT  
 CGGCCTCGACGCCGATCTTGCGC\nnAACTGGATGGCGGCCGACATGCCGAAAGCCACCGCCACCAACAGAACCTCATG  
 CGCTACGGCCATGGT\nnGGGCTGCTCTTCATTGCATTGTTAAATCCAGGCGAACAACCTCGCCTGTATCTTGTGATTTC  
 CAGTGCC\nnACCGGCTGCAGGCTTTGCCCCAGGCAAGGCCAAGGACGCATCGCCCGTCTTCCGGCCGAATTGCGCCC\  
 nCGTGGGCGAAGCAGACGATGCGACTGCCGTGCGCGTTTACATAAGCATGCTTGCGCCAGGCCATCGGGG\nnCCCCGGCC  
 ATATGCGGACAGGCATTCCGATAGCCGTACAGCGTTCCGCCCTTCGCCACCACGAACAAGCTG\nnTCACGGCCGGCGCCG  
 TCGGGGTGCAAAACCGGGCGCGCCATCGGCGATTTTCATCGCGACGGCCAGTA\nnGAACGCGCGCGGCACGGCACTC  
 AGTGCTTGCCACCAGGCGGCGGGCCGACGGCGCCCACTTTTCGCGA\nnTGCTGGAACAGGAACAGTGCAGGCGATCGG  
 CGCCATCGGCGCTTCGCGCGCGGCCACTGCTCGTCGT\nnGGAGATCCATATCGGCGTCTGATTTCGATATGGCAGCCGA  
 GCGGGCTGTTGAAGTACCAGAACCAGTTGGA\nnGCCGAACCTGTGGCGGCGCGGCCCGGACGACTGGTAGCCCTTGT  
 TCACCAGACGCGTGCCGCGCTGC\nnATTACCTCGGTGGGGCCGCCCATGTGGAAGGTGAAGTGTTCGACGCCCTGCAT\  
 n>gi|142022655|gb|EQ086233.1|521 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequence\nCGTTGTTCGCCAGGTCGTCCGCATAGCCGGC  
 CGAGCTGAACTGCGTGACATACGGGCGAATCTTCTTCAG\nnCACGTCTAGGCAGCCTGGTAGTCGGCCGGGTCTTGTCT  
 CTCCGGGTCTTGGCCATGTACTGCAGCGCC\nnGCCGCGAACGCGCGTCCGGCGCGTTCGAGCAGCGAGATGCCGAGCC  
 CTTCAGCTTCGCCGCGTTCGCCG\nnGGTCGAACAGCAGCGCCAGCTGTCCGTCCGTGCGTTGTGCGCGAGCCGCTTCT  
 GACCGCCTCGACGTT\nnGTAGCCGATCCCGTCTGTCCCCACGCCCACGGCACACCGTACTGGTTGCCCGGTTCGGCCTT  
 CGCGACC\nnATCTTCATCAGCACCGGGTCGAGGTTTCGCGAGATTTCGGCAGCTTCGACTTGTGAGCTTCTGGTACACGC\  
 nCGGCCTGGATCTGCTGCGCGAGATAGTTTCGACGTCCGCACGACGATGTCTGATGCCGAGCTGCCCGCAAG\nnCACGCTTC  
 GTCTGCAGCGTATCG\nn>gi|142022655|gb|EQ086233.1|455 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nTTCGAGT  
 CGCTCGACGCGCCTCGACGGGCAACTAGCCGCCGTTGGCCAGCCTTGAACGGATCTCGTCCGT\nnCGAGAGGTTCGCGCAG  
 ATGCGTGGCAATCTGCCAAAGATTCCCCCACGGATCCTTCACCATCGCCCGCCTG\nnTCGCCGTACGGCATGTTCGGCCGG  
 GGCTTCGAGCGGCGTTCGCGTGTGCTGCAAGCGCCCGCGATAGGTCA\nnGGTCCGCATCCTCGACATACAGTACAGAAA  
 TGCCGGCATCGGCTCACGAGGCCGTCTTCCCCGCTGAT\nnCATCAACACGGAATCACCGATCCGGATTTCGCGAGGCAA  
 GCGCGGGCGGTATTGCGCCTGCGCATGAAAA\nnACCGTCCGATGAATCCGATCAGATTGTGCGGATCCGGCACGACGAT  
 CCGGGGCGTGACGGTGTGCCATC\nnCATCGGGCACGAACCTGGTTCATGTCCGGCAGCCTCCGAAGAGATTTCCGGGGGTA  
 TTGTGCGCCGGGCGC\nnCCTCTGTGCGCAAAGGGGGCGACGCAACGCGGCAAGCGCCGTCAATCTCCGGGCGATCCACGC  
 ACCGCGG\nnGCGCGAGTCGCCATCCCCCGCATCGAGCGACCGCGCCACGCGAAACCGAGATCGTCGATTGCAAAGGT\  
 nCGGATGACTGCGACGGCGATTTCGTGCAAGGACGCCCGTTTCGACATCGGCCAGCCCCCG\nn>gi|142022655|gb  
 |EQ086233.1|229 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold,  
 whole genome shotgun sequence\nGGGTTTCGGCGTCTACTGGGGCGGCTACTGCAACCACGGCAACGTGCTG  
 TTCCCCACGTGGCACGGGCT\nnTACCTGCTGCGGCTGGAACAGGCGCTGCAGAGCATCCCCGGTTGCGAGCAGGTGATG  
 CTGCCGTTCTGGG\nnACGAAACCAGCGAGGAGTCGCTGACGAAAGGCATTCCGTGGGCGCTGACCGCCCCGGATTTCGAG

CTCGA\nCGGCGAAAAGATTCCGAACCCGCTGGCCTCCTTACGTTCAACCGGGCGATCACCGACAACATCAATGGC\nG  
ACAATCCGAACACTACAGCAAGCCGCTGCCTTACGTACGGTTCGCTACCCGTTGTGCGGGCTCGTGGGA\nnCCGACAAGG  
ACAAGGCCGCCACCGAGAAGCACAAACGCCAGGTTCCAGGACGCCAAGGAGAACGTGTCAT\nnACTGAACCGCAACATCG  
TGGACTGGCTCACGTCTACATCGTCTGTGAAGGGCGACGTGGTGGCGACCAAC\nnGTGAAGAGCAAGTTCGAGCAGTGCC  
TGGACGCGCCCAACTACACGGTGTTTTCGAACACGTGTCGCCGG\nnCCCAGTGGAACGAGAATCTTCCGGACGGCGCGA  
TCCCCGTGGTGCCGCTCGAGTCGCCGACAACGATAT\nnTCACCTTGGCTCGGCGGATTTCGACGTGCCGACGGACCGA  
ACAAGGGCGATTTTCGCGCCGATTCCGGGC\nnGCCAACGGCGACATGGGCGAAAACGACACGGCGGGCTCGACCCGATCT  
TCTATTTCCACCACTGCTTCA\nnTCGATCGCGTGTTCTGGCTGTGGCAGAAGCGGCATGGTTACACGCAGCACCTCGACG  
TCATCGCCGAGTA\nnTCCGGGCACCAACTCCGTGATGCGCAGGGGCCGACGCCGGGCGTCTGCCAACTCGTGCTCA  
CGCTC\nnGAAACGCCGCTCGATCCGTTCAAGAAGAGCGAGAACGGCAAGGAGCGCGCTACACGTGCTCGACTGCA\nnT  
CAACATCGAGGAACAGCTCGGCTACACGTACGGCCCCGGTTCGTGGAAAACCTGCCGAAGTTGTGCT\nnCGCGGCACC  
GACCGTGCCGGCCGGCAACAGCCGAAGGTGCTGCGCGTGTCGGGGCTGAATCGTGCGCCG\nnATCGCCGGGTGCTTCCT  
CGTTTCCGCGTACGTCAACGTGGACGGCGAACGGCAACTGCTCGGCACGGAGG\nnCCGTGCTGAGCCGTGGAGCGTGCA  
GTCCTGCGCGAACTGCCAGACGCACCTCGAAGTGAAGGCGTTCTT\nnCCCCTGAATCACTTCGCCGAAAGCGCGGTGCA  
GGATGCGCAGTACGACGTGAAATCCATACGCGCGAC\nnGGCGTGCGCCTGCAGACGCCACCCGCGTCGCCGCCGACTC  
GCAAGCCGTGCTGCGCGCGCGCGCCGAAC\nnCCGTACGAAACTGTTCCGCTCGAGGTCCGCTAGCGCGCGCGTGAC  
GAACCGGCACCGTCGCCCGGC\nnGCCGGTTACCTCGGTACTTCGACAACGTGACATGACGGGAGGCGCGGGCATGAGC  
TACACACCCGAAA\nnCCGGCAGCCTGGTGGGGCTGTGGACCTATCGAGTGTCTGAAACGAGCCCGACCCGGCCACCGCA  
TTCAA\nnCGACCTTGAATTGGCGCTCGGCACGATCGAAATCGCGCAGGCGCGCGGGCGGCATTTTCAAGGGTCGCAATT\nnT  
TCGGTCCGGGATGGGAGCTGCAGCTCAACGGCTGGATCAGCTACGGGAATCCGGGGACGGTGCGATTCC\nnAGGGGCGCG  
GTGTGGTCAGCGGTGAGGAATGGGTATACGACTATGTGCGGTACGTGAGCGCGCCCTGGCC\nnCAACGGCATCGACCAGC  
GCCCCGCGCTGACCGGTTTCGATCGTCCGTACGGTGCCGATGCGAGCGGCAGC\nnGGCGCGCTCGCGCCGGCCGGGGTG  
TGTGCTCGTGGTATGCGGTTCATGCGAGATCCCGCTGAGCGGCCA\nnTGACTGCGCCGACTGCACCGCCGGCCGACACGG  
CCAGGCGTCCGGCGACGGTGACGGACGCGATTTCGCGC\nnAGGCGTTGGGATCGGGCCTGCTCGAATGGCTGTTATTTGCG  
TATCCGAACGGATATGCGGCGCGATGGCA\nnTTCGCACTCCTCGGCTACCTGGCTTGCGACTTGCCGACCCGGCTGCACG  
GTGCTCCTTCCAGCTGGATAT\nnGGAACGGCATCGGCGCAATGGTTTTCGGCATCGCCGGCTATTCGATGCAAGGCGGCC  
TGTCGGCGGGTG\nnCGCGCGCAGGGCAATGCTCCTGTGCACGGGGCCATCGGCGCGGATTGACGCTCGATGCGATGC  
GGTCC\nnCCCCTCGATGCCATTCTCGACATCTGCGGCGGAACGCTCGATGCCCGCGCCATGTGGAACACACTCGTGCTC\nnT  
GCATCTCCAATGGTTCCCGATGTGATGCTGGCCATGCTGGCGCTGCTGATCTTGCGCGAGGCCGACCG\nnCAAGGACCG  
TCGCCGACGCGGCGCGCAAGCATGCTCGTTGTGGTGCCGTGCGCGTCGCGCTCGAATTC\nnGCGGGGATGCAGCTCGT  
CATGGCGCTCGGCATGACCGCGATCAGGGCATCGGCGTTTCGCTGCCGGGTGC\nnGTTGGGATACGAGCGCGGTGGCAGT  
GTCGATGCTCGCCAGCATGCTGGCGTTTCGACGATTGAGCGATCG\nnGGCGGCACGCGCAGGGGGCGTTGAATGGTTCGA  
TTGCGAAATAAACACGGCCATATTGAAGAGGCGCGGC\nnCATCCGTTGTCAGACGAGGCCGCGTTCAAGCTGCTCTGGAA  
AAGTGTGATGACGCGGTTTTCGCTCGGCT\nnGCCGCGCGCTTTGGCAGCGGCAGCCGAGTTTAAACCCGCGTCACACGG  
ACGCTGCACGGAATCCAGGCG\nnGCGCTCGTCACTATTACGATGCTCGTGCTCGTGCTCGTGCTCGCGCCACGGATCAC  
GATCGCGCGGATG\nnACGATGCCTCCACTCGTTCGCTTCCAGTAGCGATGGCCGTCCAGTAGCGATCGCCGTGCGATGC  
CGATC\nnTCGACCGATACACCGGCCCTGCGGCGGGCTCTCGCTCTTCGATCTCGGTTGCGGTGAACC\nn>gi|14202  
2655|gb|EQ086233.1|422 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
scaffold, whole genome shotgun sequence\nAGCGCGTAATAGAACTGCCGCTAGGTGCGGTTGTAATAG  
CGGATGCGCGGCCGATCGCGCCGTGCAGGT\nnCGTGACGCCGCTCGATCGGCAGGTTCCGGCATCATGTGGTGCTCCACG  
TGATAGTTGTTGCCGTTCTGTAA\nnCCACGTCATGAACGCATTGCTCTCGATCGTCCGCGTGTTCGGAAACGGGTGCGTG  
CTCGTCAGGTCGCG\nnCGGTAATGCTCGGGCATCTCGACAGCGCATGCACGGGCGCGACCACGAACAGAGCGGCACG  
ACCCACA\nnCCCACACGGCGAACCACTATGCAGCGCGATCGACACGCCGGCCAGCAGCGCATCGCCGCGACCATCAG  
nCAGGTGATCGCGGCGGATCGCACGCGACACGAGCGGGTTGTGCGCGAAGCTTGCGCCCGGCACGACGAGC\nnTTCGCGA  
CGTCTTCGCGAACTGGATATAGTGCGCGCCATCGACAGCCGCCAGGCCACAGGCCGAGAT\nnTGACGACAGGGCTCG  
CGCCGTAAGTGGTTCGCGTAGTTCGAAGAACTCCTTGTTCGCGCGTGCCGAGCAG\nnCCGGTGGTGGCGCAGGTGGCTGT  
CCTGGTACGCATGGAACGACACGAGCATCGGCATGCCGAGCAGCAG\nnCCGAACACCATGTTGAGCCGCTTGCTCCTGA

ATCCCTGGTAATGCAGCGCCTGGTGTTCAGCTCGACCC\nnCATGCGCGTACATCGCGCCGATCAGCAGGATGCCGACGA  
TCTTCAGGACGACATACGTTCGACAGCGCGAA\nnCAGCATGCCGTGCGCGATGAGCGCGACGTAGACCACGAGCTTCAGCG  
AGAAGACGACGCCGCTCGGGCGG\nnCTCTTCATGTCTGAAACCGGTTTCTGACGGGCGAGAAGCGATGAGGTGACGGACA  
AAATGCTCTCCTTGG\nnCCGGACCCTGCCCGGCACGATGACGCGGCGACACTGGCGCGCTGCCGGTTGCAATTCTTCTG  
TGAGTCG\nnATGCTTCGGACAAACGAATTTTCCGAACGCATTGCATGACGCCCCGCGCATGAGCACACCACATCGTGAAC\  
nTAGACGACCGGTGCAATGCGTACCGACCATTACCTCGCGCGGGCTCGCACGTATACGCACATGGCAGTG\nnCCGAGTG  
GATGCGTGCTGCTCATGCATAGCCTGCAAATATTCCGTTTGTGCGGTCATGTCGATCCGGCAA\nnGAATCCCCCTCACTC  
GACGAACGGGCGCGCGGCACCGTGGCGCGAGCGGCCGTGTTCTCAAGGGGAGA\nnATCCTCGTGACGAAGTGCGCATC  
CACCATTTCATCGCCTCATGCAATCGCTCGAGGACGTGCGGGCGCGCA\nnTGCATCGCATGACGTTCTGTCGACGACGCCGG  
CAACGAGTCGAGCATCACGTACCGGCGCTTCGCCGAAGA\nnGGTGTTCGCCAGGCCGGCGCACTGCGCGAACTCGATGT  
GCGCGAGAACGACCTGGTGATGTGGCGTTG\nnCCCGCATCGGTGAGCACGCGGCGCGCATGATGGCTGCGTGATGAC  
GGGCGCACTGCCCTGCACCGTGC\nnCCGTTCCCGCGCGCGCGCCACCACCGGGCGGAGGTGATCGACGTGCGATGCGA  
GCTGTACCGGCCGCG\nnCCCTCGTCATCGCCGCGGATGCGCAGTCCGCCGCGTGGCGCGACGACGCGTTTCCCGCGTCTGTC  
CACGCGC\nnGTGCTCGATCTCGACACGCTGTGCTCGTCTCGCCGATGCCGTTGCGCGCGCTGATCAGCGCAAGCATG\  
nGCCGCGATCCGCATCACGTCCAGTCCAGTCCGGCTCGACGTGCGATCCGA\nn>gi|142022655|gb|EQ086233.  
1|384 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
shotgun sequence\nnCTTCCGATCGCATGTACGCCGGCGCGACATTTGACGAACGCGTCGAGCGACACGTGAAAC  
ATGGGCAC\nnGTCCGAGCGCTGTGAGCGCCCTCGCCAGGTGCTCTTCCCTGCACTGGTCGGGCCGTGCAATACGACA  
\nGCAACGGGTTTCATGGCGATGGTCCGCGGGAGAGATTGCCGTTGTCCGGGTGACGCGCTAGATCAGCTC\nnGGCGAC  
GCGCACGATGACGAGCGCGGCCGGAACGAGCACCGCTGCGCACACCCGTGCCGAGCACGCGT\nnGCGCCGGCGAGTGT  
CGTGATGGCCCGCGAAAGCTGTTCTCCGAGATCCGCCCTTCGATGACGTGCTCGG\nnTCATGACGAAACCTGCGGGTC  
GATGACGACCGCCAGCACGACGTCGCGAAACCGTTGATGACGGACGA\nnCAGGTTGCGCATGTCACTCTCAGCTCGGG  
CTTCAGGTAGCCCGGTACAGCGACGGAATACGCCGACC\nnGTCCAGATCGCCATCGCGCGACGTTCAACGCAATGAC  
GCGCCACGACACGCCGTTGTTGCGGTCAGTT\nnGCGTGACGTTCTGGCTCGAAGGCAGTCGCGCGCCGACGCGCAGATA  
GTTGATCCCGGCGCGACTGCACGC\nnATGCAGCAGCAGGCGGGGAATCGAGCGGTTACCTGGAAGTGTTCGACGGCGCG  
ACTGAAATAGCGCTGA\nnAATGTCGGGATCAGGAACGACCCGACGATGGTCGCCACGGTGGCGGACAGCAGGAACACCG  
GAAGTCGG\nnCCAGCAACCCGCTCCCCATCCGGTGCGCGATGTCCATTTCAACGCGTTTCGCGATGAACGGCCCCCTGAAA  
\nGGAATTGGCCGTTCCGGACACCAGCGCGATGATCCCGAACAATGAAAACGACAACGCGATGCGACGCGTG\nnCGCACG  
CCGGCGATCCGCACGGCATAACGCCAGTGTGCGGATCATATGGATGACGAAGGTAAGCCCGCAGA\nnTGATCCAAAGCTGA  
CTGTCCATGCTGGTGCCGCGTGGTGACGGGATCACGAGTGATCCCGAGCGCGGCG\nnCGGCGCATCGGCTGGCGAAGC  
TGCTTCGCAAGCGTCACGGTTGGGCGACGAGCGTCAGGCGCTACCGT\nnGCGGCGCAATCCCCGTGCTTCATCCAGCA  
TCCATGTGCGGAAGGCCTGCAACGGCCCCGTTGGCCAATC\nnGATCGGTTGCGGCGACGAGGCAGAAGTTTTTCGTCA  
CCGTCTTTCCCGCCGGCCACGGTGCACACG\nnCGGCCCTGCTCGAGCTCCGCATCGATGTACAGGCGCGCACACGCG  
CGACACCGAGGCCGGCCAGCGCCG\nnCTTCGATCAGCATCGAATGAAGGTCGTAGCGCGCGCCACCGCGAGTTGGTCA  
GCACGATGCCGCACGC\nnCTGCGCATAAGCCTGCCATGCGTCCGGGTTCTGTGCGCGGTGAAGACGCGGCAATGCATCCA  
GCGACGGG\nnTTCTTCCCGCGCGCGCGGAGCAGCGCCGACTGCAGACGGGACACGCACTTCCTCCAGCAGCGGATGCA  
\nGATGCATCCCCACCCAGGCCGATGCTCGAAATGAATCGCTGCGTCGAAACCGCTCCCGGCCAGCAGGAA\nnCGGCTC  
CATCCGCTCGGCGATATGCACGGTGATGTTCCGATGCCGTCCTGAAAATGTTTCAGCCGGGGG\nnATCAACCAGCGGGT  
GGCGAAGGTGGGATCGCGGCGATATCGAGGCTCGCGCTTCGATCGGCTGCCCCA\nnTCAGGTACAGGCTGTCCCGTTC  
GAGCCGATCCAGAATCTCGCGAACCTGCACCGCATAGCGCGACCGTT\nnCGGCGCCAGCCGACCCGTTGCCGATTGCG  
CTCGAAGAGCGTGACGCCCAGGAACGCCTCCAGCCGGCCG\nnATCTGACGGCTCACGGCACCCCTCGGTGCGTGCCAGCTC  
CTCGGCGGCTCGGGCAAAGCTGCCGTGTCTGG\nnCCGCGGCTTCGAACACCTGGAGTGCGGAATTGCTTGAATCTTGCG  
GGACATGGAAGTCTCGTGGGTGCG\nnGATATGTTGCCATCAGCTTGAGTAAATGTCACTGAAGGCTGACTTTATTTCTG  
TTTATGGCGCGGAATG\nnGCGAGCTTAGCATTACGTTAGCGCACTGAACATCACATGTTTCGCGGATGGATGCGCACCCCTG  
CCCAAAGG\nnACAAACCCATGATCTACACCGTGAATGCAGCTTCGCCGACCTCGACAGCGAAGCCGAATGGAACGACTT  
\nCTACAGCCTTGAAAAGCTGCCCGCCCTGATCTCGGTACAGGGGTTTCACAGTCGACGCGTTCAAGGCG\nnATCACC  
CGCGGTTGCCCGGTCTATCTGGCGATCCACACGATCGACGGCCTCGATGTCTGACCGGCGACG\nnAATATCGCCGAAA

GGCGGCGGCAATTTTCGCGAAATGGCAACGGCACATCACCGACTGGCACCGGAATCT\nCTACAGCGATATCGGCTTCGCG  
CCGGCGGTAAAGGACGGTGAACATCTCGACTGAGCGCCGATGGCCCG\nnGATTTCGCTGCTTCGATTGGGCCTCGAACCG  
CTGGCCATGCAGGCCGTTGCGCTGGAGAAGTTTCCGGCGC\nnGCCGTTGGCTCGCCGTCGTGCCACGGAATAGCGCGCG  
CTTGTGAGGGCGTCCCCGAAGGTATTCACCT\nnTTACGCACCGATGACGGAGCAGTTGACGAGCACACACGCCCTTTTCG  
ATTGCGCAGGAGTAGACCGCGATG\nnCCGAACGTCACGATTTTCGGTCGACGCGAAGCGGATGCCGGTGATGAAAGCCTT  
GCCGAGCTGTTCGCGCG\nnACTGCGTCGAGCTCTGCACGCAGGTGCTTGAGGCTGAACTCAAGAACGTCCACGTCATTTTC  
CTGGCGGT\nnTCGGCACGGGCATGGACGTCCCGTCTTCGCCGACATCCGGTATCGCGTCGGCACGCGTCGCACGCCGGAG  
\nGTCATGAACCGGTTTCATGGAAGCGCTGGATCAGGCGATCGTTCCGCCGACCGGCCTCACGGCGCGCATTC\nnGGTGTT  
TCGGTTACACCGCGTCGAACCTTCACGCCCGCAATTAGCCGATTTCGGAGCAACACGATGTCCAC\nnCCTTACCTTTCCCG  
TTCAGCAAGTCGGCGACTTCACGATTACCGCGATCAGCGATGGATACCTCACTGCG\nnAGCCTCGACTTCCTGTGAATA  
TCGATTTCGGATGACGCATCAAAAATGCAGCGTGATGCGGGGCAGAAAG\nnAGCCGGCGGCCGTTTCATATCAATTGCTACG  
TCGTACGCGGAGCGGGCCACACCGTGCTCATCGACGGCG\nnGGCTGGCGGGTTCAAGCAATGGGGCGGCCAACTCCGGA  
CCAACCTGGCGCTTGCCGGTATCGAACCCGCT\nnGCGATCGACACCATCTTGCTTACCCACGCGCATCCCGATCACGTTCG  
GCGGGCTGGTGAATGATGCAGGAC\nnAGGTGCGTTTTCCACATGCCGAGCTGGTCGTCCATCAGCAAGAGGTCAAGTTCT  
GGCAGGACGACGAAA\nnTCTCGTTCGCGCCAGCGAGCGGGCCCGCGCAACTTTGCGAAGGCGCGTCAGGTGTTTCGACG  
CTTATGTC\nnGACAGGCTTCGCATGTTTCGACGATGGACAAGTG\n>gi|142022655|gb|EQ086233.1|280  
marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequence\nnGCAGCGACGAGCCGCCAGATCTTGATCCCGCACGCGCAGCTCGCCGGCCCGACGTCGACCGTGACGCT  
\nGAGCGGCTCCATCGCCGATGCGCGAGATACGGAATTCGTAATCTGCCAATGCGCGTGGCCGCCGGC\nnGGCGCG  
CTGCCGGCGTCGGCCGTGCGCACACCGTCCCCGGCGTGCCGGCGAGCTGCTTGACGCCTGCA\nnTCTGTTTCGGATGAC  
GTGACGGTCGAGCCGGCATCCTTCCACGTGATCTTCAGCGCGTCGCGGCCGAGCTT\nnGGCCGGCCAGTAGCCGTTTCGCG  
ATCACGGCCACGCCGTGCCGCCGCGATCGGTTCGGATCTCGACGACC\nnTCGACCACGCCCTTGACGGCACGGGCCGCC  
GCCGCGTCGTAGCTCGCGACCTTGCCGCCGAAGCGCGCG\nnGGCGCGGACACGGCCACCATCGTGTCCGGCAGACGC  
CAGTCCATCCCGTACTTGAGCGTGCCGTGAG\nnCATCCCGCGCGCTCGATGCGCGGCGTCGGCTTGCCGATGATGCGG  
AACTGGTCCGGCGACTTCAACGAC\nnACCTGCTGCGGCACCGGCATCGCCATCGCGGCCGGTGCGAGTTTCGCCGTAGGTC  
GCGCGGTAGTTGCCGG\nnAGGTCACGACACCGTTTCGCCGTCCGGCAGTGGCGGGATCGACGTTCCAGCGCTGCGCGGCC  
GCCGCGAC\nnGAGCATCGCGCGGGCGGATGCGCCGAGTTTCGCGGTATTGCGCGAACGAGTGATTGACCGACGTCGAACCG  
\nGCCGTTCATCTGGATGCCCGTCACCGGATCCTTGTAAGGCTCGCCGGCCGGCGCGAGCACGGTGCGCACGT\nnTGCGCC  
AGTGGCGATCGAGTTTCGTGCGCCAGCGCCATCGGCAGCGCGGTGCTGACGCCCTGGCCGGACTC\nnCGTCCGGTTGACGG  
CAACGGTCACGGTGTGTCGGCGCGATGATCAGGAATGCTTGCGGAGGCGAAGCC\nnTGAAACGCTTTGCCGGCTTCGG  
CCGTGGCCCGGTCGGCGGATGCCAGCGCAAGGCAGCCGCTCGCCG\nnCCATCGTGAACCTCAGGAAGCCACGCCGGC  
CGAGCTTCGGAGAACCTGTTCCGGCGACCGGTGAGGCGTC\nnGGAGTGGGGGGTGTCGGTGCTCGCCGGCGCCCGGCTG  
TGCGAATACGCATGGTCAATGTCTCCCGGCAG\nnGCTCAGGCAAGCGCGCTGGCCGCATCGTGAATCGCGCGCGGATGC  
GCTGGTAGGTGCGCAACGGCACA\nnGGTTGCCGGCCATCGCGCGATCGATGTGCGCGTCGGTCGGCGCCTTGTTGCCGT  
GCAGCAGGCCGGCTGC\nnGCTCATCATCTGGCCGCTCTGGCAGTAGCCGATTGCGCGACGTCGTGCCGGATCCAGGCGT  
CGAGGACG\nnGCCTTGCCGACCGCATCGTCGGCGAGATGTTTCGACGTCGTGATGTGCGCGCTCCCGACCGCCGACAGCG  
\nCCAGCGTGACGCTTCGGGCCGGCCGGCCGTGCAGATGAACGGTGACGGCGCCGAGTCGCCGGTGCCGCA\nnTCCGAA  
CTTGGTGCCGGTCAGCCCGAGGTTGTCGCGCAGCACCCAGAGCACCGGTGTCGAGGGATCGGCG\nnTCGATCTCGGCCGG  
CCGGTTGTTGATGTTCAACGTCACCATCGTGAAGTTCTCCATTGGACACCTCGTCA\nnTGTTTCGCGAGGCCAAAGCGTA  
ATGCGGCGGACGTGAATCAATTAAGCAGATTTCGAGAAAAACCCAGCA\nnTCTGCAAATTCGATTTTTTTTGCTTTCTTA  
ATGATTTTCGGCCGAAGATTAATAATAAAATTTACGTTT\nnTCCGTACAAATCATCGAATCTTTTTGAAAGATTGTTGA  
GTTTTGACAAGATCTTTGCCGACCCGTGGCAA\nnACGTTTGCCAGTCGAGCCTATATTTCACTCGCGATCAATGCGGAAG  
GCGTGCAAATCGGATAACAGAAAC\nnCGAAGATTCAATATCCCGGAAATCTGTACCTACCGGGTTGACCGGATGATTGG  
CCTGTTTCGTTTACCGG\nnGCGGTGATCCCCGTCCCGTCTGGAGCAACGGGACGGGCGGTGCTTCGGACGGAGACATCGA  
GACACGGG\nnCGGGCATGGGTGCGCGACGGTAATTTCAATCGTCCGATACAGCCGCTCTAAATGAAAAATGGGAAAAA  
\nTTACCGGGCAGGGGTTCAAAAGTAAAAAATGCCAATTAATAATTCTCGTTGATTAATTCTCGTTTGT\nnGCCAAT  
TCAGCCATCCGCTCGTGTTTGGTGACGGGAAGATATTTTCAAAAAATATCGAATAACGCATTC\nnCGGTATTTTCTTC



GGCATTCCGATGCGCGAAAGGGAGTTTCGGCGGCTTCGATTCTCTTTATGGGAAGG\nTGTGATTATGCGTATGCGCATG  
GGCTCGAACAAAAAGCGCGCGAAAGCACGTCCGGCATTCTCGTCCCG\nACCGGCATGCTGGTCGCGCTGGTCGGCG\n  
>gi|142022655|gb|EQ086233.1|158 marine metagenome JCVI\_SCAF\_1096627390048  
genomic scaffold, whole genome shotgun sequence\nGCCCCGTTTTACGCGGAGATCGGGCGATTG  
CATCACCTGAAGGCATCGTGGTCGGTCGACGGGATGCTT\nGGGTGCGTGCAGCGCGAGCCATTCTGAAATTGCGATTTT  
TCGGTTGGATGAAGCAGTCTTCGAAACAGTG\nCCGGTATGAATGCGAGCAAGAGGGCGATTTCGATGCCTCGTCGCGATC  
TCGTATTCGTGCAAACGACAAGG\nAGTCGATCGTGAGTGTGCGAAATCAAGGGAATCCGAAAAATCGCGCTGCTTTATAC  
CGGGGCCGTGGTCTG\nGACGCCAACCTACCTGCTCCCATTCAGCATCGAGGAATCGATGGCCCCGGTTCCGGCTGTCCGA  
AAGCACC\nGCTGGCTGGCTCGCATCCGCGGTGCTGCTCTGTCTTTTCGCTTTCCATCATCGTTTTCGGTGCTCACACGG\  
nCGACGTTGAACAAGAGAACGGGCGCGCTCGCTGCGGCAGCGTTCCGCCGTTCTGTGACGTGTGCGATGTT\nTTTACAT  
AATTTCCATGTATTCTGTCGCCGGAAGCTGATTCTCGGCGCCGCGTTGGGCGTCAGTTGCGTG\nTGCGCGTATGGGGTG  
ATTGCACACGTCAAGCACCCGGAGAAGGTCGCGGCGCAAGTAGCCGTTGCGATGG\nCCATCATATTTTCTGTGCGATG  
TACGTGCTACCGATCGTGAATGTGAAGATGGGGCATATCGGGGTGGA\nTGCCGTCCAGCTTGCCGTGGTCTGTGGCGGCG  
CTGTTTTTCGGCGCGTTTCATGCATCCTGCCGTCAACACA\nAGCGAGGAAGCGGTGCGAAGGACGCCATCAAATCCAGT  
GACGTCCGAGGGATTCTGATCAGCGCATTCT\nTCATCTACGTCTCGCAGACGGCCCTGATGGGATTGCGCGTGAGGTC  
GCTGCAACCCGGGGCGTCAAGC\nGGAGCAACTCGGCATGCTGTTTCATGCTCAATGCGGCCCTCCAGTTGCCGGTGGGG  
GTGGCGGTCAACTGG\nCTGGGCGATCGTTTTGGGCTGTTCAAGCCGATCGCCCTTGGAATCGCGCTGTTGATCGCATGC  
AGCCTGG\nGCATGTATTGCGTGGGCGGAAGTGGGCGTTTTCTCGTGTGACGGCGCTCGTCAGTGCCGGCGCCACGCT\  
nGGTGGCTCCTTATATGTTGGCGGCACTTTTCGAAATGGATGCCAGCGGAAGAAGCACCGCGACCTGTGCG\nTCGGCAA  
TCAATCTTGGCCTGACGTTGGGCCCCGCGATTGCCGGCACGTTGTCAGCGTGGCTGGTTTGC\nGTGCGGTTGGCTGGC  
TCTCCGTGCGGGTTGCTGGTGGTGCCGATCTTCTGACCAGCGTGGCGATCCGGCA\nGTTCAAGGCAAAACATGCCGAAG  
CCGCTGTGCGCTGACGTGCGCGGGAAGCGATCAACTGAGCGCGGCC\nGCACTGACCGCACCGGTACCAGGGCGGCGCC  
AGCATGCCTGACATGCCATTGCGGAGCAGACATCGCTCG\nGAAGAGCCCGCTCGCGTCGACAGGCAGCGCAAAAACTG  
CCGTGAACGTGGATGACGAACGCAGCGGCAG\nACAAATATCTCGACGATCATGGATTTCTGCATGGAATCAAATTCTA  
TCTTTATTGCCTGGCTGATCT\n>gi|142022655|gb|EQ086233.1|59 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCGTGAAT  
TCGATCTGCGGCAGCACGATCGCATGCATCACGCCCCGCCAGTCGATCTGCACGAAGCCGAGT\nTCGAGCACATAGGCG  
ATCAGGCCGAGCGTCATCCATTTGAGCCAGCGCGCATAGCGCTCGTACGACAGCG\nTCATCTGCAGCGGAGCGACGCG  
AGCCCGTAGAACGGCGCATACCAGTAGCTGCGCCCCGCCGCGACGAG\nCGCGGTGCGTTCGCCCATCGCCGCGAGATCG  
GCGCCGACGTTGACGACGTTGACGATCGCGACCATCACG\nATCAGCGTCTTGGCGAGCCGGTCGGAGAAGTGCTCGCGG  
ATGGCGGCGACGAGGCCCGCGCGTTGCGGC\nGGCCGATGAACGCCGACGCGAGCTGCACGGCGACCATCATCGGCAGC  
GTGACGACGATCATCCACAGCAT\nGTTGAAGCCGAACCTGGCTACCGGCTGCGCGTAGACCGCGATGTTGCCCGGATCG  
GCATCGGACGCCCC\nGCGAGCAGGCCGGGCGGAGCCGCCGTACACGGCGTTCGCGGTGCGCGCGCGGCCCGCG  
GCCGGCT\nGCGCGTCTGCGCCGCGGTGTTTCATCGGGTGTCTTCCCTTGCGCGTGCCGGCTGGCGGATCGCGGTGCG\  
nCGGTGCGGTTGCGCGGCCAGGCCGATGCGGTGCGCAATGCGTCGACGAGCGTCCGGTACGCGTGTCTCGC\nGCGAATC  
GGCGCCGCGCGTTCGCGACGACGTACGACGGGTGATAGGTGGCGACGACGAGCCGGCCTTCCGC\nCGTGCGAATCGGTGC  
GCGCGCGGCTCCAGCGTCGCATCGTTGTCTTGCAGCAGCGCGCGCAACGCGGTG\nCCGCCGAGCGCGACGATGACGCG  
TGGCCGTACCGCGTCGAGCTCGCGCTCGAGCCAGTAGCGGCACGCGG\nCCACCTCGCGTTGCGCGGCGCTCTTGTGCAG  
CCGCCGCTTGGCGCGCGGCTCCCATTTGAAATGCTTGAC\nCGCGTTTCGTGAGATAGCACTGCGCGCGCTCGATCCCGGC  
TTCATCGAGTGCACGTTCCAGCATGCGCCCG\nGCCGCGCGGACGAACGGCAATCCGTTGCGGTCTCTGGTTCGCCCCG  
CTGCTCGCCGACCAGCATGATCG\nTCGCGTCGACGGGCGCGCGCCGGGCACCGGCTGGGTGCGATGCTCCACAGCGC  
GCAGCGGCGGCACGC\nGTTCGAGTGTGCGCGGTGCATCGTCGCGAGCGGGCGGCGTCCGGCGCGTCATGTGACGACGCCC  
GTGCGGC\nACGTGGCGCGGCAGGTGCGCCATGCAAGCGGGCGGAACGGAAAGGAAAGCGCAGCGTACGCGACATGAA\  
nTGCGTCCGGTACGACGATTCGACCTGTGCGCAACCGCCGTTCCGCATCGGCCGAGACGGGCGCGTGGTT\nTGCGTGG  
AGGCGGAGGTCAACGACAGGAGAAACGCGATGCTTCCCGATTCCGCAACGATCCGGCCGCGG\nACGACGTGCTGCTAT  
GGCGCGACCCGACAGCGTCCGGCGCGGTGCGATCGCGTGTGATCGTCGACGA\nCTATCGCGACGCGGCCGATGCGC  
TGCGCCTGCTGCTCGAGGCGCGCGGCTTCGAATGCCGCGTTGCCGTC\nGATCCGTTCCCGCTGCGCGATCGCGCGCG

ACTGGCAGCCGTTTCGCGGTCTGCTCGATATCGCGATGC\nCGGGCCTCGACGGGCTGCAGCTTGGGCGCGGGCTGCGTG  
GCGATCCGCACACCGGCGACATGCTGCTCGT\nGGCGTGCAGCGGCTTCGCGTTCGCGACGCGACTGCGAGGCGGCGAAGC  
AAGCCGGGTTTCGACGCGCACTGT\nGCGAAGCCGCTGACGCCGACCGGCTGCTCGGCTACCTGGATGCGGCGAGCGGCC  
GCGCGCAGCGGGACG\nGGGAGGTTACACGGCGCTTGTAACCGCCGGGGCGGGGTGGCGCAGGCATCAACCGTCCG  
GTGCGCG\nnTGGCGGCGCGCGCGGGCGACCTGCCCCGCGACGACGGCCAGGCCCGCGCACGCGAGCAGTGCCGCGCT\  
nCATCGGAAGGTTCGCGCGGTAGCCGTGCAGGTGCAACAGCATGCCCGGATCGTTCGCGCCGAGCCCGATC\nnGCGAGTT  
GCACGACCGCCACCATCAGGCCGCCCGGCTTCCGCATCGTTCGCGCAGCGTACGGGCGAGCC\nnACGTCCACCAGCCGA  
CCGGCGCCGCGGTGCCGAGCAGGCCCATGCGCCGAGCAGGCAGGCGGTTGCCGC\nnCGTGGAGCCGCCGAAGAGCGTCA  
GCGCTACCGCAATCGCGGCCATCAGTGCCGGCGTGCCACGAGTGTC\nnCGGTAGAGCCCGTGTTCAGGACGCCGCCGA  
TCAGGACCGTTCCCGCGAAGCCCGCGACGCCGATCACGA\nnGCAGGATCAGCGACAGTGTGACGGTGCTGCGTGCGTGA  
CGTTTCGAGGAAGGGCGCAGGTAGGTGAA\nnCAGCGTGAAGTGGCCATGAAGAACCGCTGACCGCCGCGATGCCGA  
GCGCGACCGGGCGGCCGGGCCAGC\nnAGCCTGAGCGCGTTGGCCGACGTGCCGCGCGGCGGCGACTGCATCGACGGCAGGC  
TGACAAGCTTCCAGC\nnCGAATGCGATCGCGGCGACCGGCACGACGCGAGAAGAACGCCCGAGCGCCACCCGACGATCCCGC  
CCAGAAA\nnGCTGCCGAGCGGAGCCGCGACGACGGTCGCCAGCGCATTGCCACCATTACGATCGCCAGCGCGCGCGGC\  
nACGTGCGTTGCCGGCACGAGGCGCATCGCGGTGGCGGCCGACATCGACCAGAAGCCGCCGATCGCCACGC\nnCGATCAG  
GGCGCGGCCGGCGATGAACGCCCGTAGTTCCGGCGCAATGCGACGAGCGTGCCGGACACGAT\nnCGTCAGCAGTGTCAA  
CACCAGCAGCAACGGCTTGCGATCGCGCCGGCCGGCGAGCGATGCGATGAACAGG\nnCTCGTGACCAGCGCGAAGCGGCC  
GGATACGGAGATGGCCTGGCCGGCTTGGCCTTCGCTGACGTGCAAGT\nnCGCCTGCAATCGGCGTGAGCAGGCTGACCGG  
CATGAACTCGGAGGCGACCGCGGAACGCGCCGAGCGA\nnCATCGCGAGGACGGCGCCCCACGCGGCCGACCGTCAGG  
TGCGGTGCCGCGATCGGGCGTGAGACGGAG\nnGGCATGTGCGGAACCGGATCGGAACAAGGTGTGGGTGCGCGCTCAG  
TCGCAACGCGTCTGTGGCGCCGTG\nnCACGGCGAGCTGCGTGCGCGCTCGTTCGGCGCGCCCGGCCTGCACGCGGGCCGCG  
GCGGCCTGGGCGTCTG\nnTCGACCCAGCTGCGCGGGTTCATGCGCGCCGATTGAAGCCGGCGGCCTCGTAGCGCAGCATG  
TCCTGGC\nnGGACCTGCTCACGCGTCAGGCCCGGTGCGGATTGCGCGAAGTGGGCCGATGCGGCAAGCGCGGCGAC\  
nGGCGGCGGCGAGCGGCGCGAGAAAGCGGGTAGTCATGAGGTGGGTCCGGAATGCGGGGTAGTCCGAGCG\nnTCCGGGG  
CTTTGCCGAGGCGCATCGGGTCGATGGCTTCGCAGTCTAGCGAGCCCGCGGATCGATTAAG\nnTAGCGGGCACGCTT  
GGACTCATGAGCCAACTGATGAATCGCCGCCGGCGGTGCTACGATTGGCGTCCG\nnTCCCGACCGAGCTTTTCGACAT  
GGCGCGGAAAAATTTGAACGACCTGCTCGTCTTTCTCGCGGTGGCCC\nnGCGAGCGCAGCTTCACGCGCGGGCCACGCG  
GCTCGGCGTGTGCGAGTCGGCGGTGAGCCATACCATCCG\nnCGATCTCGAAGCGCGGGTCGGCGTGGCGTCTGTGACCGG  
CACGACGCGCAGCGTGCGGACGACCGACGCG\nnGGCGAAGCGCTGTTCAGGCGGTGGCCCCGCGCATCGACGAGATCAC  
CGCGCAACTGGCCGCGCTGTCGG\nnATTTAGGGACAAGCCGGCCGGCGTGGTGCGCATTACGGCCACCGAGCATCCGAT  
CGATACGGTCATCTG\nnGCCGAACTCAGGCAGGTCTGCCCGACTACCCGGACATCCGCGTCGAGCTGTCGGTTCGATTA  
CGGGCTG\nnTCGAACATCGTCGAGGAGCGCTACGACATCGGTGTGCGTACGGCGACCAGGTGGCGAAGGACATGATCG\  
nCCGTGCGGATCAGCCCGGATGTCCGATGGCGATGGTTCGCGGCGCCGGCCTATCTCGACGGCCGCAAGCC\nnGCCGAAG  
AAGCCGAGGACCTGCTCGATCACGATTGCGTCACGTTGCGGCTCGCCACGGCGAAGGGCATC\nnTATGCGTGGGAGCTG  
AAGAAGGGCAGGAACGACGTGCTGGCGCGGTGAACGGCCGATCACCTGCAACA\nnCGCAGCCGCACATGGTCCAGGCG  
GCGCTCGACGGCTTCGGCATCGCGTTCGTACCGAGGACATCGTGCT\nnGGAGCATGTCCGACGCGCCCGCTGCGGATC  
GTGATGCCGACTGGTGCCCGGTTTTTCCCGGCTACCAC\nnGCGTACTATCCGAGCCGCCCGCCGGCGTTCGCTGCGTTT  
ACCGTCGTGATCGACGCGCTGCGGCATCGCG\nnCGTAGCGCGTGTGCGTTCGGGGCCGCCGTCGCGACGCCCGGAAGAA  
AAAAACGCCCCGCGGGCGGGGC\nnGTTTCGACGCGGCTGCGCCGGCGAACCAGCGCGCGTGGCGGCGCGATCAGTGACG  
GTCGCGTGCAGCGG\nnATCGTTGAGTTCTCGGTGCGGTGCGTACGCCGAGTTCTGTCGCCGCGCTCCAG\nn>gi|14  
2022655|gb|EQ086233.1|319 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
scaffold, whole genome shotgun sequence\nnAGCGGCGACGAACATCGCGGCCGGCCGGGCGGCGCGC  
ACCCGGCGTGAGGTCTTCGGCTCCATGGC\nnGGGCTGACGGCGGTGCTCGCACCCGCGCTCGACCGAAGCACGTGCG  
CGAAGCACTGCGCGCGCGCCGCA\nnCGTGGGCGACACCGGCGAGCACAGCGCGTCTGCTGCGCTGCGGCGATCACTG  
GCAAGGCGACGATT\nnCGCGCACCATGGGCCGGCGCGGCTCGACTACCGGTGCTCGGCCAGTCAGGTGGGAATACGT  
GGCCGCG\nnTACGATCACGACGCTGCTTTGGGAGCGCAACCTGCATGCGGAACCTCGGTACGCGGAGCGCCTCGTCC\  
nGGGTGCGCTGGGGCGGCGCCCGCATTCGCGATCGCTACCGTGGGCTGCGTGGCGCGCCGGATCCGGAT\nnCGTCAAC

GGCACGATCCATCGTTTCGGGTCCAACGGGTTCGAGCATGTCGAGGAATCGGCATGGCGCACC\nGGCGCGACCGACGTC  
GAATTCACAGCGACACCTACGGCGACGCCGACAGCATCGAGCTCGACATCGGCA\nnATCTCGCCGCCGCGACGATCGTG  
GTCGAAGGCACGATCGACAGCTTCGCGAAGGTGGCGATCCGCTGCA\nnGCGTAATCCATTCTCCATGCGCCTGAATTC  
CGTTTCGAGATCAGCGGTGCGGAACTGCTCGCGCGTGGT\nnACGGCGACGCATGCGCTTGGCGGAACGGAGCTGTTCTGTC  
GCGGTGCGAGCGGCTGACCGACAGCCGCTGC\nnCCGTCGATCTCTCCGGCAGCATCGAGGTGGCGCCCGGCCATGCGGAA  
TTCGGCTACCGCCCCGCTCTATTT\nnCTTTGGGCGGCAACGCAACGACAGCAAGGTGTGGTCGTCGGCGCAGTTCATCACG  
TTCGGCTAGGTTTTTC\nnAGCACCACTGCGGTGCTCGTTGCCGCCGTGCAACTGGAAGACATTCTTAGCCGATCGATG  
CCAACAA\nnCCGCGACCATCATGACGGGAGCGCCTGTTCTGTTGACGTGGTGCACGCACAATCCTACTTCTGCACGACTG\  
nCGACGTTTCGGGTTGGGGCGTCCATCCCATTGCTTTTCAGGCCGCGTAAGCTAATCGGCCGCCGGGTGCCC\nnGGTGCCG  
GAGGCTGCGATCGCTGCAACGGAGCGGATCACGATGACGATGCGGGAACCGAACAGGTTCAAG\nnGTCGTTTCAGGACGTA  
GCCGACGGCAAGCTCAAGCCATGGCGTGCGGCAGAACGGCTGGCGTCAACGACGC\nnCGCAGGTTTCGCCGCTGGCCGCC  
CGGTACGTGAGCACGTCAGGCAAGTTCCGTTGTCGGGACACCGATC\nnGAAGCTCGGCAACCGTCGTTCTGGACCCAGGA  
GCCGCCGATCGTGCGCTGTGATCATCCGCGACCGCTAC\nnGCCGATTTCGGGCCGACGCTGCCTGCGAGAACCCGTCGC  
GACGCTTGACGGGACGCAGCCGCCCTGGCAA\nnCGGCATGCCGGAACAGTCACGCGGTATCCGCCGTGGCGAAGTCGCC  
CGGGGGCAACGGTCGGGTGCGGA\nnCATCGTGCGCCTCGGCGCGGTGAGGCCGAGCGTCTGCAACACCAGCGCCGCCAC  
CCGCTCCGGCATGCC\nnGCTCGCCGACGCAGCCGTGCGGTGACACTCCGCGGCGATCGCCGCGAGAATCGTTCCGGTCAT  
CGTAGC\nnACGCTCATCAGCGGATCGTCGACGACGAAGCGCGTGGCGCGATGCCGCTTCGCTGTCCCGCCGCAACC\  
nGCCGACCGAGCCCGTGGTCGAGCGCGCACCGAGAAGCCCTCGCGCATCAGGAACCGGG\n>gi|142022655|gb|  
EQ086233.1|438 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole  
genome shotgun sequence\nnCGGGGTCTGATCCGGGAGTCCGGGCCGCCGGCGTATCTTCGCGTGTGATCGGC  
ACCGGCGCGGAGATG\nnGCCGCTTTCGTGCTGCGCTGCTTGACGTGATGGCGCTCGAATAAGCGACCGGCCATCCGGTC  
GACGATG\nnCGATGAAGCGCATGCCGCGCGGTATCCGACGATGTGCTTACTTCATCGTCGGAACCGCATTTCCATCG\  
nCGTCGATCTTCGCGAGCACCGCTTCGCTGAGCGTCTGACGTGGCGCGGCACATCGGGCGTCGACAGCAG\nnTTCCTCG  
ATTCCGCCGCGCCAGTAGCTCGGATGCCGCACGCGCCCCGCACCGCCATCGACACATCGTCA\nnGCGTTCGTCTCGCTG  
TCCAGGAGAGCCACCATTTCTCTGAATGTGAGAGAGCTCTCGTTCTGCATAGTCTC\nnGCGCCAGCATAATGATTGAAATT  
CCCCGAAATACGACAGCAGGCTTCCGTCATCAACGCGCCATCGTG\nnTCGCGGTGCAAAGCATTACCGCACACGTTT  
GTCGCCATCGTACGAGAAGCCGTGTCAGTAACTGCGT\nnGGGCATTTGCTGATTTGATACGGAATGGGGAGCCTGTC  
CGAATGTTCTGACGACTTCCGGCCCCGATGC\nnCTGTAGTCAATCGTAGGCTTACTGTTGGGCGATGACATCTCCCCAA  
ATGGAGATGCGGGGCGATATCGG\nnCGGGGCGGTGTTTCCGGCATTTCAGGTGCTCGGACGTCGCGCGGAAGGCATGCG  
GTTGCCGGCTGGAGC\nnAAATTCGCGATCAGCTCACTACCTCGACTCCACGCGGTCCGGCGCATGACTGACCGGCAAGC  
ACGAGGC\nnGGGCATTTGCCGTGTGCACATTTTGCTTTCCGCATTGCTGGCGCGCGCCTTCACCAGCAGGCGTGTCTGTG\  
nCATCGGGTGTGAAATGTTTCTGTTGAACTACATACTTGAGAGACACGGCCCCGCTGAAGTCAGATACCGAC\nnATCGCCA  
TGCTTCAGCCCGAAAATGAAATCGACCTCCACCGAAGCCCCCTTTGGGCAATTGCAGCACACCG\nnACCGACGTACGCGTA  
TGAACGCTGCGTCGCCAAGCACCGCATACAACATGTGCGATGCACCATCGGCGA\nnCTTCACTTTGATCGTGAAGCCG  
GGTGCGCTTCGCACATAGACGGTGATGCGCGGCACTGTGCTATCGC\nnGTCGAGCGTGCCGAGTGCTTCCTGATCAAC  
GACAACGCACGCAGCGCGGATACGCCGGCGGGCCCCGGCGA\nnCTTCGTGAGCGAAACGGATTCCCCCACTATGCCACG  
AAATGCACCACGTGCGGATCCGCGGAATCT\nnGGCCCCGCAACGTAGGCCGTATTCCCGTCGACCAGAAATGGGCGTGTAC  
TTTCCGCCAATCCTGATTTCTC\nnGTCCGATTGAAGCCGAATTCGGTCGCGACCTGTGTGAGCTTATCGTCTCGGTTT  
ATCAGTGAATCCCCC\nnGATACTTCAGAAAAATGTAAAAATAATTAAGCCTCGCCCGCTTGTCTCGAGGAAGAAATCCC  
GCCGCGA\nnACGATCAAAGAGACGTGCCGCTGCCGTTAAAGTTTCCGGGTAATCCTCTACGTTTTCTCATACATGACA\  
nCACTTTTCAGGTGTTTCAAGCAGTCCAGCATCGCTTTATCGATGCCAAATCTCGTAGTCGACGAGAATTT\nnGCTCAAA  
ATTTTCCGCTGAAAACGCTATCATATGTGTACTCAACCTGAAGAGTGATATATGGCAACGTAC\nnAAGGAACTGAAAGCT  
CAAAATGGAAGCTTTGGCCGAGAAAGCCGAAGCTGCACGTGCTGCGGAATTCAGG\nnCGATCGTCGACGACATCCGCACC  
AAAGTTGCGGAATTCGGCATTACCGAGAAAGACATTTTCCGTACGCG\nnTCGTGGTTCGACCGGCCAAGCAAGCGTCCGCC  
CCGGTGCAAGGAAATATCGCGACCCGAAGACCGGAGCA\nnACCTGGTCCGGCCGCGGTGCGCGCCCCGATGGATCAAG  
GACGCGAAGAATCGCAATCGATTCTGATCC\nnAGGAATAAGATTCTCGACGCGCACCGGCGCGCTCTGGCGCGCGT  
CCTGCTATTTGAGCATCGATACC\nnGGCGGTATTAAAGCCGCTCCGGTGCGGTGAAGGTTTGTGCCATCCGGCGCGTGGT

TGCACGCTACGCGTC\nGGATTGCGTGCTGCCCCGATCTTTCCCGATTCAACGTTTCGTAATAGATCGCAAGCCAGACGGTT  
TGCGTTT\nnCTTCGTGCGTCCACGCGACGCGATGACGGCAATGCGGCTCGATCAGCACGTAGTCGCCGGGATGCATTTTC\  
nATGACGCGCCGGGCTGCTTCGAATTCGAGTACGGCCGCGCCGATAGCAGCACGACCCATTTCGGCGCGT\nnGCGTTCGT  
CATACCAGAAGCCGTCGGGGCTCGTGTGCCCCATCGATACGATTTCGCTCGACGTTTCAGACGCT\nnGCCCCGTGACAAGGA  
GGTCGATTTGTTTCATCGTCGCGCGCTGGCCCTTCGCTGGTGAACAGGTTGCCGCT\nnTTGAAGGTGCATGGGGCGGCCTC  
GTCGATGATCGTCGGTGAAGGAATCGCGCGTTGAGTGACGCGCAAT\nnATTACCGTAGTCTAACGATCACCGTCGACG  
CTGACTCGTTTCGTGCGAAAACGCATCGTAGCGTGTGCG\nnGTGCCTGCAATCGGGCATCAATCCGTTCAAGCGAGATCC  
CGTACTGTTTCGGGATCTCGCCACCGAACGT\nnGTTGCGATTGATGCCGTCAACATCAGCATGCTTTTCATACCGCCTCCC  
GTTACACTCGATCCAGCGGAAC\nnGGTAAGCAGGTCGATCACCGCTTGCGTCTCGGGGCGACGCCGCGCCACCACGCA  
AAGGCTTCGCGCGCC\nnTGCTCCACCAGCATGCCGACACCATCGGCTACGCCGGCGACCCCGCATCGCGCGCCAGTCGC  
AAGAACG\nnGCGTGAGGCGCTTGCCATAGGCCAGTTTCATAAGCCGTACCATGCGGGCTGAAAACGCCGGGCGGCACGGG\  
nCGGCAAAATCGCCTGTCTAGACTCGCGGACGTGCGGTTGACCACCAGATCGAAGCGGGCCATGCGGTTCGAGC\nnTCCCCGT  
AAGCGCACGCCGTTATCGGGCCGTACGAAGCCAATTGCAATGCCAGCGCTTGCGCTTTCTCGA\nnCGTGACGATTTCGCCA  
TCACCAGTTGCGCGGGGCCCGCCAGGAACGGAAGCAACGCGCCGCGCGCCG\nnGCCCGCGCACCGAGCAGCAACA  
CACGCTTGCCCGCGAGCGGCGTGTGAAGATTACCTCGATATCGCGC\nnAGCAGACCGATGCCGTGGAAGTTTTCGCGA  
GAATCCGATCCCCCTCGAACTTGAACGCATTGCGCGCGC\nnCTGCGAGCGTCGCGCGCTCGCTGCGCTCATCGGCCATGT  
CGAACGCCGCGAGCTTGAACGCTGC\nn>gi|142022655|gb|EQ086233.1|210 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnGAGCGAC  
GGGAAAAGACGTACGGGAACTCACGCCGTGCGAGGCGCGCCGGCGAAGGAGCTCCTGCTCG\nnAGCATCTCGACGGCA  
ACCTCGGGCTGGCCGACGTGGCGAGCGCTGCGGTCTGTCTCGCAGCTATTTTCGT\nnGAAGGCGTTTCATCAGACGACCG  
GGCTGCCGCCGCATCGCTGGCTCATCGTGACGCGCTCGAACGGGCT\nnAAGGAGTGGATGCGCAATTCGGGCCTGCGGT  
TGAGCATGATCGCCGACGCGTTCGCGGTTTCGCCGATCAGA\nnGCCACTTCTCGCGACCTTCACGCGCTGACCGGCGTCA  
GCCCTGCGCGCTGGCGCGCCGACAATACCTG\nnATCGTCTGATGCGGGGCGGGTAGGTGAGGTTTCGCTTCGTGCTTCG  
GTATCGGAGCGTATGCATCCTGA\nnCCGGCGCGATTGGGCGGAGTGCAGGCGGGCGCGGAAGCGGCGTGCAGGGTGC  
TTATGCGTAGAAGCC\nnGATCAGCATCTCGAGATTGCGTTGCGCGTGTGCTCGAGCGGCGCGAATTCCGCCGATCGAAA  
CAACGGG\nnGCCGTGCGCGCGCAGATGCTCGAGCGCCTGCCGTGCGCCGGCGCGGAAGGCTGCCGACGGTGCATCGGGAT\  
nCTGCGCCCACTCGCGCGGATCGCGCGGTTCGATTTCGCGGAGCCGTCGGGTGCCGCCGCGAGGATCGC\nnGAGATCG  
GCGTCGAGGAAAATTTGCGCGGCGCGCTGCAATTCCGCATCGTCGCGGAACCCGTCAGGCAGC\nnCGGTGCGCACTTCGTC  
GCCAGCACCGGTGCGGGGAACGACACGTGCGATGCATGCGCGTGAACCCAGG\nnCTGCGTCGCAATGTTTCGTGCGCG  
ACCTGCGCGAGCCATTGCGCGCTGAGCGCTTCGTTGTCCGCATAATC\nnCGGCAGTGTCTGTCGATAGACGACATCGTGC  
GCCACACGGCGAGCTCGATGGCCGGCCATAGCCGGTCC\nnGCACGATACGGCGCGAGGTGCGCGAACAGTTCCGCCAGA  
TGCGCGAGCGTGTGATAGAAGCGCTGCGGCT\nnCGCCGTACGCGCGCTCGACGAGCGGCCACACGTCGTGCTGCGGAAG  
CACGCGTCGAGGCGGCGCGCAG\nnCGTGCGCGGATATCGGAATCGAGGGCGTTTCATGGCGGCGATCGGCTCAGGCGGG  
CGATGCGCGAAGCAT\nnAGCAAACGTGGCCGCGAGCAGGCATGCCGCTTGCGCACGCGTACCGATCGCACCGTATCGGAG  
GTTGCCA\nnTGAAACGAATCCGGCCGCGCTGCCCGCTGCTCGCGATCTGCGCGTTTCGGGGCGCTCGCGCTGCACGCAGA\  
nCTTCGCATCGGCGGCGCCCGACGTATCGATCAGCCGCGCGGATGACGGAACGACCGCGGCGCGGACCGG\nnGTGAGCC  
GGTACACGCTGGCGAACAAGCGCGCGTGACGCTGAAAATCATC\nn>gi|142022655|gb|EQ086233.1|237  
marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequence\nnGGACCGGATAGCCCTGGTCGCGCGCTTCGCTGAGTGCCGTGTACGCATCCGAGCCGAGCAGCAGCAGTA  
\nCGGGCCGTCGACGCCGGAAGCCGACGCGCTTCGAGCGCGTTCGCTGATCGCGTCGGGATATGCGCTGACG\nnTCGGCC  
GGCAGCGTGAGCTTGCGGTTTCGACGTGCCTTCGCGGATACCGACGATGCCGGCGGCCGGGTAGC\nnCGTCGAAAATCGCG  
CCGTCTCGGCGAACGCGAGCCGCTGCGCGGCGTCTTCGCCGTTGCCAGTCGGC\nnGTCGCGCGCGCCGCGCTCGACG  
CTGTCGATCGCATCGCGGCTCAACTCGAACGGCACCGTCAGCTCGACG\nnATCGTGCGGACTTCGCGCAGCCGCGCATTC  
ACCAGCTCGCGCGGGGCGGCGACGTGACGAGGTGCCCCG\nnTGCCGACGGCCGACAGCTCGGGCCCTCGGGGCCGTGCG  
ACATCGACGACGCGGCGGCCGCGGATCG\nnTTTGAACGTGCGTGCCACTTCCTCCTCGATTTGCTCCAGGCGGAG  
GACGAGATCGGGGCGAGTTTCGCGG\nnTGCAGGTTATTCATGGGTTTGGGTTGCTCAGCTCAGCGGAAACACTTCATGGT  
GGGCCGATCGGCTGAC\nnGGACTGCGCTGAGCCCGGCGGCGGATCGCGGACGAAAATCAATATTAGCAGGCTCTATTGA

CACCGCTG\nCAGCACAAATCCGGACGACATCCTGCGGCGGCGCAACTGTTCTGCAATTCGCGTGTGGGGCGACGGAGCG  
\nTATGCGCTGAACGGTGCACCGGATGACGATACCCCGGGGCGCGCGCGCTCGCGTCGCGCGGGCCGG\nTGGCGCGCCGAACCTGTTGGCGCGGACGGCCCATCGCGTGTCTGCAATTCTTGCGAACATTCA\nTCCGGCAGGGTTCCCGACGCGGGGCTCGGCGGTATTAGTTGCACGGACCGTGCCCGGATGAATCATA\nnGACGCAGCCTATTGAAGCTTCA CGACAAGTGAGGGGCGTTTCATCGCATGGCCGCGCGTCTCGCCCCGGCG\nnGCGGTGGGCGCGCGCGGACTTTGGCGATGGGCGTTACGACGGCAGCCTTTTCAGCGAGCCGTGCGATGCG\nnGTGCGAATGACGCCGGTTCGGCGCGCGCGGCATCGGCGGCCGGCACGAAGAACAGCGAACCGGTAACCGC\nnGCGGTGTAGTCGAGCAGGCGGTCTGTAATTGCCGAATTCGGCAGCTATATCACTGGATATGCGCGCTCGC\nnCGGCACTGATCAAACAGATACTCGAGAACGTGTTCTGTTGCCACCTGCGTCGGTA GGTTCCAACGTAGCAAC\nnTCTTTTTCGGGTTGAAAAGGTTGGCAGCCTGTGCCTAACACCGGCTGCCAACCGGATGATGA AACGAATA\nnAATTGCTGTGTGATTTAGTCGAGCGGACCTCTCCGTTTCATCCTAAGATCAAAAACGACTTGGCCAGTTT\n\nCGGTATCCACCGTCGCAGCGTAAGACAGCAGTTCTGTTCCCGTCGACATTGAAGCCTTCACTGAACACGTG\nnCAGTAC CGAAGCGAGCGGCACTTGAATTCGTGATGTTGGGCATCTGCGCCGTAGGTGTGACGTGCCCG\nnACAGATTGGGTCTGATACTTCGTTCATGGTGATATCTTGAGAAAGTAAATCTGCTCAGACTGCGCCTTCTG\nnGAAAGTCGTTCCCGGAGGCATTGATTATCTGAGGGGTTGAAACGAAGAGATCGTCAAATTCGGACGCTTC\nnGACTGTGGCCCATCTTGCGAAGTTTCGACCGTGCTTTGTGCGAGAGATGGCTCCCGATCTTCTCTATCACA\nnTCTGCGTTTTCGATAACAAGTTGTTTTGTAGTGTCTATAAGTTTCGCAGCGTAGTAATTTTACCGCTCGAT\nnCGCGGGCTGCCGGTTCACATAATGCAGGATTTGATGCACATTGCAG GTTGACCAGAAGCTCAAGATGCTC\nnAGCTGCTTTAGCAAAATCTGATACTGCGGTTTCAATTTACGCTTCCGAGCGATCTCGATTTGAACTTGA\nnCCGTTTGAGAGCGATTCTGCGAGTACCCCGCGAATAGCACTTGGACCCGATCTCTGAGATA TCCAGAAA\nnTTGCATCTGTCTGCGGACGGGACGATGAAGTATGACTGTGGCGCTTCCGTCGTGCGGTACATCTTCCGT\n\nCTCATGCCCCGCTGCCCTGGAACGAAGTTCGAGCGGGCCCATCATGAAGCCCAGCCTTCGAGCAACGACG\nnTGGTGCCGGCTTCATGTTGGCCGATTCTGTGATTGAGATTCTTCTTTTCTTCCCATTCTGGAAGGG\nnTGTCGTAATTCATGGGCTTATCTTCTGTCCGCTGTGTTGTACGTAGCGCATCTCATACTTGAGGTGG\nnTGACATCTTTGGGTGCGTAACGATTGGAACCAAGTCACTACGCTCCAGCCGACGTAGCGGCGCGGATTTC\nnTCATTCTGACGTTTCGGATCCGAGCAGGATGCTCACGTCTTCTAGCAGACGGGAATTGGGACGGGACTG\nnCGTAGTCAAGTAAGCAGTCCCGGTCTCTCGAATTCAATGCGACATTTTCGGGGCGGCCGAAAAGCGCAA\nnCTTTCACCGCTACCCCTTTGATTGACCAAACCTTGGCTGCAATTTCCGGCATCGCAGCTTGACGTTTTTC\nnAACCGCATCCGTGTTCCAAGTGCTAAGCTTGCTTACCCCGTTCGTAAAATGAAATGCTGAAGTTGAGAAA\n\nCTGGATTTTCTTTGAGAATTCCTCGTTCCCAATGGTGCTGTTGATTTTGTCTCCCATGATCGCCATGT\nnTGCCCTAACCGCTTTTGATAAGCTTTGACATCGTCAAGTGACCATGTCTTCATCCATTTTTCAGACGGATT\n\nGATGGGGAGAATGTGCTCTAAATTAACCTTGACCAGCGTCGGTGCTTGGCACAAGCTCGCCGAGGCGCCA\nnGGCACA CCACTCTCAAGCATCCGAGGTAGTATCTGGCAAGTACTGCTTCGATACTGTAGCTACGGA\nnACGCTGCAATGAATGTGGGGTCAGTCGGTAGAGTGGTAAATCCTTTCAATATATCTTGGATTGAATTGAG\nnCTTTCCTCTGTAAACCCCTTTGCTGCATCGCTGTAGATTGCTCTAAGGTGCCACCGCCTGCGCCACCA\nnACGATTTGGAATCTTACGGCGACCGCAACAAGCTTTTTAAATGCTGCTGATACTTTTTGGGCTCGAATT\n\nTATCTAGGATTGCAAGCAAAAGCGGGCGAACTTGAACCATTCCAAGAAGGTTCAATT\nn>gi|142022655|gb|EQ086233.1|507 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCAGGCAGCGTTGCGCGCGTGACATGAATCATCGGGCGGCTATCGAGGTACAACCAGTAGCCGTCGACCCT\nnGAACGGCGGTCGAGGCCGGTATCGAGACCGACGACCTCATCGAAAAAGCGGCGTGACCGTCAATGTCA\nnTCGGTAACGATGGTGGCGTGGTCGAAATGCATGCTTCGTGTCCATGCTGGAAGCCTGGTACGATTCTG\nnTCGCACGTGGTGGTCTTGATAATTGGCTTCGCCATTGAATGATCTTCAATGGCGACAAGAAGGTGGG\nnGGCCGCGACGGGCGCCGATCCTACAGCCCCTGCGGCGCATCTCGCAGCACGGCGGTAGCGCGGCATCG\n\nAACGGCGTCGCGTGGTTCGATCCCGGCGGCACGCACGCGCGATGCGTCAGGTGGCAGTCATATGGGCGCG\nnGTGTGCGGTGGTGGTTCGATTGCTTGAGCGGGCTTCGATGCCGAGCGCGGCCGCGATCCGCTG\nnCGCAATGTCGTACTTCGTTCATCGGCTCCTCGCCGACCAAGTGACGAAGCCCGTGATCGACGCGCCCCGCG\n\nAGATGGCGCAGCGTGAGGTGCGGATGACGTCCGCGACATCCGGCGTATACGTGCGGTAAACGGATCGCCC\n\nATGCGTCCATGCCGACCGCGTCGGCGCCCGGCGAGCGGCAGCGACGATCGCCGCGACGAGGCTCGTGAC\nnGGCCGATTCGCTCCAGTCGACGATCGGGCCGAAGAGCAGCGGCAGGCGAGCACACGACAGCGGCGAC\nnGCGGCGAGCAGCGCCGCCTCGCCTTCCAGCTTCGTGCGGCCGTAAATGTTTCAGCGGGTTCGGCGCGGCGT\n\nCCTCGCGGTACGGGCGAGCCTTGCCGTCGAAAACGTAATCGGTGGAGATGCCGAGCGTCCATGCGCCGTA\nnGCGCGCGGCGAGCACGCCGATGCGGGCCTGG

GCCGTGACGTTGATGGCGCGTGCGGCCGCGGGATCGCGT\nTCGCAGACATCCGGGCGGCGCTCGGCCGCGCAGATGATC  
 ACGGCGGCCGCTCGGCAGGTGGCGAAGAGGT\nGTTTCGAGGGCTGGCTCGTCGAGGACGTCGAGTTCGGCGAGGTTTTCC  
 GGGGCGAGCGCCAGGCGTTTTGC\nGCCTGCGCTCTGCGGTTCCGGATGGTCGCGACGAGTTCAGCGACGATTCACGG  
 GCCAGGGATGCGGCC\nACC GCGCGCCGAGCAGGCCGCGCGCGCCGATGAGGAGGACGGTCGGGAGGTCAGTCGACGAG  
 GTCGGGT\nTTGGCATGGGCGGTCAATTCGGGTGACTTGGAAGTGACGAGGTCGCGTTCGACGCTCGGGCCCGGGTGGC\  
 nGGGAGCATAGCCGAGTATGACAGTCGGTACGGCGTGATGACAGATCGGCAACCCCGCGACGGACGTGCG\nTCATGTG  
 CGCCGCGGCCGCGCTCGGCACGGCGCCCCGCTCCACCGGACGACGCCGCTCTGCCGCGCA\nCGCCGACACGCCGCG  
 CGGCCGCCCTGCGCGGAGCAAGCCGCGGGCGCGCCCTCACGAAGATCGCCC\nAAAAGCAAAGCCCGGTGCGCTTG  
 CACGCACCGGGCCGTACCTCTGCGGAAAACCGCGCCCCGCGTTACT\nTCCGGTCGAGCGAATAGCGCCCCGGCCCGGT  
 CAGTGCGAGCAACAGCAGCCCGCGATGATGCTAACGTT\nCTTGTAAGATTGATCATCGCCATGTATTGCTCCATCCC  
 TTGCAACGCCAGTAGCGATGGCCGATCAAC\nGCGGTGCGGAGCGTATAGGCGGCAACACCAAGTCGAGCGGACGCGT  
 GTAGAAGCCGATCGCGATCAGCA\nGGCCACCGGCCAGCTCGACCGCCACCGCGATCGCGGCCGCGAGTCGGGCGCCGG  
 GTTCCCCGTGCGACG\nCATGTACGCGACCGTGCCCGAGAAGCCGGTCAGCTTCTGCCAGCCGAACAGCACGAACAGGAT  
 CATCAAC\nAGCACGCGAGCCGCCAGCAGCAGCTCGTCTTCTTCGACTCCAGTGAAACGTAACGCATGACAATCACCC\  
 nTTGAATTGACGTTTTGATGCGTCCGGTGCCGAGATGCGCCGAATCGGCGCGCCAGGTCGGCAACGAACCA\nCTCCTCG  
 CACGATCGGACAGCAAG\n>gi|142022655|gb|EQ086233.1|350 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGCAATGG  
 GTGTGAGGAAGCGGCGCCGACGAGCCCCGTATTTTCGAATCTTCAATCCTGTTTTGCAGGGG\nGAAAAGTTCGACTCG  
 AGCGGCGAATACACGCGCGGTGGGTGCCGAAGTGTCTGCACTCCCCGCCAAAA\nCACTCCACAGCCCATGGACCGCT  
 TCCGCTGAAGCGCTGCGGGAAGCATCCATCGCGCTCGGCGCAACTTA\nTCCCTTCCCGATCGTGGATCACCAGGTTGCA  
 CGCGCGGAGCCCTCAGTTCGGTAGAGATCCTGAATACA\nCAGAGCCAAATGAAGTGACCTTCGATCGCTAGCCGGGCT  
 CGACCGTGCGACGGGGCGCCCCGACGACATA\nCCGGGGGCGACGTGATGTGCAAGCGCGCACGGCTGCTTCCGCACAT  
 CACGCCCTTCTGCCGATACAGGC\nTGAAGCATCAGGGCATCAGCACGGAGTCGACCACGTGAATCACGCCATTGGAAGT  
 CATGACATCGCCGAT\nCGTGACGTGCGCGACGTTGCCCTTGTCGTGCGGTACAGGCCAGCCCGTTGCGGCCACGGCTGAC  
 GATCAGG\nGAAATCGCCTTCGACCGTCTTCAGCGTTGCCCTTGCCGCCACCTTGCTCGACCGCCCTGGCAAGGTGCAATG\  
 nCCGTGAGTCGTCCGATACGACGTGATACGTCAACACCTTGACGAGCATCGCCTTGTTCTCCGGCTCAA\nCAACGTC  
 TGCACCGTGCCGGCCGGCAACGCGGCGAATGCGTGGTTGGTTCGGTGCAAAATACCGTGAAGGGC\nCCCTTGCTCGACAAT  
 GTGTCGACCGAGCCACCCGCTTGACCGCAGCGACCGAGCGTCTGTGATCGTGAG\nAATTCACCGCGTTTTTCGACGATG  
 TTCTCGACGGGTACATCGCAGCGCCACCCACCTCGACCGTCTGCTC\nCATCGACATGGATCCTGCCTCGGTGGCAGGC  
 GCGTACACGCACGTTGCGAATGCAAGGGCGAGCGCCGCA\nCCACCCAGTTTAACTTTGCTGCAAGATTTCATTTTCGTT  
 CTCCGAAAGATTGCCGCGTTGACTGGAAGC\nAATCGGAGATACGAGTCGAACCGCGCAGCGGATGCGCGATGAGTACT  
 TCATCGCGGACAGATCCGCGATC\nCAGGGAATGCCGAGGGCGTCGACAACCGTTTCGCATCTTCAATCGCCACTGCATCC  
 GTTCTCGCTCGGCCA\nCCGTATCTGCAAATGTTCATCCAGCCAAGGTGCTTGGCGTCCACCTCGGGAATGGCGTCGCG  
 ATGGCAT\nTCAGAGGAGACAATCATGAAGCTTTTCATCGGTATCGTTCGTTGCCGAGCGGCGCGGTTTCGCGCTCAGC\  
 nGCTTGCGGAGGCGGAGACGACATTGCGCCCGTGGCAGCGCAACATCCGGAGAATCTGGCCATCTGGACCG\nCGCTCGG  
 CGGTGATTTCGGCGGCCCCCGCGTGTAGTGCCCGGCGTGGTCAATGCGGCTGTGCGAGGCCTTTT\nGGCGGACCCGGTTCGA  
 AGCGCCCTATTTTTCGGCTGTAACAGGTAATGCCACCCACAGCAGTGACGCCCAT\nGACACGCCAGCACGGCTTGAAGC  
 CTGTCTACGCTGCAGTTCGAAGCGCTCCTCGGCGGTCCGTACACCT\nACCCTGGCCCTGTCAAGGTGGCCGAGACAT  
 CGACGCCAGCCGAGATGTGCCAGGACATGACGGCTGC\nCCATAACGACGTGGCGTTCCAGGCTGCGTGT\n>gi|1  
 42022655|gb|EQ086233.1|245 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
 scaffold, whole genome shotgun sequence\nCCAGCCGAGCCGAGTGCCAGCACGGCGCAGGCGGCGAG  
 CAGCGGCACCGTGACGATCATCGCGAGCAGC\nCCGAGGCGGCACGGCACGAGCCACAGCCGTCGGTAAGCGGTACGCG  
 CAGCCGCGAGCCGATCGCGAACA\nGCGAGACGAGCAGCGGATCTCGACGATTTACGCGAGCAGCCGCGGTCGCGTTC  
 GAGGTGCGAGGTTTCAG\nCAGGCCGCGCGCGGCGCGGCGGAGCGCGATGCCACCGCGAGATAGATCATCGCAGTGCTGAC  
 GGGCAGG\nTGCGCGCAGCGCGGACGTGCGGACGCCCATGCCAGCAGCACCGTGCCGACGATCAGGTACCAGAGGGTTT\  
 nCGTGATGTGTGCGGTAACAGATCGATCAGTGCCGCGGTGCCGTACGGCTTGCGGCCGAACGCCTCGC\nGCGAGCTT  
 CCGCTTCGCGCGCTCGAGCGTTGCGCGGCGCGCTTTTCGGCAGGTTGCGGCCCGCGCGGTTTCAC\nGTAGAAGTTCAGCAT

CGACATCGCCGACTGGAACGGCGACGCCTTGCGGGCGGCGCTGCGTTCGGCCGAA\nTGCTTGAGCGACGGGCGATCGC  
GGCCGGGTTCATCGGACTTGAAGATGTCGGGTTCGATGTCGAGCGCAT\nnCGCTTTTTTGCATCACGTCCGCGGACCAGCG  
CTTGCGGTGGTGCGGATCGGTGCCGCGGGCGGCCCCGG\nnGCGCCGAGTCAGGCGCTTGCCGTGCGCGTGGGTGGCGGT  
AGCGGAAGGCATTACGTGGGTCCATTTGGG\nnGGACAGGGGTGTAGTGACGCGTGTGGCGGCGCAAGCCCCGTGCCGC  
GCGGCGCGCTCATCCGTACGAC\nnGAACCCGTGCGGGGCGGCACGGGTGTTGCTGGAAAGGACCGGCAGGACATGCGAT  
CCCGCCATCACGCAG\nnGAGGACACCATGCAATCGCATTCGCCGATCTGGCTGGAAAACGACGAGATCGACAACGATGCC  
GATGGCG\nnACACGATCGACGCGGACGATCTCGATACGACGGTCCATGTGACGTCGAAACGGGCAGGATCACGTTTCA\  
nGGCGGCGTGGGCGGAGGCGGCCGATCCGCCGCCGAGACCGCACGGCATTTCGGTCGAGCTCGTGCTCGAC\nnTGCGACA  
CGATGGAGCGCTATGCGGATCTCGACGAAAGCACGCGGATGCGCGTGACGCGATCCTGCACG\nnCCACCGTGGAGGATC  
TGTCGACCGCCTGCCGACGAACACGCGACGCTGACGATCGAGCTGACCGATGC\nnAATGCTCGATGCCGCGAGCCGTC  
TGCAGTAGCCGCGCTGATCAGCCTTTCGATACGCGATTTGCG\nn>gi|142022655|gb|EQ086233.1|279

marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequence\nnAGGCTTGACCTTGTGCGGCAGCCGGCAGCCGCCGCGGCGGCGAGGTTGCCGTACGGATCCTCGTCGGCCGCC  
\nGTGCGGGGGGCCGCCGGCATCACGGGGGCGGTGCGACGCGACCAGCATCGGGGGGAACGGCAACGCATGCA\nnCCCCT  
CGCCGTCAGCACGAACGCACCGTTGGCCACCGCCGCGCGATCGGCGGCAGCCGGCGTCGCT\nnCAACCCGGTCGGCT  
GGGCGTCCGACGGCACGAAGAAGACATCCACCGGCGGCGCTTCCTGCGATGCGTATC\nnGGCGAATAGTCGGCGAAACCGG  
CGTTGCGGACCGCGCCATGGTCGACGTCGATCGCGAAGCCGGGCTTCG\nnTCGTGCGGAGACCGAACAGCGCGCCGCCCT  
GGATCTGCGCTTGCGCGCCGGTTCGGTTGACGATGCGGCC\nnCGCATACACCGCGCCGTACGCGATGACGCGCGGTT  
GTTGCGCTTCGATCGACACTTCGTCACGTAC\nnGCGACGACCGAGCCGGCCGTTTCGTGATCGCGACGCCCCACGCGT  
GCCCCGCCGCGCAGCGTGCGCGCGC\nnCGTAGCCGACTTGTGACGCGCCAGCGCGAGCGCCTGCCGATGCGCGGCGTGCT  
CGGGGCCGCCAGCCG\nnCGTCATCCGGTAGGCGACCGGATCCTGCCGCGCCGAGTGCGCGAGCTCGTCGACCAGCGTTT  
CCATCACG\nnAACGCCGTATGCGAGTTGCCGCCGAGCGCCACGTCTGGACCGGCACGTGCGCCTCGGTCTGATGAACCG  
\nATACCTGCATCGGAAGCCGTACGGGCTGTTGTCACGCCTTCGGTCAGGCTCGGATCGGTGCCGCGCTT\nnGAGCAT  
CGTCGTGCGCTCGAGCGGCGAGCCCTTCAGCACAGACTGGCCGACGACCACGTGCTGCCAGTCG\nnCGCACGGCGCCGCT  
GCCGTCCACGCCGATGTCGACGCGATGCAGCACCATCGGGCGGTAATAGCCGCCGC\nnGCAGATCGTCCTCGCGCGTCCA  
GATCGTCTTGACGGGGCCGAGATGGCCGGCCGCGAGGTACGCGCGGA\nnCACGTGGGCGGCTTCGACCAGTAGTCCGA  
CGTCGGCGTCGAGCGCCGCCATAGTCGCCGCCGAGGTC\nnAGCGTGAAGATCTGGACTTTCTCCGGGGCGACGCCGAG  
CGCTTCGCGACCGCCGCGCGGTTCGGTCGTC\nn>gi|142022655|gb|EQ086233.1|378 marine

metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ce\nnCATGGCGCGCCTGCTGCACGCGCAACGCCTGTGCGATCCAGAGCGCGGCGAGTCGACGCGATCACCGCGAT\nnGCCG  
ATCGACGCGGCGAACAGCGCGGCATCGTAGCGGATGCCGGTTCCATGCGCATCGCGGCCATCCCC\nnGTGTAATGCATG  
CCGGCGATCCCGGCGCCCATCAGTGCGCGCCACGAGCAGCCGCCGAGCCAGCC\nnGTGCGCGGTGACGACGGTC  
AGCGCGAAGTACGACACGAGCACGGCAATCGCCAGCGACGCGCCCGTATC\nnCGGCCATGCGTAGCCGAGCGGAATCGGC  
AGCGAAAACGCGAGCATGCCGACGAAATGCATCGACCAGATG\nnCCGGTGCCATGGCCACCGCGCCGCCGAGCCAC  
GCACGCTTGAGCTTCGATTGTGAGCAGCGAAA\nnTGAAGCGGCCAGGTGAGTGTCGTATAGGAGGCCAGCGTCGCG  
ATCGCGAGCGACAGCAGGACGAGCGG\nnGAGATTGTAGGTGCCGTACATGATCGAGCGCCCGATCGAGGTGAGTACGTT  
GCCGCGGCGCGCA\nn>gi|142022655|gb|EQ086233.1|101 marine metagenome

JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnCACATCG  
ACACGAAGATCACGCGCATGCGTTGCTGATCACGCTCGTCAGCGCGCGTGCTCGGACATGA\nnAGCGATCGATGCCGA  
CCAGCAGTGCGACACCGGCGACGGGCAGGTGCGGCATGACGACGAGCGTGCGGAC\nnACGCGCAACCAGCCCGCTTCGG  
AAACGCCGGCCGCGCCCTTGACGTGAGCAGCATGATGGCGAGCATC\nnACGGCGATCTGCGACGCGAAAGGGGCACGT  
CGCACGCTGCGGATGAACAACGCGGCGAGCGTCAGAT\nnAGATCGCGGTACCGTCCAGATTGAACGAATAACCCGCCG  
GCAGCACGAGCCCCACGACGCCCTTGTCGCA\nnCCCGAGCGATTCCAGCTTGACGATCAGGCGTGGCAGAACGGGCTCCG  
AAGAGGACGTGCGGAGGACGATG\nnAGCAACTCTTCGCGCAGGTAGCGCAAGAGCCGCCACAGCGCGAAGCCGTGCAGCC  
GCGCGAGCGGGGCGA\nnGCACCAGTGCGACGAACAGCCCGCAGGCCACGTAGAAGGACAGCATCAGTTTCGCGAGCGAGC  
CGATCGA\nnGCCGATTCCGAAGCGGCCACCGTGAAGGCATCGCGCCGAATGCGCGAGCGGCGGAGCCGATGATC\  
nATCGCGAGCACGCGAAAGACGACCTGGGCGACCCGTCGATCAGTGCAAGAACGGGCCGCCCGGCCGCG\nnGGTGTGC

```

GTTTCAGCGAGAAGCCGAACAACAGCGACAGCAGCAGCACCAGCAACACCTCGCCTTTCTCGAA\nCGCGCCGAGCATCGT
ATCGGGGATCACGCTCAGCCGAACGCGACGAGCCCGTTTCGGTTGCGCGTCCCTC\nACGTACGGCGGAGGATGCGCGA
ATCGAGATGGTGGACGTCGATGTGCATGCCGGCGCCCGGTTGCAACA\nCGAACGCGGTGATGAGCCCGATGGCGAGCGC
GACGGCGGTGAGCAGATAGAAGAGCGCGAGCGCCCGCAC\nGATCGTGCGGCCGATTGCCTGCCCGTTGCCAGCGACGT
GATGCCCAGAGACGATCGTGCAGAACACGATC\nGGCGCGATGGTGCATCCTCACGAGGCCGACGAACGCGTCGCTGAGCGG
TTTGAACATCGCGCCTGCGTCCG\nGCCATACATGACCGACGGTCACGCCGAGCAGCATCGCGAGCAGGACCTGCACGTA
GAGCGATTTGAGCAG\nCCTGGCCAACCTCACGGAACGGTCTTCCTTCGTGTCTCTTGTGTCGTTGTCACACCGATGCAT
CGCTGTT\nCGCGATGCCGTGCGCTTCGCGTCGACGATCATCCGGTGAAATGCTTCGGCGGCCGCGGTGAGCGGCCGT\
nCCGCGCCGCTTGACGATGCCACGCGCGCTTTGACGACCGGCTCGACGAGCGGCACGCTCGTGAGAATCG\nGGTGGTC
GTGCCCGGGCATCGCCATCGACGGCACCGCGCGACGCCGAGCCCCGCTCGATCAAGCCGAG\nCAGGGTCGTACGTCG
GCGCGCTTCGCATACGCTCGGCCCCGCGGCGCCACGGCGGCCAGCGCCTGGTCG\nAGCAACAGGCGGTTCCCGGAAGT
CTTGTCACCGACACGTAATCGTGCTCGTACAGTTCGTTCCAGGTAA\nCGCGCTTCTTGCGCGGAGCGGATGGTCGCG
GCGGCAGGCAGCGACGAACCGCTCCTGGAGCAACATCTT\nGAACTCGATGTGCGATTCTGGCTGCCATGAAGCTCAC
GCCGAAATCGGCTTCGCCGCTGATGACGGCG\nCCCAGCACCTCGTTCGCGCTCGCGTCCAGCAGCTTGACCCGGATGCG
CGGAAAGCGCTGATGATAGCGCG\nCGATGATGGCCGGCAGAAAGTAGTAGGCGACCGAGGGCACGCACGCGATGGTCAC
ATGGCCCAGGCGGCT\nCGACGACACGTGCGGAATGCCGAGCAGCGCCGATCGAGATCGTCGAGCAGCTGTTGCGGCGT
CTGGGCG\nAACACGCGGCCGACCGTGGTGAGCGCGACGCGACGCGTGGTGCGCTCGAACAGGCGCACGCCGAGCGCTT\
nCCTCGAGCTTGTGATCCGGCGACTCAACGCGGGCTGGGAAATGCTGACCGATTCCGCGGCCTTGCGGAA\nACTGCCC
GTTTCCACGACCGCGCGAAACGCTGCAAGTCGTTCAAGTCGAAGTTGATCCCCACGGGCGCG\nTCTCCCCATCTCAGA
TGGGGCGTATTTTGATGATTTGCGCGGGCGGCCGATCGGCGCGGCACGCATTC\nGCGCCACCCTCGATCGCAACCGC
GTGCGTGAGCGTGCGGCCTGGCCGTCGATCCGCCGCGCGTCGATGTC\nGATGGGGGGCGTCGAGCCTGCGGGCCTTACG
GCGCAAGCAGGCCGCGTGCGGCCAGATTGGCATAACAGCG\nCGCGCACGCCGAAGGTCCAAGGCGCGATCTCGGTGCACA
GCCGCACGGTGTTGACGAGTGCGCCGAGCG\n'

```

```

[19]: c = a.count(">")
      c

```

[19]: 25

```

[15]: b = a[120:]
      b

```

```

[15]: '\nTCGGGCGAAGGCGGCAGCAAGTCGTCCACGCGCAGCGCGGCACCGCGGGCCTCTGCCGTGCGCTGCTTGG\nCCATG
GCCTCCAGCGCACCGATCGGATCAAAGCCGCTGAAGCCTTCGCGCATCAGGCGGCCATAGTTGGC\nGCCAGTGACCGTA
CCAACCGCCTTGATGCGGCGCTCGGTTCATCGCTGCATTGATCGAGTAGCCACCGCCG\nCCGCAAATGCCAGCACGCCA
ATGCGTTCTTTCATCCACATAGGGGAGCGTTACGAGGTAGTCGCAGACCA\nCGCGGAAATCCTCGACGCGCAGTGTCGGG
TCTTCGGTAAACGTTGTTTCGCCGCGCTGGCACCCCTGGAA\nGCTGGCGTCGAAGGCGATGACGACGAAACCTTCCTTG
GCCAGCGCCTCGCCATACAGTTCCCCGATGTT\nTGCTCCTTGACGCTGCCGATCGGATGCGCGCTGATGATGGCGGGA
TATTTCTTGCTTCGTGCAAGTTCG\nGCGGGAAGTGATGTGCGCTGCGATATCCAATACACATTCTTGATCTTGACG
CTTTTCATGACAGCTCC\nGTTTCAGGGGGAGGGGGTAAGTTCGCCAGGCCGAATCGTTGGTAGCCAAGCGGCAACGACTC
GAATATAGA\nGAGCCGATTGGAATTCGTAAGATCGCAATCTGGACTACAGTGGTATCTTCAAATTGACAATGGCACCT
A\nCATGGATCCCTCACTGCTTCCGTCTCTCGCGTGGTTCGCCCACGTCGCACATCATCGTAGCTTCACGAAA\nGCGGC
TGCGGAAATGGGCGTTTCTCGAGCAAACCTGTGCGAGAACGTGAAGGCGCTCGAACGCCGTTGA\nACGTCAAGCTGCT
GTATCGAACGACTCGCGACATGTGCTGACCGAGGAGGGGCGAGCGGCTCTACGAGGT\nGTGGTATCCCGCGCTGGTCGC
GGTCGAGCGGACGGTCGACGCGCTGCACGAGGAGCGCGACGAGCCGTCG\nGGGTTGATT\n>gi|142022655|gb|
EQ086233.1|160 marine metagenome JCVI_SCAF_1096627390048 genomic scaffold, whole
genome shotgun sequence\nATTGGGGAGGAGGCGAGTTGAGCGGCGGCAGTTCGCCTGCGTGCGCGGCG

```



TCGACATCTGATCCG\nCACCATGGAAATCCCCGCTCAATCTTTGGAGCAGGGATGCGGGGCGATCAAGATGGGGATGCG  
GGATGGG\nnGGCGACGGTGTATTTCCGCCAGAAGATTTCCGCCGCGGAGCTCGCGGTGCGTACGTGCATGTTCAAACGC\  
nACGGTGCGCGCATGGCAGTGGCAGACTGATCAACGCAGCTGGAAGCATCCGAAGCGCGCGGGCACGCGTG\nnTCCTCGA  
CGCGTGGCCTCACATGCTGTCTGGGTCTGGTTCAAGACCGAAAGCCACCGACCGACGCGCGAGCA\nnATGCGCTACGCGGAT  
CGCGTTCGACACGAGCCGCGCGGAGGCAAGGCCGACGTATTTCGATCTTCCAGAG\nnGAAGCCTATTGGCTCGAGTTCGTA  
GTGCTCGATATGGTAGAGCAACATGAATCCCCGGCTAAGTACAAGAA\nnGTAACCCGGCAACGAGTGAGATTGCGACGAA  
TAAACGCTTCACCATGATCGCGCTCCTGAGTTGGTTGAG\nnGTGAATTGGAAGTCGATTCTGCGGGGATCATTCCCGGC  
AAGGCGCGCAATCCCCGATTGTTCTCAAGA\nnTCGCAACGCGATTTCGTCAGGCCGATCTTCATGGGGTGTCTCGCTGGT  
AGTGATTCCGTGCTGGCCCGCGC\nnATGTGCATGACGGCATCCGGGGAG\nn>gi|142022655|gb|EQ086233.1|4  
1 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
shotgun sequence\nnGACCTTGATCGGCTGATCGCCGAGCGTGCCGACGTATTCAATCACAACTGCGGCCGGGG  
CGGCTGCG\nnGATTATGGGCTCGACGCAAAACAGCCTGCGAATTACGAAGCCGGAGTTGATCTGCTGCGAGATCGGTGCAT  
\nTCGGCCATGTCTGGACCGATGAATACGCTTCCCGGTTACGACCCGCTGATGCAGGCATTTTCCGGGATCAT\nnGAGCAT  
CACGGGCGAAGAAGGTGAGGCACCGGTGCGAGCAGGGGTGTCGATCGTCGATTTTCGGCGCGGGC\nnATGTGGGCTGCCAT  
GGGCATTCTGGCAGCGCTGTACCGGCGCCAGATGAAGCACTGCGGGGCGACCGTCA\nnATGCCTCCCTGCTGGAGACCGC  
ACTGGCGTGGATGTCTGATCGGCATCGGAACATAACGCCGACGGCGA\nnACCGGGTGGCCGGCATGGGTGCGGCGTTCGCG  
GTTTATCGTGCCGACCGTTCGCTTCGCGACCGCGGATGGC\nnTTCTCATCGTGAGTTGCGGAACGATCCGCTCTTCGC  
GCGGTTGTGTGCCGCGCTCGACCATCCGGAAT\nnGGGCGACGGACGAGCGGTTTCGCGACGAACGCGGGCGCCTGAAGAA  
CCGCGCCGAGATCGACCGCTGAT\nnCGGCGACCGGTTGTCTGACGGGAACGCGGGAGTTCTGGACGGAACGGTTGCAGGC  
CGCCGGAATTCCGGTT\nnGCGCCGATTCTGACGACCGCGGAGGTGCTGGCGCACGAGCAGACGACGAGCGCTCGGCATCAT  
CGAGAAGC\nnCGTCTGACGACGAGATCGGTATCGTTCGGCTGCGCTTTCGTTTCGACGGCAAGCGTTCCCCGCGCTGCA  
\nTCTGGCGAAGGATACCGGCGTGACAATTCTCTCTCGACAGTCTTCTGAAGGTCAGTCATTAGCCGCTT\nnTTTCCG  
AAGCAGACGATGCGGCATGGCGGATACAGTGCCGTTTCGAGTCGTCGAAGCGCCACTTGGTGA\nnTCCCGCCAGCCGCG  
CAATTTGGGTCTGGTTGGCGGGTCATTTATTGGCGTACCGACAGATGACGTTGGTG\nnTACGATCAGGGTCGACGTCAGT  
CGACGAGGAAGCACATTTCTGCTGCCATCGATGCGTTATCGTGGTGAGG\nnCGTTTGAGAGTGGAAGTGCACGCCGTGTGA  
CATGACGCTGACAACGTCATGAAGACGTGCACGTGAGGA\nnGACATGAGTACATTCAACCGATTGCACCTGATTCTGTC  
GGTGGATTTGTTTACGCTGAAGCTGTTCTGTT\nnCGGCGGTTGAGGAGCAGCAGATCGGCCTGGCCGCGATCCGGGAGAA  
CATCGCCGATCCACTGCAACCAA\nnGCGTATTCTAGGATCTCGAAGATATTGCCGGCGTAAACTTCTCGAACGAGGCCC  
GAAAGGCGTCTCGCCC\nnAGCCCGCGCGGTGCACTACTCCTGCGCTATATCCGCAAGATTTTCGACGATCTGGACGATAT  
GCGATCCG\nnAGATCGCGTCTGTTACCGACGGCATGCAGGGCGAGTGACGATCGCATCCGCACGGTCGATCATCGTTCC  
\nGTTTCTATCGCGCAAGTAGGCGAGTTCAGTCGGGAATACCCGCTTGTGCAACTTGTGATGCAGGAGCTC\nnGAAAAC  
GCTGAAATCGTGAGGCCGTTGTTTCGGGGCGAAGCCGATGTCTGGGGTGTTCGCCGCGACACACG\nnAGCTGGATCTCGGC  
GGTGTGATGTACGCCGTACCGGGAAGACAGGCTGATTGCCGTCTCCGAAGGG\nnGCATCGCCTGAGCGAGCAGGCC  
AGCGTGACCTTCAAGGATCTCCTGCCGAGAATGTCTCGCCGTACG\nnGCGATGCTGGGGCGGTACGTGCGGCCACG  
CGGCGGCTGGGCGAAGAGTACAAGCCGCGCTACAGCGTCC\nnGGAGCGCGGGTGTGCGAGTCAGCCTGGTGAGGCCGGG  
TTAGGCGTCAACATTCAACCGGAATGCCTCGT\nnCAGTCATGAAGTTCAGCCGGGTTGCCGTGCTGGAATTGGCCGAA  
CCATGGGCGTTGCGACGCATTAT\nnATCGCCACGGCGCGCGGCCGAGACCAGAATCCGATTGTGCGTGCACTGATCAAG  
CAACTGCTGGACCGGC\nnCTCGCGACTGAGTGGCCGCGTCGCGCTGGATGGCCACGTTCCCGCACTTTCGACATCGAAA  
TCACAACG\nnACAGACAAGGAGAGGCCAGATGCAGCCATTGTGCTTGGCAGGTTGGAGATTGAGAAGGTTGTGAGATG  
\nGAAACGAGCCTTCCGATCGAAGTGTGTTTCCGGGCGTCACTGCAGGCGACCTTTCCAGCTGAAGGGCT\nnGGTACT  
GGGACGATGTGCTTTCTGCTCGATCCGGCTCAAGCGGGGTGCGGCCTGAGCATGCACAGTTACCT\nnCGTACGTTTCAGCG  
GAAAGACGATCCTGATCGACAGCTGCAACGGGAACGACAAGGATCGCTCGACCGAA\nnGCCTTTCATCGGCTGAATACGC  
CATGGCTCGACAGGCTGCGGAACTCGGCATTGCGCCGGAAGATGTCG\nnACATGGTCATGTGCACGCACCTGCACGTGG  
ATCATGTGCGGCTGGAACACGCGCCTCGACAACGGGCGCTG\nnGGTACCGACTTTCCGAATGCGCGCTATGTCTATGAACA  
GGCGCGATGTGACCATATCAGTTCGCCGAGC\nnACCCATGCGATGATTCGCGAGCCGTGGCTCGATTCCATCTCGCCGG  
TGATCGATGCCGGGCGAGCCGATC\nnTCGTGACGAGACGAGGTCAATCTCGGGAAGGTGGGTGACGGCATGTGGCTGC  
GGCCGGCGTATGGCCA\nnCTCGCCCGCAGCTGCATGATCCACGCTTGCTGCGGTGGTTTCGCCGCGGTATTTTCTGGCG

ACGTGGTG\nCACCATCCAGTTCAGCTGATCCGCCCCGATCTTCATTTTCGTGTACGACTGGAACGGCGATCTGGCGGAAC  
\nAGGTGCGCATCGCGCTGCTCGAGGAAATTGCCGGAACCGATACGATGCTGTTTCCGGCCCATTTCAGATA\nTGGATC  
GGGCGGGCATGTCTGAGCGGGACGGCGACAAGTATCGCTTTCGCTTCACCCAGGTCTGACGAGTC\n>gi|142022655|  
gb|EQ086233.1|221 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold,  
whole genome shotgun sequence\nGCCCGGGCGATTTGACGTCTGCAGCCTCACGTCCAAACTCAATTTGAGAT  
TCTTGACAGACACGATGGGCTA\nATTAAAGTAGGTGATCGCATTATGGTGCAACATACGATCACCTGCCGATGCCAGTCTG  
AGTCAAGTGTATA\nNGGGGTAAAAATAGTCAAAAATATGTAATTAAGGGGCGTGATATGTTTCAAGCCGCCAACCACAC  
CGCCA\nCCCAGCATCCCACCTAGCCCAGCGCCGCCCAATTCCCACCTGTGGCCAAGCTTGTGAGGGCCGATCCAA\nT  
ACCACCATAAACAGCACCAGATGCAATTGCATAAACCAATCCGCTTACGCCCCCCCCGATTCTGGTTTG\nGTTGGCCTT  
CTTGCTCCTCTCCCCCGATACTAGCAACAGTTCATAAGCGGTTAATCTTGAATGCAGCC\nAAATCTCCAATGAATTTT  
CAATTGCCGAAGAGTGATTATCATTTTTTTCTTCCGTGTATTTGCATAGTTAA\nAAAGAACTTCCGCCAAAACAACGGA  
TATAGCAGCTGCAAAAACATCTCTCAAAAATCGATCCTTCTAGAAT\nTTCAATTTTCCCAAAAATAAACCATATCGCAAC  
GCCAAGCAAAAAAATCAGAAGGAGTTTTGCGGCAACC\nNGCGTTAATGTCCATATCGATAATTTTTTATTTCTCCAGATC  
ACGCCAGAAGGATGAGCCGACACTGTCAT\nACGATATTCATGTGGAATAACAAAAATGTAAGCTCAGATCGTCGATTCT  
ATAATCGCGCCACGAAGTCA\nATACACGTTTTCTACCGTAGCACTCCGTATTAATGCTCGCAACCTCCTTTTGGAGGCA  
ATCCTATCAAGTT\nAAGCCGGAACCTTGTAACAAAGGCTCGCTCATTGACGATGCGCCTCAACGCCATCTTCAATAT  
TTCCG\nGATATGAAATTGGCTTGACCCCATCTTGGCCAAGCTCCTCGAAGGCCTTCATACCGCGCATCTGTGCACT\nC  
GCCATGTGACGATCCGCGTCGCTTCAGCCGACGCCGACGCTCGATTTTCGTCAACCTGAGCTGTTGGT\nTTCTAGCGC  
CCCGCGCGGGCCATACAGACCGAACTCGACTGTTTCTTCGCCATGCTGATCGATCAGGCT\nGAACTGATCCGCCGCGT  
CAGTGCGCAAGCCGTGTCCACGGCGGACAGCACTTGCCGTTCTCCGTATTCA\nCCGACCTCAATGATCAACTGCTCAC  
CAGCATCGAACAGTACTTGTTACAGCTCCGATGGCCCCGCGCTGCG\nGGTCGTGCGGGCAAAGGCTCCGACCTGCGCCT  
GACCTCGCGGGATGCCACGGACGATGATCACC GCC\nTGATCCGCGTCCGCGCGTGCTCACGACGAATCTGCTCGACG  
CGCACGCCTATTCGGCCGCCGACATTGCT\nNGCGCTTTAACGCCGCTCCAAGGTACGAAGTCGCTGGCATCGTTTCTCG  
ATCTGTTTAAACAGCGAGTTG\nCACGTCCGTTGCCCTTTCGACAAGGCTGACTTGCCGTGATTGTGCGCAATATTGACG  
GACCGCGATGTCT\nTGGTCGTGGGAGGTTTTCCGCAGAAAGGCGAAGCTAACGATGCTCCGTATTTTTGCTCGATGGGC  
CGGCC\nTGTCGATCGATTTTCTGGTCGGGAAAAAGAGAAGTTTTGCGTGCTCTATCGCTAACCTTCATTAGCGAT\nNA  
GGGTTACGTTTTGGACAAAACCTGCCATTTTGCAAAACGCGTCACATCCACCTGTACACCGCTGTCAT\nCCCCGAAGCC  
TATAACTTTTGGATGCCTGTTCTTCCGGACTCGCTCTCCTTTCCCGACGCCGCTTCCCTGG\nCCGTCTTGCGTACCTGG  
TACGCGAGCGTCGGCTCGCGCGAAGCCGTCGAGCGTTATTGTCCCGACCAGCT\nCGGAGAAGGAAAGTCGGCCCCGTGGT  
GTGACTGGTCGTATCCGACGGAACCTCGTCGAGTTTGCCATCAGT\nCGTCATCGTGACGCTCTGGCGAAACTGTTTCAG  
TGCGAAGCAGGCGAACGATTCTGTCACCGCAAAGCGG\nTGCGCCGGGGCGCTCGCCGCGTTGCATACGAACGGCATCAGG  
ACGCTGGCCGAGCTGATTGTGCGAATTTCT\nCCGGCGCCGTCGCTGGTGGCTCGGCATTCCCAGTCTCGGGGAGCGTAGC  
GCGCACCGAGTCGAAGCCTTC\nTTTGCGGCTTGTTCCGCGCTGACGGAGCGGGCTCGTTGCTGGTGGTGGCGACCACC  
CCGGAATCCGTGA\nCGCCGTGGGAAAAACATCCGCGTCCCCATGATGTGACGTTTCGCGCGCATCTCCGTGCGCCG  
AAGGC\nGATGTGCGCACGATCCCAATCAGGCGGAACGCTCCGGCGTACGCCAGGGTCATAACTGTTTTCCCCGGGA\nNG  
TAGCGAATGGCATCGAGCCAAGGCACAGTCGATTTTCATCGTCGAGCAGATGGCGGGCGCCGGTACGGTA\nTCGGGCGCGC  
AAGATGTTCCGGCAATACGGCATCTATTGCGACGGCAAGATGGTCGCGCTCGTCTGCGACG\nATCGGTTGTTCTGTCAG  
CCGACGGCCGAAGGCCGCGCTTCTCGGCACGTGCGAAGAAGGCCCGCGTA\nTCCGGCTGCGAAGCCGACCTCGTC  
ATCGCCGGCGAGCGCTGGGACGATCGGAATGGCTGTGCGCGCTG\nATCCGGATCACCGCCGCGCAGTTGCCGGTGCCG  
ATGAAGCGGCGCGGATAACGCTGATTTCGCGTGCCAT\nGTTGCGCTGCGGCCAGGCGTATTCTGTGACACATGTGCGCG  
ACCATCGCGCATCCGGCGGGAGCACAAGA\nATGAAGACATCCGGACATTCGCGCGGCACGTGGCTTGCTCGCCGCGCCC  
GGCGCGCTCGTC\n>gi|142022655|gb|EQ086233.1|294 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGATCAGC  
CCCGCATACGCGTACGCGCTGCGACGCCGAAGAGCGTCTGCTGCCGAGGATCTTGCGCAAC\nTCGAAGCCAATGCCT  
GCCATCAGCGGCCCTCGTTGCGCGAGCCGGATGCAGCCACGCCCGACCCGCTTG\nATGCACCGGGCAACCGGCTTCGA  
ACGGACGCCCGGACGATCCCGTCGGGGGCGATCGCGAGCGCATCG\nTACAACGCCCGATACCGGTCGATCATCTGTG  
CGCGCGTGTAGAAGCGTCGGACCCGGGAGAGCCGGCTT\nGCTGCGCGGCCTGCCAGCGCTGCGCGTCCGCGAGCAGGT

CGAGCGACGCCTGCGCGAACGCGGCAGGATT\nCGCGATCGGCACGACCGCACCGGCGGCCCCGAACGCGCGGTCTGTCTT  
CTCCGAGCCCTTCGATCAGTTGC\nCGGCACGACCCGACGTCCGTCTTAATCGACGGCAGCCGCGCGGAAGCCCTTCGA  
GCACGACGAGCGGGA\nGCGCCTCGCTGATCGACGTCAGCGCGACGTCGACTTTCCGGCAAATGTCGTGAATCCGCT  
GAAAGCC\nGAGAAAGCGCACCGTCGTGCGAGCCCGAGGCTCTGCGCGAGCGCCCGGCATTCCAGCGCATACGCGGGA\  
nTCCTCTTCTTCCGCCCCGACGATCCAGCCTTCGATTTCCGGCATCGTCCGCGATGCGATGAACATCGCGC\nnGGACGAA  
CGTCTTGATGTCCTTGATCGGCACCACACGGCCGATCAGCGCGACGATGCGGCGCGGGCCGCC\nnGTCCCGCTTGTCAG  
CAGCGGTGCGAAGCCGTCCACGTCGACGCCGTTTCGGAATGTTGCGGGTCCGGGCG\nnGCATCGGCCCCGTCAGCGATCTG  
ACGGTGTGATTGGCCTCATAGAGCGGACGATCTCGTCGGCCGCGT\nnCGTATGCGAGCTTGCCGAGCGCTTCAAAGAA  
ACGACCCACAGCGCGCGAAATAGCTGATCGTCGATAC\nnGTCCCGCTCGAACGCGCCCCGTTGTGCGGATCCACTC  
GCTCTGCAGCAGATCGATCTTGCGTTCCTTG\nnGTATAGATGCCGTGCTCGGAAACGAGCAGCGGTGCGCCCGTTTCGGTA  
ATGCAACAGCGCACCGAGAAAGC\nnCCGCATAACCCGTCGACACGGTGTGATAGACCGCGCCGGCGGACGCGCGGTGCG  
GATCCGCGCAAGCTG\nnCCACAGCGCTTGTGCATGATCCGAACCGTCCAGAAGTAGTCGGTGAAGGAAGGGTTCGGTGCA  
GTAGCGG\nnTGATAACGCTCGACGATGAAGTCCACGCGCGCGGTGGTGGAAAAGTCTGCTCGTTCAGCGGTCCGT\  
nCATCGCGGATCAGTTGGACCAGGTGCGCCATCAGTGCAGCCGGTTTCGCGCGACGCCGTCCGCGCTCAG\nnCCCTCGC  
CACGCATCGTGCAGCGCAGCCGAACGCTTGAACGCTTCGGGGTTCGCGGGAATCTGCCGTGCG\nnCCGCCCCGATCGGCA  
GGCAGTGCCTATACAGATAATGGGATTGGAAGTGAACGACGTTGTGCGGCAACG\nnCGTATGCGGCGCCATGATAGTCG  
TCTTCCCGGCTGCGGATGAACACGATCGCGAAACGCTTGTGCGGATA\nnTGACGCGATCATCTCGTTGACCCAACTCGAT  
ACGCCGCTCGCACATACGGGAACGTGCCTTCGAGCAGC\nnAGGCAGACGTGATCGCGGTGACGGACCGGCACGGCGG  
TGAATGGACGATTATGACGACCAGTAGCG\nnAACGACGGGATTGAGAACGGGCAGCGCGCCCGCTTGGCGAGCGACG  
CGAGCAGCTCGGCCGACGCGAA\nnTACTCGCCGCGCAGGAACGCGCCTCCGCAAGCCACGGCACGAGACGCTCGCGCG  
GAAAGCCCCATTGCA\nnGCGCGCGGCTCATGTATTGCACTGCTGCATCCGCGTGC CGGTGCGGAGTGGAGGCGGCCG  
GAACGAG\nnCCACATCGCCGCGTGCCTTCGTGATCGCGAGCGCGCCTGCGCGTGGCGGTGCGCCTGTCGCGCGTG\  
nTGCCCGTGCACCGCGCCGAGCAGGATTCTGATAGACCAGCTCGAAGTACAGCTCGGCAAGCCGCGCAT\nnGGGTGCG  
CTGGCGGTTGCGTCTGTTGGGCGGCGTGCAGCGTCTGCTGCGCGGGAAGATGTTCTGCGC\nnGATCTCGTTCTCGGC  
GCGGTGCGAGGTGCCGTATGCAATGAGCCGACGTCCTTCGAGAGGATCGGCGAGC\nnAGATCGCGCAGCAGCGTGCCGT  
CGTGCGGTGCGCATGCTCTGGATCGCCACCAGCGCGAAAGACGAT\nnCCGACGCGGCCACCCGCTATTCACGAGACG  
CGCCTGCAGACGCGCACCGCCGCGGTGCGAGACGCGCGA\nnCACGAGATAGGTGACGAATTCGGGCAAGGCACGTGCGC  
GTATGCGTCTGTCCTGCTTCGCGGGCCAG\nnCAAGCCGCCACACACCACTCGCCAGCACGACAAGGCCCGCCGAC  
GGGCACGAACGTGCATGCGAGCC\nnAGAGCGCCGCGAGCGTCCCGTTGCGCGCGTTTCGGTAACGTTGCGGCAGCAACTG  
CCGGTACAGCACCGC\nnCTGCGCGCGCGCCCTGCCACCATCGCGCGGAACACGGGCCACAGGGGTCAACCGCGGTGCG  
GGGAACG\nnGGCATCGCCGCCCGCCTGGGCGAACCCGCCGAACCGCGGAGCACGGCATACTGCAGCGCCGCGACGA\  
nTCGCAATGCGGACGAGCAGCGGAGTGC GGCCAGCATCGACGCGCGGACCCGTAACGCGTGGCGCCC\nnCGGACTG  
GGCGCCACGAACCGGATCGCAGCCTACACATGAACGCCGTTCCCTTCAGCAGGCGCTGCAA\nnTGCCGGCCCCGGTTC  
GAGGCTGCCAGATGCAGCGTATGCAGCCGATCCGCGCGCCTTCGAGATCGAGA\nnCCGAACGCGCATCGAGGCTCGC  
CTCGATGCGTGCAGATAGCCGTGACGCCGCGCTATCGGTTGCCG\nnGCATCAGGTTACGAGCACCGATTGCGTCGG  
CGTCCTCAGCGGCCATATGAGATCGAGCGAACGGCGCCC\nnGCGCACCATGCTCGAACAGCGAATCGTCCGCTCGTC  
GTGCTGGAACACGAGCGCAACCAGCGACGAC\nnGCGATCCCCGACTCGCGCTGCAGGCGGTGACGCGCGATGTCGAG  
CGGAATTCGTACGGGCATGCCG\nnGTACTGCCGCCACCAGTCGCGCACGCACGCCGATGCTCGACGCCGTCCGCGTA  
ATAGCCGAGCAGCAC\nnCAGCAGCAGTTGCAGATTGTGCAAGTTCAGCGACAGAAACGGCATGCGCTTGATCGCGAGCAA  
CGCGATC\nnAAGCGGCCGTGCGCGCACACGACCGGCGCGGACACGAGTGCCTATTTCGGCGCCAGTCTGCTGCTGC\  
nTTTTCGGATGGGCGACGATTTCTGTCGAGCGCTCGGGCCACGAGGCCGTGCGCGGATCGAAATCGAA\nnTGATCG  
CCGAGCCGCGCAACCGTTTCGCACAGATTCTGGGCGAGTTCGTGCGACCTTCGACCGGATAC\nnAGCGCCGCCACCTCG  
ATCTGACATGCCTGTGCGACGAATTCGAGCAGCGGTGCGCACCGGTAGCCGCA\nnCCTTGCTCGCCGCGCCCTGCGGC  
GGCGCATCGAATCCGTGCGGAGCGACAGGCGGCGCAATTCGGTGTAT\nnCGAATCCCGCAGCGTGGTGGCTTCGACAGC  
AGATCCTTTTCCAGCCGCTCGTGCAGAGGCGCATCAGG\nnTAGTGGCTCTTCGTGATCGCGACGAGCCGTGCTTCAGA  
TAGTCGTTGAGCG\n>gi|142022655|gb|EQ086233.1|323 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nACGCCCCG

GCGCACCGCGAGTACCGCGCCGCGGGCACTCCTTGACCCCGCATGATCGATTCCCGATGAAA\nCCCCGAAAACCTCGTC  
GCCTGCCACGAATGCGACCTGCTGTTTGGCGGCGCCGCGCTTGCGCGCGCTGG\nnCTGCGCACTGCCGAGGTGCCGT  
GCCCCGCTGGGCGGCAGCGCGCACGGCCGTCCGGCGCTCGACCGCG\nnGTGCGCGATCGCGCTCGCCGCGCTGTTACG  
CTCTTCATCGCGCAGGCCTTTCCCATCGTCGCGCTCGAC\nnGCCGCGGCATCGCATCGCACGCGACGCTGGCCGACGCG  
GTGGCCGCGTTGCGCTTGAACGGGCAACCGG\nnCGGTGGCGCGCATCGTGTTCTGCACGACGATGTTGTTCCCGCTGCTG  
GAACTCGCCGCGTGGCTGTACGT\nnGCTCGTACCGTTGCGCGCGGGCCGCGTACCGCCCCGCTTCGAGCCGGTCCTGCGC  
AACATGCAGCGGCTG\nnCGCCCGTGGAGCATGGTCGAGGTGTTCTGCTCGGCATCCTGGTCACGATCGTCAAGATGACG  
AGCCTCG\nnCGCACGTGATACCGGGCCCCGCGCTGTTTGCCTTCGCGCGCCCTACCGTGTTGCTCGGCTTTCTCGCGTC  
nATTGACCCGGGCGGCCTGTGGGAAGCGCGCGACGAAATCATCGCGCTGCGCGGCGGCGGTACGTCCGCC\nnGCGGTAT  
CGCGCCGGCGGCACACGCCGCGACGCGCTGCACCGGTGACGCCCGACACAGCGGACGCAACGA\nnACGCGACCCGGCGGA  
CCGGGCGCCACCGCTCGGCAAGCGTCACGGCCCCGCGCCGCGGGCTGGTCGCATG\nnTCATACCTGCGGACGCGTGCAAC  
CGCATACCGAAGCCGCCGACGCGCGCTGCACGCGCTGCGGTTTCGACG\nnCTGCACGAACGCCGCCCGCGCAGCGCCGCAC  
GCACGGGCGCGCTCGTCATCGCGGCCGCGCTTCTCTACA\nnTCCCGCGGAACCTGCTGCCCATCATGCATGCGACCTCGC  
TCGGCCGCGCCGAGGACGACACCATTTCTCGC\nnCGGCGTCGCGTACTTCTGGACGTCCGGCGACTGGCCGCTCGCCGTGCTG  
TCGTGTTCTGTCGCGAGCGTGCTC\nnGTGCCGATGCTCAAGCTCGCGATTCTCGCGTTGCAGGCCATTGCCGCGCATCGCG  
GCACGCCATGGCGCC\nnCGCTGGAACGCGCGCGGCTCCATCGACTGGTGAGCGGTGGGCCGCTGGTCGATGCTCGACG  
TGTTCTG\nnCGTCGCACTGACGATCGCGCTCGTGCAATTCGGTTCTGTTCCCGGAGATCACGCCCGGCCCGCGCACTG  
nGCATTCCGGCGCAGTGGTCGTGCTACCATGTGTGCGTCGATGCAGTTCGATCCGCGACTCATCTGGGACG\nnGCGCGCA  
CCGCTGCGCCGCCAGCCCTCGATCCTGAACGGGCGCAAGCCCGTCAACAACGCTACGTCCATT\nnACAAGATGCTTCCCC  
TCGATTTGCTGAACCCGAGATCCGGCCGCGCAGCCGCTGGATCCCGTCCCTCGT\nnCTGGATCGTGCCGCTCGTGTCG  
CGCTGATCGGGCTGGCGCTCGTGTAACCGGGCATTGCCGCGACCGG\nnCCGACGATACCGGTCACGTTCCGGAACGCAG  
AAGGCCTCGAAGCCGGAAGACCAAGGTCCGTTACAAGG\nnACGTCGATATCGGCTCGGTGCAGGCCATCACGCTCACGC  
CCGACTTCAAGCGCGTGGTCGTGCGCATCCA\nnGCTCACCAAGGACGCCGCGCAGTTCGGAACCCGCGATACGCGATTCT  
GGGTCTTCGGCCGCGTATCGGC\nnGCCACCGGTATCTCCGGGCTCGGCACGCTGCTCTCCGGCGGTACATCGGTGCGG  
ACCTCGGCCGCTCGC\nnCGCACGAACAGACCGACTTCGCCGACTCGAGACACCGCCCGCCGTCACCGCGACACGAGG  
GCCGCCG\nnCTACCTGCTGCACGGCGATTGCTCGGCTCGATCGACATCGGCTCGCCGGTGTCTATCGGCACCTGCAG  
nGTCCGCCAGGTGGTCGGCTTCTCGCTCGACAAGGATGGCACGGGCGTCGACGTGCAAGCGTTCTGTGAAGG\nnCGCCGTA  
CGACAAATACGTCGGCACCCAAACGCGCTGGTGGCACGCAAGCGGCTTGACCTGCGGCTCGA\nnTTCGAACGGCCTGAA  
GCTGAACACGCAATCGCTCGCGACGGTGGTCGTGCGCGGCCTCGCGTTTCAGTCG\nnCCGCCCGGCAAGCCGATGCACC  
GCCGGCAGCCGACAACGCGTCGTTCCGGCTCGCGCAGGACGAAGCCG\nnATGCGATGCGCGACCCCGACGGCACGCCGCT  
GATGGTCGTATGCGCTTCGACCAGTCGCTGCGCGGACT\nnGTCCGTCGGCGCGCCGGTCGATTTGCGCGGCATCGCGCT  
CGGTCAGGTACGAACATCGGGATCGAGTAC\nnAACGAGCACTCGCGCACGTTTCAGCATGAAGGTACGATGGCGCTGTA  
TCCGTCGAGACTGAGCCGGCACA\nnGCGACACCGCGCTGCCGGCCCCGATACGGCCGGCGGCCACGACCTCCTCGAGCA  
CCTGGTGCTGCAGGG\nnGCTGCGCGGGCAGTTGCGCACCGGCAGCTGCTCACGGGCCAGTTGTACGTGGCGCTCGACAT  
GTTCCCG\nnAAAGCGCCGCGCGGAGCGTGGACGTGCACCGCACGCCGGTCGAGCTGCCGACCGTTCCGAACACGCTCG  
nACGAACTGCAGGTTTCAGCTCGCCGACATCGCTCGGAAGCTCGACAAGGTTCCATTTCGACCAGATCGGCAA\nnCAGCCTG  
AACGGTGCGCTCGAGCATGCGAACCAGTTGTTTCGGACACCTCGACGAGCAAGTCGTGCCGCG\nnGCGCGGACACGCTG  
GCGGCGGCGCAACGCACGTTTCGACGCCGCGCAGGCGACGCTGCGGCAGGATTCGC\nnCGATGCAATCGGACGTTTCATGAC  
GCGATGCAATCGCTCACGCGACGCTCCAGTCGCTCAATACGCTGGC\nnCGACTATCTCGAGCGGCATCCGGAAGCGCTG  
CTCTTCGGCAAGAAAGGAGAACCGAAATGACGCCGCAATT\nnCGTTTCGCACGTTCCGGATGCCGATCCGCATCGCGACGT  
GCATCGCGCTGGCCGTGCTCGGCGCTGTAC\nnGTGCGCGCCGTACGGTTCATACGCTCGGGATGGCGGATGGCGCGG  
GCGGCGACACCGACGCTCGCGT\nnCCCGCATGGCTGATCGACATGCAGCGCGTGACGTGGCGGCGCCGGCGGACGGCA  
ACCGGCTCGCGGTGC\nnAGCGCGGCCCGAACGGGTGCACATCCTGGAACAGGAGCGCTGGGTTGCGCCGCTCGGCGACG  
AGATGCG\nnCGACGGACTGTGACGCGCGCTCACGTCCCGGCTGAACACGTTTCGACGTTACCGCGCTCGCTCATCCGGAT  
nGGCACGCCGGTCTATCGGGTCGCCGTGGACGTCCAGCGTGTGCAATCGTGGCCGGCGTCTCACGTGCTGC\nnTCGATGC  
GACGTGGACGGTGGACGCCGGCTCAGGACAGCCGGCACTGACTTGGCGCAGCATCGTTCCGGG\nnCGGTGCGTCGGCGGG  
CTACGACGCGCTCGTCGACGCGCATCGCCATGCGCTCGACACGCTCGCGCTCGGC\nnATCGCCCGCGCATTCGCGCGGC

GGCCACGCACGCGCCGCGCCGCGCGGATGCCGCGCGGGAACACCCG\nnACCGACGGATCGCGCCAGCCACGCGCGAACG  
CGGCGCGCCGGATCCGCTCATCGATTGAATCGGGCAGGG\nnAAAACGCGCGCGCGGGCCGGGTGGGTGGGTGGGCGA  
TCGGCCTCGCGCCGGTGCCCTGCCCCACCCA\nnGCGACATCGTCAGGCGAAACCGTCCCGACGAACCACGAATCGGCGCG  
GCGGCGTTCCCGCGTCGATCCGG\nnTATTTGCGCGCGCGCGGTGTTGCGCCGCTTCGATCATTTCCACCTGCATGGACCA  
CCCTGGATGAAACCC\nnGCCCCGACTCACGCGCGAGCAACGCAGGCTCGACACCCGCGAACGGCTCATCGCTACCGCTCGC  
GCACTAT\nnTCATCGCGAAAGGCTTTGCCGATACGAGCGTCGAACATATCGCCGAGGGCGCCGGGTACACGCGCGGCGC\  
nGTTCTATGCGCACTTCGAACACAAGCGCGAGCTGTTGATCGAGATCCTTCGCCGCGACCATGATCGAATG\nnCTCACGA  
AAGGTCGGCTAGTGGCACCAGGTGAGCGGCAGCGCTGCACCACGGCCGACTACGCGGAGATTG\nnCCGCCGGCTGGGAAT  
GTTTCCCGCTGTGGGTGAGGTTGATCTGTACGCGCTGCGCGATGCCGGGTTTCG\nnGCACATCGTCGACTGCCTGCACA  
CCGAACCCGCGCGCTACCGGAGACCACAAGCGGCACGGCGGGCGCC\nnACCGATCGGATGCCGTCCGCGCGCGGCTGGG  
CGGCCGCGCTCGGCGTCGCACTGCTGAGCACCGGCGCGC\nnGGGATGCCCGCGCGACGATCCGGCGTAGCAAGCAGCCA  
CGCGAGCGCCGTTCCGCCCATCGCCCCGGC\nnTATAATTTCAATCGTTCCCCGACGAGTATCCGCACAATGACGGACGC  
CAACCTTCCTCACGTGCTGGTAG\nnTCGACGACGATCCGGCAATCCGCGAACTGCTGTCCGGCTATCTGTCCGACAACGA  
CATCCGCGTGACGGT\nnAGCCGATTCCGGCCAGGAAATGGCGGGCGCGCTGGACACCTTTGCAATCGATCTCGTGATTCT  
CGACCTG\nnCGCATGCCGGGAGAAGACGGCGTGAGATCGCCCGGAGCTGCGC\nn>gi|142022655|gb|EQ08623  
3.1|564 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
shotgun sequence\nnCTGGTAGCCATGCAGCAAGGCTGGCGCTAGATGTACGGCCAGATTGGCGAGCTGGCCAATAA  
CGCGGAA\nnCGCTGTGACATCAATCGACACTCTGAGCGAAAGCGCTCAACAAGCCGACAGCGCCGCATCCAGATGGC  
\nGAGCGGGATGCAACGGTTCGGCGCCAGATCCGGCATATTCCTCGTAAGCCGCCTAAGTGGAGGACGCG\nnGTGCGG  
CGGAAGAACGGCCGGGGCCGGTGGTGGCGTGATCGAGCAGACCATCGCCAGCATCTACGCCATC\nnGTCAACGTGATCAA  
GGACATCGCCGATCAGACCAGTCTGCTGGCGTTCAACGCCTGTATCGAGGCGGCGC\nnGCGCTGGCGATCAGGGGAGCGG  
CCTCGCCGTGGTGGCCGACGTAGTGCGCAATTGGCGGAGCGCACCCG\nnCCAATCCGCCGACGACATCGCTCGACCGT  
CATATGCGCCAGTTAATGGCTACGGCGTCCGACGGCATGC\nnAGTAAGACGGCACCGCGCCCGCCAGATCGGCGCCGAC  
GCGCGGCCGCGCTGCCATCGCCG\nnGCCAGGTGAACCGAATCGCCGGTGGCAGCCGCGAAAGCAGCAAGGCTGCCCGGC  
GACACTGCCGAGGCGG\nnCGCGCTTCTGGATCGGCTAGTGATGAAATGAAGCAGATATGGCGCGAAACCCCTCTGC  
ATAAACAA\nnCAAGGCCCGCTTTCGGCAAGGAAAGCGGGCCTTATGCCGATCTGGATAGGGTCATGGGCCAGCGCCTGC  
\nAGCGACTCGCGCATGATCTGGCGGTGCTCGTCCAATCTCGATCAATCGGCCTCGACGCCGATCTTGCGC\nnAACTGG  
ATGGCGGCCGACATGCCGAAAGCCACCGCCACCAACAGAACCTCATGCGCTACGGCCATGGT\nnGGGCTGCTCTTCAT  
TGCAATTGTTAAATCCAGGCGAACAACTCGCCTGTATCTTGTGATTTCAGTGCC\nnACCGGCTGCAGGCTTTGCCCA  
GGCAAGGCCAAAGACGCATCGCCCGTCTTCGGCCGGAATTGCGCCC\nnCGTGGGCGAAGCAGACGATGCGACTGCCGT  
CGCCGTTAGATAAGCATGCTTGCGCCAGGCCATCGGGGC\nnCCCGGCCATATGCGGACAGGCATTCCGATAGCCGTACA  
GCGTTCCGCCCTTCGCCACCACGAACAAGCTG\nnTCACGGCCGCGCCGTCCGGGTGAAACCACGGGCGGCGCCATCGG  
CGATTTTCATCGCGACGGCCAGTA\nnGAACGCGGCGCGGCACGGCACTCAGTGCTTGCCACCAGGCGCGGGCCGACGG  
CGCCCACTTTTCGCGA\nnTGCTGGAACAGGAACAGCTGCGAGGCATCGGCGCCATCGGCGCTTCGCGCGGGCCCACTG  
CTCGTCGT\nnGGAGATCCATATCGGCGTCTGATTTCGATATGGCAGCCGAGCGGGCTGTTGAAGTACCAGAACCAGTTGGA  
\nGCCGAACCTTGTGGCGGCGCCGCCCCAGAACGACTGGTAGCCCTTGTTACCAGACGCGTGCCGGCTGC\nnATTACC  
TCGGTGGGGCCGCCCATGTGGAAGGTGAAGTGTTGCGAGCCCTGCAT\nn>gi|142022655|gb|EQ086233.1|521  
marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequence\nnCGTTGTTGCCAGGTCGTCCGCATAGCCGGCCGAGCTGAACTGCGTGACATACGGGCGAATCTTCTTCAG  
\nCACGTGCTAGGACAGCCTGGTAGTCGGCCGGGTTCTTGCTCTCCGGGTCTTGCCCATGTACTGCAGCGCC\nnGCCGCG  
AACGCGGCGTCCGGCGCGTCGAGCAGCGAGATGCCGCAGCCCTTCAGTTCGCCGCGTTCCGG\nnGGTCGAACAGCAGC  
GCCCAGCTGTGCGTTCGTTGTCGCGGAGCCGCTTCCTGACCGCTCGACGTT\nnGTAGCCGATCCCCGTCGTGCC  
CACGCCCACGGCACACCGTACTGTTGCCCGGGTCGGCCTTCGCGACC\nnATCTTCATCAGCACCGGTCGAGGTTTCGCG  
AGATTGGCAGCTTCGACTTGTGAGCTTCTGGTACACGC\nnCGGCCTGGATCTGCTGCGCGAGATAGTTCGACGTGCGC  
ACGACGATGTCTAGCCCCGAGCTGCCCGCAAG\nnAGCTTCGTCTGACGCGTATCG\nn>gi|142022655|gb|EQ086  
233.1|455 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole

genome shotgun sequence\nTTCGAGTCGCTCGACGCGCCTCGACGGGCAACTAGCCGCCGTTGGCCAGCCTTGA  
ACGGATCTCGTCCGT\nCGAGAGGTGCGCGAGATGCGTGGCAATCTGCCAAAGATTCCCCACGGATCCTTCACCATCGC  
CCGCCTG\nnTCGCCGTACGGCATGTGCGCCGGGGCTTCGAGCGCGTTCGCGTGTGCTGCAAGCGCCCGCGATAGGTCA\  
nGGTCGGCATCCTCGACATACAGTACAGAAATGCCGGCATCGGCTCACGAGGCCGTCTTCCCCGTGAT\nnCATCAAC  
ACGGAATCACCGATCCGGATTTCCGCAGGCAAGCCCGGGCGGTATTCCGCCCTGCGCATGAAAA\nnACCGTCCGGATGAAT  
CCGATCAGATTGTGCGGATCCGGCACGACGATCCGGGGCGTGACGGTGTGCCATC\nnCATCGGGCACGAACCTGGTCATG  
TCGGGCAGCCTCCGAAGAGATTTCCGGGGGTATTGTGCGCCGGGGCGC\nnCCTCTGTGCGCAAAGGGGGCGACGCAACGCG  
GCAAGCGCCGTCAATCTCCGGGCGATCCACGCACCGCGG\nnGCGCGAGTCGCCATCCCCCGCATCGAGCGACCGCGCCA  
CGGAAACCCGAGATCGTCGATTGCAAAGGT\nnCGGATGACTGCGACGGCGATTTCGTGCGAAGGCAGCCCCGTTGACAT  
CGGCCAGCCCCCG\nn>gi|142022655|gb|EQ086233.1|229 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGGGTTCG  
GCGTCTACTGGGGCGGCTACTGCAACCACGGCAACGTGCTGTTCCCCACGTGGCACCGGGCT\nnTACCTGCTGCGGCTG  
GAACAGGCGCTGCAGAGCATCCCCGGTTGCGAGCAGGTGATGCTGCCGTTCTGGG\nnACGAAACCAGCGAGGAGTCGCTG  
ACGAAAGGCATTCCGTGGGCGCTGACCGCCCCGATTTCGAGCTCGA\nnCGGCGAAAAGATTCCGAACCGCTGGCCTCC  
TTCACGTTCAACCGGGCGATCACCGACAACATCAATGGC\nnGACAAATCCGAACTACAGCAAGCCGCTGCCTTACGTCACG  
GTTTCGCTACCCGTTGTGCGGGCTCGTGGGGA\nnCCGACAAGGACAAGGCCGCCACCGAGAAGCACACGCCAGGTTCCAG  
GACGCCAAGGAGAACGTGCTCAT\nnACTGAACCGCAACATCGTGGACTGGCTCACGTGCTACATCGTCTGAAGGGCGAC  
GTGGTGGCGACCAAC\nnGTGAAGAGCAAGTTCGAGCAGTGCTGGACGCGCCCACTACACGGTGTTCGAAACACGTGCG  
TCCGCGG\nnCCCAGTGAACGAGAATCTTCCGACGGCGGATCCCCGTGGTGCCGCTCGAGTCGCCGACAAACGATAT\  
nTCACTTGCCGTGCGCGGATTTCGACGTGCCGCACGGACCGAACAAGGGCGATTTCGCGCCGATTCCGGGC\nnGCCAACG  
GCGACATGGGCGAAAACGACACGGCGGGCCTCGACCCGATCTTCTATTTCACCACTGCTTCA\nnTCGATCGCGTGTCT  
GGCTGTGGCAGAAGCGGCATGGTTACACGCAGCACCTCGACGTCATCGCCGAGTA\nnTCCGGGCACCAACTCCGTCGATG  
CGCAGGGGGCGACGCCGGGCGTCTGCCCCAACTCGTGGCTCACGCTC\nnGAAACGCCGCTCGATCCGTTCAAGAAGAGCG  
AGAACGGCAAGGAGCGCGCTACACGTGCTCGACTGCA\nnTCAACATCGAGGAACAGCTCGGCTACACGTACGGCCCCG  
GTTTCGCTGGAAAACCTGCCGAAGTTGTGCT\nnCGCGGCACCGACCGTGCCGGCCGGCAACAGCCGCAAGGTGCTGCGCG  
TGTCGGGGCTGAATCGTGCCCG\nnATCGCCGGTCTGTTCTCGTTTCCGCGTACGTCAACGTGGACGGCGAACGGCAAC  
TGCTCGGCACGGAGG\nnCCGTGCTGAGCCGCTGGAGCGTGACGTCTGCGCGAACTGCCAGACGCACCTCGAAGTGAAGG  
CGTTCTT\nnCCCCTGAATCACTTCGCCGAAAGCGCGGTGCAGGATGCGCAGTACGACGTGAAATCCATACGCGCGAC\  
nGGCGTGCGCCTGCAGACGCCACCCGCGTCCCGCCCCGACTCGCAAGCCGTGCTCGGCGCGGCGCCGCAAC\nnCCGTACG  
GAAACTGTTCCGCCTCGAGGTCCGTAGCGGCGGCGCTGACGAACCGGCACCGTCGCCCGGC\nnGCCGGTTACCTCGG  
TACTTCGACAACGTGACATGACGGGGAGGCGCGGGCATGAGCTACACACCCGAAA\nnCCGGCAGCCTGGTGGGGCTGTGG  
ACCTATCGCAGTGTCTGAACGAGCCCCGACCCGGCCACCGCATTCAA\nnCGACCTTGAATTGCGCCTCGGCACGATCGAA  
ATCGCGCAGGCGCCGGCGGGCATTTTCAAGGGTCGCATT\nnTTCGTCGGGATGGGAGCTGCAGCTCAACGGCTGGATC  
AGCTACGGGAATCCGGGACGGTGCATTCC\nnAGGGGCGCGGTGTGGTCAGCGGTGAGGAATGGGTATACGACTATGTC  
GGGTACGTGAGCGCGCCCTGGCC\nnCAACGGCATCGACAGCGCCCGCGCTGACCGGTTTCGATCGTCCGTACGGTGCCG  
CATGCGAGCGGCAGC\nnGGCGGCGTCCGCGCCGGCGGGGTGGTGTGCTCGTGGTATGCGGTATGCGAGATCCCGCCTGA  
GCGGCCA\nnTGACTGCGCCGACTGCACCGCCGGCGGACACGGCCAGGCGTCCGGCGACGGTGACGGACGCGATTTCGCGC\  
nAGGCGTTGGGATCGGGCCTGCTCGAATGGCTGTTATTTGCTATCCGAACGGATATGCGGCGCGATGGCA\nnTTCGCAC  
TCCTCGGCTACCTGGCTTGCGACTTGCCGACCCGGCTGCACGGTGTCTCTCCAGCTGGATAT\nnGGAACGGCATCGGCG  
CAATGGTTTTCGGCATCGCCGGCTATTTCGATGCAAGGCGGCCTGTCCGGCGGGTG\nnCGCGCGCAGGGCAATGCTCCTGT  
GCACGGGGGCCATCGGCGCGGATTGACGCTCGATGCGATGCGGTCC\nnCCCGTCGATGCCATTCTCGACATCTGCGGCG  
GAACGCTCGATGCCCCGCGCATGTGGAACACACTCGTGC\nnTGCACTCTCCAATGGTTCCCGATGTGATGCTGGCCATGC  
TGGCGCTGCTGATCTTGCGCGAGGCCGACCG\nnCAAGGACCGTCGCGGACGCGGCGCGGCAAGCATGCTCGTTGTGGTGC  
CGGTGCGCGCTCGCGCTCGAATTC\nnGCGGGGATGCAGCTCGTCATGGCGCTCGGCATGACCGCGATCAGGGCATCGGCGT  
TCGCTGCGGGGCTGC\nnGTTGGGATACGAGCGGCGTGGCAGTGTGATGCTCGCCAGCATGCTGGCGTTTCGACGATTGA  
GCGATCG\nnGGCGGCACGCGCAGGGGGCGTTGAATGGTTCGATTGCGAAATAAACACGGCCATATTGAAGAGGCGCGGC\  
nCATCCGTTGTGACAGAGGCCGCGTTCAAGCTGCTCTGAAAAAGTGTGATGACGCGGTTTTCGCTCGGCT\nnGCCGGCG

CGCTTTGGCAGCGGCAGCCGAGTTTAAACCCGCGTCACACGGACGCTGCACGGACTCCAGGCG\nGCGGCTCGTCACTATT  
CACGATGCTCGTGCTCGTGCTCGTGCTCGCGCCACGGATCACGATCGCGCGGATG\nACGATGCCTCCACTCGTCGCGTT  
CCCAGTAGCGATGGCCGTCCCAGTAGCGATCGCCGTGCATGCCGATC\nTCGACCGATACACCGCCCCCTGCGGCGGGCT  
CTCGCTCTTCGTATCTCGGTTGCGGTGAACC\n>gi|142022655|gb|EQ086233.1|422 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ce\nAGCGCGTAATAGAACTGCCGTAGGTGCGGTTGTAATAGCGGATGCGCGGCCGATCGCGCCGTGCAGGT\nnCGTG  
CAGCCGCTCGATCGGCAGGTTCCGCATCATGTGGTGCTCCACGTGATAGTTGTTGCCGTTCTGTGAA\nnCCACGTCATGAA  
CGCATTGCTCTCGATCGTCCGCGTGTGCGGAACGGGTGCGTGCTCGTCAGGTGCGAG\nnCGGTAATGCTCGGGCATCTC  
GACCAGCGCATGCACGGGCGCGACACGAACACGAGCGGCACGACCCACA\nnCCCACACGGCGAACACGTATGCAGCGC  
GATCGACACGCCGGCCAGCAGCGCATCGCCGCGACCATCAG\nnCAGGTGATCGCGCGGATCGCACGCGACACGAGCGG  
GTTGTGCGCGAAGCTTGCGCCCGGCACGACGAGC\nnTTCGCGACGTTCTTCGCGAACTGGATATAGTGCGCGGCCATCGA  
CAGCCGCCAGGCCACAGGCCGAGAT\nnTGACGACAGGGCTCGCGCCGTAAGTGGTCGCGCTAGTCGAAGAACTCCTTGTT  
CTCGGGCGTGCCGAGCAG\nnCCGGTGTTGGCGCAGGTGGCTGTCTGGTACGCATGGAACGACACGAGCATCGGCATGCC  
GAGCAGCACG\nnCCGAACACCATGTTGAGCCGCTTGCTCCTGAATCCCTGGTAATGCAGCGCCTGGTGTTGCAGCTCGAC  
CC\nnCATGCGCGTACATCGCGCCGATCAGCAGGATGCCGACGATCTTCAGGACGACATACGTCGACAGCGCGAA\nnCAGC  
ATGCCGTGCGCGATGAGCGCGACGTAGACCACGAGCTTCAGCGAGAAGACGACGCGCGCTCGGGCGG\nnCTCTTCATGTCT  
GAAACCGGTTTCTGACGGGCGAGAAGCGATGAGGTGACGGACAAAAATGCTCTCCTTG\nnCCGGACCCTGCCCGGCACGA  
TGACGCGGCGACACTGGCGCGTGCCGGGTTGCAATTCTTCTGTGAGTCG\nnATGCTTCGGACAAACGAATTTTCCGAAC  
GCATTGCATGACGCCCGCGCATGAGCACACCATCGTGAAC\nnTAGACGACCGGTGCAATGCGTACCGACCATTCACCT  
CGCGCGGGCTCGCACGTATACGCACATGGCAGTG\nnCCGAGTGATGCGTGCTCATGCATAGCCTGCAAAATATTCCG  
TTTGTGCGGTATGTGATCCGGCAA\nnGAATCCCCCTCACTCGACGAACGGCGCCGCCGGCACCGTGGCGCGAGCGGCC  
GTGTTCTCAAGGGGAGA\nnATCCTCGTGACGAAGTGCGCATCCACCATTTCATCGCTCATCGAATCGCTCGAGGACGTC  
GCGGGCGCCA\nnTGCATCGCATGACGTTGTCGACGACGCGCGCAACGAGTCGAGCATCACGTACCGCGCTTCGCCGAA  
GA\nnGGTGTTTCGCCAGGCCGGCGCACTGCGCGAACTCGATGTGCGCGAGAACGACCTGGTGATGCTGGCGTTG\nnCCCG  
CATCGGTGAGCACGCGGCGCGCATGATGGCCTGCGTGATGACGGGCGCACTGCCCTGCACCGTGC\nnCCGTTCCCGCGC  
GCCGCGCCACCACGGGCGCGCAGGTGATCGACGTGCGATGCGAGCTGTACCGGCCGCG\nnCCTCGTCATCGCCGCGGATG  
CGCAGTCCGCCGCTGGCGCGACGACGCGTTTCCCGCGTCTGCCACGCGC\nnGTCTGTCGATCTCGACACGCTGTGCTCG  
TCGCCGATGCCGGTGCGCGCGCGTGATCAGCGCCAAGCATG\nnGCCGCGATCCGCATCACGTCCAGTTCAGTCCGGCT  
CGACGTGCGATCCGA\n>gi|142022655|gb|EQ086233.1|384 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCTTCCGA  
TCGCATGTACGCCGGCGCGACATTTTCGACGAACGCGTCGAGCGACACGTGAAACATGGGCAC\nnGTCCGAGCGCTGCTG  
GAGCGCCCTCGCCAGGCTGCTCTTCCCTGCACTGGTCGGGCCGTGCAATACGACA\nnGCAACGGGTTTCATGGCGATGGT  
CCGCGGGAGAGATTGCCGTTGTCCGGGTGCAGCGCTAGATCAGCTC\nnGGCGACGCGCACGATGACGAGCGCGGCCGGA  
ACGAGCACCGCCTGCGCACACACCGTGCCGAGCACGCGT\nnGCGCGGCGAGTGTCGTGATGGCCCGCGAAAGCTGTTT  
TCCGAGATCCGCCCTTCGATGACGTCGTGCG\nnTCATGACGGAACCTGCGGGTCGATGACGACCGCCAGCACGACGGTC  
GCGAAACCGTTGATGACGGACGA\nnCAGGTTGCGCATGTCACTCTCAGCTCGGGCTTCAGGTAGCCCGGTACAGCGAC  
GCGAATACGCCGACC\nnGTCCAGATCGCCATCGCGGCGACGTTCAACGCAATGACGCGCCACGACACGCCGTTGTTGCGG  
GTCAGTT\nnGCGTGACGTTCTGGCTCGAAGGCAGTCGCGCGCCGACGCGCAGATAGTTGATCCCGCGCGACTGCACGC  
nATGCAGCAGCAGGCGGGGAATCGAGCGGTTACCTGGAAGTGTTTCGACGGCGCGACTGAAATAGCGCTGA\nnAATGTGCG  
GGATCAGGAACGACCCGACGATGGTCGCCACGGTGGCGGACAGCAGGAACACCGGAAGTCGG\nnCCAGCAACCCGCTCC  
CCATCCGGTGCGCGATGTCCATTTCAACGCGTTTCGCGATGAACGGCCCCCTGAAA\nnGGAATTGGCCGTTCCGGACACCA  
GCGCGATGATCCCGAACAATGAAAACGACAACGCGATGCGACGCGTG\nnCGCACGCCGCGATCCGCACGGCATACGCCA  
GTGTCGCGATCATATGGATGACGAAGGTAAGCCCGCAGA\nnTGATCCAAAGCTGACTGTCCATGTGGTGCCGCGTGGT  
GACGGGATCACGAGTGTAACCGAGCGCGCG\nnCGGCGCATCGGCTGGCGAAGCTGCCTTCGCAAGCGTCACGGTTGGG  
CGACGAGCGTCAGGCGCTACCGT\nnGCGGCGCAATCCCCGTGTTTCATCCAGCATCCATGTGCGGAAGGCTGCAACGG  
CCCGTTGGCCAACTC\nnGATCGGTTCCGGCAGCACGAGGCAGAAAGTTTTTCGTACCGTCTTTCCCGCGGCCACGGTGC  
GACCAGC\nnCGGCCCTGCTCGAGCTCCGCATCGATGTACAGGCGCGCACACGAGGCCGCCAGCGCCG\

nCTTCGATCAGCATCGAATGAAGGTCGTAGCGCGCGCCACCGCCGAGTTGGTCAGCACGATGCCGACGC\nnCTGCGCA  
TAAGCCTGCCATGCGTCCGGTTCTGTGCGCGGTGAAGACGCGGCAATGCATCCAGCGACGGG\nnTTCTTCCCGCGCGCG  
GCGAGCAGCGCCGACTGCAGACGGGACCCAGCACTTCCTCCAGCAGCGGATGCA\nnGATGCATCCCCACCCAGGCCGGA  
TGCTCGAAATGAATCGCTGCGTCGAAACCGCTCCCGGCCAGCAGGAA\nnCGGCTCCATCCGCTCGGCGATATGCACGGTG  
ATGTTCCGGATGCCGGTCTGAAAAATGTTTCAGCCGGGGG\nnATCAACCAGCGGGTGGCGAAGGTCGGGATCGCGGCGATA  
TCGAGGCTCGCGCTTCGATCGGCTGCCCCA\nnTCAGGTACAGGCTGTCCCGTTCGAGCCGATCCAGAATCTCGCGAACC  
TGCACCGCATAGCGCGCACCGTT\nnCGGCGCCAGCCGACCCGGTTGCCGATTGCTCGAAGAGCGTGACGCCAGGAAC  
GCCTCCAGCCGGCCG\nnATCTGACGGCTCACGGCACCTCGGTGCGTGCCAGCTCCTCGGCGGCTCGGGCAAAGCTGCCG  
TGTCTGG\nnCCGCGGCTTCGAACACCTGGAGTGCGGAATTGCTTGAATCTTGCGGGACATGGAAGTCTCGTGGGTGCGG\  
nGATATGTTCCGCATCAGCTTGAGTAAATGTCACTGAAGGCTGACTTTATTTCTGTTTTATGGCGCGGAATG\nnGCGAGCT  
TAGCATTACGTTAGCGCACTGAACATCACATGTTCCGGATGGATGCGCACCTGCCCAAAGG\nnACAAACCCATGATCT  
ACACCGTGAATGCAGCTTCGCCGACCTCGACAGCGAAGCCGAATGGAACGACTT\nnCTACAGCCTTGAAAAGCTGCCCG  
CCCTGATCTCGGTACGCGGGTTTCACAGTTCGACGCGGTTCAAGGCG\nnATCACCCGCGGTTGCCCGGTCTATCTGGCGA  
TCCACACGATCGACGGCTCGATGTCTGACCGGCGACG\nnAATATCGCCGAAAGGCGGCGGCAATTCGCGAAATGGC  
AACGGCACATCACCGACTGGCACCGGAATCT\nnCTACAGCGATATCGGCTTCGCGCCGGCGGTAAAGGACGGTGAACATC  
TCGCACTGAGCGCCGATGGCCCG\nnGATTCGCTGCTTCGATTGGGCCTCGAACCGCTGGCCATGCAGGCCGTTGCGCTGG  
AGAAGTTTCCGGCGC\nnGCCGTTGGCTCGCCGTCGTGCCACGGAATAGCGCGCGGCTTGTCGAGGCGCTCCCCGAAGGTA  
TTCACCT\nnTTACGCACCGATGACGGAGCAGTTGACGAGCACACGCGCTTTCGATTGCGCAGGAGTAGACCGCGATG\  
nCCGAACGTCACGATTTCCGTTCGACGGAAGCGGATGCCGGCTGATGAAAGCCTTGCCGAGCTGTGCGCGG\nnACTGCGT  
CGAGCTCTGCACGCAGGTGCTTGAGGCTGAACTCAAGAACGTCCACGTCATTTTCTGGCGGT\nnTCGGCACGGGCATGG  
ACGTCCCGTCTTCGCCGACATCCGTATCGCGTCGGCACGCGTCGCACGCCGGAG\nnGTCATGAACCGTTTCATGGAAGC  
GCTGGATCAGGCGATCGTTCCGCCGACCGGCTTCACGGCGCGCATT\nnGGTGTTTCGTTACACCGCTCGAACCTTCA  
CGCCCGCAATTAGCCGATTCCGAGCAACACGATGTCCAC\nnCTTACCTTTCCCGTTTCAGCAAGTCGGCGACTTCACGAT  
TACCGCGATCAGCGATGGATACCTCACTGCG\nnAGCCTCGACTTCTGTGCAATATCGATTGCGATGACGCATCAAAAAAT  
GCAGCGTGATGCGGGGCGAGAAAG\nnAGCCGGCGGCGGTTTCATATCAATTGCTACGTCGTACGCGGAGCGGGCCACACCGT  
GCTCATCGACGGCGG\nnGGCTGGCGGGTTCAAGCAATGGGGCGGCCAACTCCGGACCAACCTGGCGCTTGCCGGTATCGA  
ACCCGCT\nnGCGATCGACACCATCTGCTTACCCACGCGCATCCCGATCAGTCGGCGGGCTGGTGAATGATGCAGGAC\  
nAGGTGCGGTTTCCACATGCCGAGCTGGTTCGTCATCAGCAAGAGGTCAAGTTCTGGCAGGACGACGGAAA\nnTCTCGGT  
CGCGCCAGCGAGCGGGCCCGCGCAACTTTGCGAAGGCGGTCAGGTGTTGACGCTTATGTC\nnGACAGGCTTCGCATG  
TTCGACGATGGACAAGTG\nn>gi|142022655|gb|EQ086233.1|280 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnGCAGCGA  
CGAGCCGCCCCAGATCTTGATCCCGCACGCGCAGCTCGCCGGCCCGACGTCGACCGTGCAGCT\nnGAGCGGCTCCATCGC  
CGCATGCGCGAGATACGGAATTCTGAATCCTGCCGAATGCGCGTGCCGCGCGGCG\nnGGCGCGCTGCCGGCGTCGGCCGT  
GCGCACACGGTCCCGGCGTGCCGGCAGCTGCTTGTACGCCTGCA\nnTCTGTTCCGATGACGTGACGGTCGAGCCGGC  
ATCCTTCCACGTGATCTTCAGCGCGTCGCGGCCGAGCTT\nnGGCCGGCCAGTAGCCGTTCCGATCAGGCCACGCCGGT  
GCCGCCGCGATCGGTGCGGATCTCGACGACC\nnTCGACCACGCCCTTGACGGCACGGGCCGCCGCCGCGTCTAGCTCGC  
GACCTTGCCGCCGAAGCGCGGCG\nnGGCGCGGACACGGCCACCATCGTGTCCGGCAGACGCCAGTCCATCCCGTACTT  
GAGCGTGCCGTCGAG\nnATCCCGCGCGCGTCGATGCGCGCGCTCGGCTTGCCGATGATGCGGAACGGTCCGGCGACTT  
CAACGAC\nnACCTGCTGCGGCACCGGCATCGCCATCGCGGCCGTTGCGAGTTCGCCGTAGGTGCGCGGTAGTTGCCGG\  
nAGGTACAGACACCGTTCCCGTCCGGCACGTGGCGGGATCGACGTTCCAGCGCTGCGCGGCCGCCGCGAC\nnGAGCATC  
GCGCGGGCGGATGCGCCGAGTTCCGGTATTGCGCGAACGAGTGATTGACCGACGTGCAACCG\nnGCCGTATCTGGATG  
CCCGTACCCGATCCTTGTAAGGCTCGCCGGCCGGCGGAGCACGGTGCGCACGT\nnTGCGCCAGTCGGCATCGAGTTTCG  
TCGGCCAGCGCCATCGGCAGCGCGGTGCTGACGCCCTGGCCGACTC\nnCGTCCGGTTGACGGCAACGGTCACGGTGTG  
TCCGGCGCGATGATCAGGAATGCTTGCGGAGGCGAAGCC\nnTGAAACGCTTTGCCGGCTTCGGCCGTGGCCCGGGTCCGGC  
GCGATGCCCAGCGCAAGGCAGCCGCTCGCCG\nnCCATCGTGAACCTCAGGAAGCCACGCCGGCCGAGCTTCGGAGAACCT  
GTTCCGGCGACCGGTGAGGCGTC\nnGGAGTGGGGGGTGTGCGTGTGCGCGGGCGCCCGGCTGTGCGAATACGCATGGTC  
AATGTCTCCCGGCAG\nnGCTCAGGCAAGCGCGCTGGCCGCATCGTGAATCGCGGCGCGGATGCGCTGGTAGGTGCGCGAA



CGGCACA\nGGTTGCCGGCCATCGCGCGATCGATGTGCGCGTCGGTTCGGCGCCTTGTTGCCGTGCAGCAGGCCGGCTGC\  
nGCTCATCATCTGGCCGCTCTGGCAGTAGCCGATTGCGCGACGTCGTGCCGGATCCAGGCGTCGAGGACG\nCGACCGCATCGTCGGCGAGATGTTCCGACGTCGTGATGTGCGCGCTCCCCACCGCCGACAGCG\nnCCAGCGTCAGCTTC  
GGGCCGGCCGGCGTCGACATGAACGGTGCGAGGCGCCGAGTCGCCGGTGCCGCA\nnTCCGAACCTTGGTGCCGGTCAGCC  
CGAGGTTGTGCGCGCAGCACCCAGAGCACCGGTGTGAGGGATCGGCG\nnTCGATCTCGGCCGGCCGGTTGTTGATGTTCA  
ACGTCACCATCGTGAAGTTCTCCATTGGACACCTCGTCA\nnTGTTTCGCGAGGCCAAAGCGTAATGCGCGGACGTGAAT  
CAATTAAGCAGATTTCGAGAAAAACCCAGCA\nnTCTGCAAATTCGATTTTTTGCCTTTCTTAATGATTTTCGGCCGAAG  
ATTAATAATAAAATTTACGTTT\nnTCCGTACAAATCATCGAATCTTTTTGAAAGATTGTTGAGTTTGTACAAGATCTTT  
GCCGACCCGTGGCAA\nnACGTTTGCAGTCGAGCCTATATTTCACTCGCGATCAATGCGGAAGGCGTGAAATCGGATAA  
CAGAAAC\nnCGAAGATTCAATATCCCGGAAATCTGTACCTACCGGGTTGACCGGATGATTGGCCTGTTCTGTTACCGG\  
nGCGGTCATCCCGTCCCGTCTGGAGCAACGGGACGGGCGGTCTGTCGGGACGGAGACATCGAGACACGGG\nnCGGGCAT  
GGGTCGCCGACGGTAATTTCAATCGTCCGATACAGCCGCCTCTAAATGAAATGGGAAAAA\nnTTACCGGGCAGGGGG  
TTCAAAAGTAAAAAATGCCAATTAATAATTCTCGTTGATTAATTCTCGTTTGT\nnGCCAATTCAGCCATCCGCTCGTG  
TTTGGTGACGGGAAGATATTTTCAAAAAATATCGAATAACGCATT\nnCGGTATTTTCTTCGGCATTCCGATGCGCGA  
AAGGGAGTTTCGGCGGCTTCGATTCTCTTTATGGGAAGG\nnTGTGATTATGCGTATGCGCATGGGCTCGAACAAAAAGCG  
CGCGAAAGCACGTCCGGCATTCTCGTCCCG\nnACCGGCATGCTGGTCGCGCTGGTCGCGC\nn>gi|142022655|gb|EQ086233.1|158 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGCCCCGTTTTCACGCGGAGATCGGGCGATTCCATCACCTGAAGGCATCGTGGTCG  
GTCGACGGGATGCTT\nnGGGTCGCTGCAGCGCGAGCCATTCTGAAATTGCGATTTTTCGGTTGGATGAAGCAGTCTTCGA  
AACAGTG\nnCCGGTATGAATGCGAGCAAGAGGGCGATTTCGATGCCTCGTCGCGATCTCGTATTCGTGCAACGACAAGG\  
nAGTCGATCGTGAGTGTGAAATCAAGGGAATCCGAAAAATCGCGTGCTTTATACCGGGCCGTGGTCTG\nnGACGCCA  
ACCTACCTGCTCCCATTACGATCGAGGAATCGATGGCCCGTTTCGGGCTGTCCGAAAGCACC\nnGCTGGCTGGCTCGCA  
TCCGCGGTGCTGCTCTGTCTTTTCGCTTTCCATCATCGTTTTTCGGTCGTACACGG\nnCGACGTTGAACAAGAGAACGGGC  
GCGCTCGCTGCGGCAGCGTTCGCCGTTCTGTGACGTGTGCGATGTT\nnTTCACATAATTTCCATGTATTCTGTCGCCGCG  
AAGCTGATTCTCGGCGCCGCGTTGGGCGTCAGTTGCGTG\nnTGCGCGTATGGGGTGATTGCACACGTCAAGCACCCGGAG  
AAGGTGCGGGCGCAAGTAGCCGTTGCGATGG\nnCCATCATATTTTCTGTGCGATGTACGTCGTACCGATCGTGAATGTG  
AAGATGGGGCATATCGGGTGGA\nnTGCCGTCCAGCTTGCCGTGGTCGTGGCGGCGCTGTTTTTCGGCGCGTTCATGCAT  
CCTGCCGTCAACACA\nnAGCGAGGAAGCGGTTCGGAAGGACGCCATCAAAATCCAGTGACGTCCGAGGGATTCTGATCAGC  
GCATTCT\nnTCATCTACGTCTCGCAGACGGCCCTGATGGGATTTGCCGCTGAGGTGCGTGCAACCCGGGGCGTCGAAGC\  
nGGAGCAACTCGGCATGCTGTTTCATGCTCAATGCGGCCCTCCAGTTGCCGGTGGGGGTGGCGGTCAACTGG\nnCTGGGCG  
ATCGTTTTGGGCTGTTCAAGCCGATCGCCCTTGACTCGCGCTGTTGATCGCATGCAGCCTGG\nnGCATGTATTGCGTGG  
GCGGCAAGTGGGCGTTTCTCGTGTGACGGCGCTCGTCAGTGCCGGCGCCACGCT\nnGGTGGCTCCTTATATGGTGGCGG  
CACTTTCGCAAATGGATGCCAGCGGAAGAAGCACCGCGACCTGTGCG\nnTCGGCAATCAATCTTGGCCTGACGGTGGGCC  
CCGCGATTGCCGGCACGGTGTTCAGCGTGGCTGGTTTGC\nnGTGCGGTTGGCTGGCTCTCCGTCCGGTTGCTGGTGGTGC  
CGATCTTCTGACAGCGTGGCGATCCGGCA\nnGTTCAAGGCAAAACATGCCGAAGCCGCTGTCCGTGACGTGCGCGGC  
GAAGCGATCAACTGAGCGCGGCC\nnGCACTGACCGCACCGGTACCAGGGCGGGCGCCAGCATGCCTGACATGCCATTGCGG  
AGCAGACATCGCTCG\nnGAAGAGCCCGCTCGCGTCGACAGGCAGCGCAAAAACTGCCGTGAACGTGGATGACGAACGCA  
GCGGCAG\nnACAAATATCTCGACGATCATGGATTTCTGCATGGAATCAAAATCTATCTTTATTGCCTGGCTGATCT\nn>  
gi|142022655|gb|EQ086233.1|59 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCGTGAATTCGATCTGCGGCAGCACGATCGCATGCATCAC  
GCCCCGCCAGTCGATCTGCACGAAGCCGAGT\nnTCGAGCACATAGGCGATCAGGCCGAGCGTCATCCATTGAGCCAGCG  
CGCATAGCGCTCGTACGACAGCG\nnTCATCTGCAGCGCGAGCGAGCGAGCCGTAAGCGGCATACAGTAGCTGCG  
CCCGCCCGCAGAG\nnCGCGGTTCGCGTCGCCATCGCCGCGAGATCGGGCGGACGTTGACGACGTTGACGATCGCGAC  
CATCAGC\nnATCAGCGTCTTGCGGAGCCGGTCGGAGAAGTGCTCGCGGATGGCGGCGACGAGGCCCGCGCGGTTGCGGC\  
nGGCCGATGAACGCCGACGCGAGCTGCACGGCGACCATCATCGGACGCGTACGACGATCATCCACAGCAT\nnGTTGAAG  
CCGAAGTGGCTACCGGCCTGCGCGTAGACCGCGATGTTGCCCGGATCGGCATCGGACGCCCCC\nnGCGAGCAGGCCGGGG  
CCGAGCCCGCGGTACCACGGCGTCGCGGTGCGGGCGCGCGCCCGCGCGCGGCT\nnGCGCGTCTCGCCCGCGGTGTTTC

ATCGGGTGTCTTCCCTTGCGCCGTGCCGGCTGGCGCGATCGCGGTCG\nnCGGTTCGGGTTTCGCCGGCCAGGCCCGATGCGG  
TGCGCAATGCGTCGACGAGCGTCCGGTACGCGTGCTCGC\nnGCCAATCGGCGCCGCGCGTGGCAGCACGTACGACGGGT  
GATAGGTGGCGACGACGAGCCGGCTTCCGC\nnCGTGCGAATCGGTGCGCGCGCGGCTCCAGCGTCGCATCGTTGTCTT  
GCAGCAGCGCGCGCAACGCGGTG\nnCCGCCGAGCGCGACGATGACGCGTGGCCGTACCGCGTCGAGCTCGCGCTCGAGCC  
AGTAGCGGCACGCGG\nnCCACCTCGCGTTGCGCGGGCGTCTTGTGACGCCCGCTTGCCGCGCGGCTCCCATTTGAAAT  
GCTTGAC\nnCGCGTTTCGTGATAGCACTGCGCGCGCTCGATCCCGGCTTCATCGAGTGCACGGTCCAGCATGCGCCCG\  
nGCCGCGCCGACGAACGGCAATCCGTTGCGGTCTCTCTGGTCGCCCCGCTGCTCGCCGACCAGCATGATCG\nnTCGCGTC  
GACGGGGCCGGCGCCGGGCACCGGTGGGTGCGATGCTCCACAGCGCGCAGCGGGCGCACGC\nnGTGAGTGTGCGCGG  
TGCATCGTCGGCAGCGGGCGCGCTCCGGCGCGTCATGTGACGACGCCCGTGCGGC\nnACGTGGCGCGGAGGTGCGCCAT  
GCAAGCGGGGCGGAACGGAAGGAAAGCGCAGCGTCACGCACATGAA\nnTGGCTCCGGTACGACGATTGACCTGTGCAG  
CAACCGCGGTTCCGCATCGGCCGAGACGGGCGCGTGGTT\nnTGCGTGAGGCGGAGGTCAACGACAGGAGAAACGCGATG  
CTTCCCGATTCCGGCAACGATCCGGCCCGG\nnACGACGTGCTGCTATGGCGCGCACCCGACAGCGTGCCGGCGCGGTGCG  
CATCGCGTGCTGATCGTCGACGA\nnCTATCGCGACGCGGCCGATGCGGTGCGCCTGCTGCTCGAGGCGCGGGCTTCGAA  
TGCCGCGTTGCCGTC\nnGATCCGTTGCCGCTCTGCGCGATCGCGCGCACTGGCAGCCGTTCCGGTCTGCTCGATATC  
GCGATGC\nnCGGGCCTCGACGGGCTGCAGCTTGCGGCGCGGCTGCGTGCGATCCGCACACCGGCGACATGCTGCTCGT\  
nGGCGTGACGCGGCTTCGCGTCGCGACGCGACTGCGAGGCGCGGAAGCAAGCCGGGTTTCGACGCGCACTGT\nnGCGAAGC  
CGCTGACGCGCACCGGCTGCTCGGCTACCTGGATGCGGCGAGCGGCCGCGCGCAGCGGGACG\nnGGGAGGTTACACGGC  
GCTTGTAACAAACGCCGGCGGGGCGGGTGCGCGCAGGCATCAACCGTCGGTGCGCG\nnTGGCGGCGCGCGCGGGACCTG  
CCCGCGACGACGCGCCAGGCCGCCGGCGCACGCGAGCAGTGCCGCGCT\nnCATCGGAAGTCGCGCGGTAGCCGTGCAGG  
TCGAACAGCATGCCGCCGATCGTCGCGCCGAGCCCGATC\nnGCGAGTTGCACGACCGCCACCATCAGGCCGCCCGCGGT  
TCCGCATCGTGCGGCAGCGTACGGGCGAGCC\nnACGTCCACCAGCCAGCCGGCGCGCGGTGCCGAGAGGCCCCATGCG  
CCGAGCAGGCAGGCGGTTGCCGC\nnCGTGAGGCCGCCGAAGAGCGTCAGCGTACCGAATCGCGGCCATCAGTGCCGGC  
GTGGCCACGAGTGTC\nnCGGTAGAGCCCGTGTTCAGGACGCCGCCGATCAGGACCGTTCCCGCGAAGCCCGGACGCCG  
ATCACGA\nnGCAGGATCAGCGACAGTGTGACGGTGCTGCGTGCGTGACGGTTTCGAGGAAGGGCGCAGGTAGGTGAA\  
nCAGCGTGAACCTGGCCCATGAAGAACACGCTGACCGCCGCCATGCCGAGCGCGACCGGCGGCCGGGCCAGC\nnAGCCTGA  
GCGCGTTGGCCGACGTGCCGCGCGCGCGGACTGCATCGACGGCAGGCTGACAAGCTTCCAGC\nnCGAATGCGATCGCGG  
CGACCGGCACGACGCAAGAAGCGCCAGCGCCACCCGACGATCCCGCCAGAAA\nnGCTGCCGAGCGGAGCCCGGACGA  
CGGTGCCAGCGCATTGCCACCATTACGATCGCCAGCGCGCGCGC\nnACGTGCGTTGCCGGCACGAGGCGCATCGCGG  
TGGCGGCCGACATCGACCAGAAGCCCGGATCGCCACGC\nnCGATCAGGGCGCGGCCGGCGATGAACGCCGCGTAGTTG  
GCGCGAATGCGACGAGCGTGCCGGACACGAT\nnCGTCAGCAGTGTCAACACCAGCAGCAACGGCTTGCGATCGCGCCGGC  
CGGCGAGCGATGCGATGAACAGG\nnCTCGTGACCAGCGCGAACGCGCCGATACGAGATGGCCTGGCCGGCCTGGCCTT  
CGCTGACGTGCAGGT\nnCGCCTGCAATCGGCGTGAGCAGGCTGACCGGCATGAACCTCGGAGGCGACAGCGCGAACGCGC  
CGAGCGA\nnCATCGCGAGGACGGCGCCCCACGCGGCCGCACCGTCAGGTGCGGTGCCGCGATCGGGCTGGAGACGGAG\  
nGGCATGTGCGGAACCGGATCGGAACAAGGTGTGGGTGCGCGCTCAGTCGCAACGCGTCGTGGCGCCGTG\nnCACGGCG  
AGCTGCGTGCGCGCTCGTCGGCGCGCCCGGCTGCACGCGGGCCGCGCGGCGCTGGGCGTCG\nnTCGACCCAGCTGCGC  
GGTTTCATGCGCGCCGATTGAAGCCGGCGGCGCTCGTAGCGCAGCATGTCCTGGC\nnGGACCTGCTACGCGTCAGGCCC  
GGTGCGGATTGCGCGAACAGTGGGCCGGATGCGGCAAGCGCGGCGAC\nnGGCGGCGGCGAGCGGCGCGAGAAAGCGGGTA  
GTCATGAGGTGGGCTCCGGAATGCGGGGTAGTCCGAGCG\nnTCCGGGGCTTTGCCGAGGCGCATCGGGTCGATGGCTTCG  
CAGTCTAGCGAGCCCGCGGATCGATTAAG\nnTAGCGCGGCACGCTTGGAATCATGAGCCAACTGATGAATCGCCGCC  
GGCGGTGCTACGATTGGCGTCCG\nnTCCCGACCGGAGCTTTTCGACATGGCGCGCGAAAATTTGAACGACCTGCTGCTCT  
TTCTCGCGGTGGCCC\nnGCGAGCGCAGCTTACGCGCGCGGCCACGCGGCTCGGCGTGTGCGAGTCGGCGCTGAGCCATA  
CCATCCG\nnCGATCTCGAAGCGGGGTGCGCGTGCGGCTGCTGACGCGCACGACGCGCAGCGTGGCGACGACCGACGCG\  
nGGCGAAGCGCTGTTCCAGGCGGTGGCCCCGCGCATCGACGAGATCACCGCGCAACTGGCCGCGCTGTCGG\nnATTTTCAG  
GGACAAGCCGGCCGGCGTGGTGCGATTACGGCCACCGAGCATCCGATCGATACGGTCATCTG\nnGCCGAACTCAGGCA  
GGTCTGCCCCACTACCCGGACATCCGCGTCGAGCTGTGCGTGCATTACGGGCTG\nnTCGAACATCGTCGAGGAGCGCTA  
CGACATCGGTGTGCGCTACGGCGACCAGGTGGCGAAGGACATGATCG\nnCCGTGCGGATCAGCCCGGATGTCCGGATGGC  
GATGGTCGCGGCGCCGGCCTATCTCGACGGCCGCAAGCC\nnGCCGAAGAAGCCGAGGACCTGCTCGATCACGATTGCGT

CACGTTGCGGCTCGCCACGGCGAAGGGCATC\nTATGCGTGGGAGCTGAAGAAGGGCAGGAACGACGTGCTGGCGCGCGT  
GAACGGCCGGATCACCTGCAACA\nCGCAGCCGCACATGGTCCAGGCGGCGCTCGACGGCTTCGGCATCGCGTTCGTAC  
CGAGGACATCGTGCT\nnGGAGCATGTCCGCAGCGGCCCGCTGCGGATCGTGATGCCGGAAGTGGTGCCCGGTTTTTCCCGG  
CTACCAC\nnCGGTACTATCCGAGCCGCCCGCGGCGTCCGTTACCGTCTGATCGACGCGCTGCGGCATCGCG\  
nCGTAGCGCGTGTGCGCTCGGGGCCCGCTGCGGACGCCCCGGAAGAAAAAACGCCCCGCGCGGGCGGGGC\nnGTTTCGAC  
GCGGCTGCGCCGGCGAACC GGCGCCGCTGGCGGCGCGATCAGTGACGGTCTGCGTGC GCGCGG\nnATCGTTGAGTTCCTC  
GGTCGGGTGCGTCAGCCCGAGTTCGTGCGCGGCGCTCCAG\n>gi|142022655|gb|EQ086233.1|319 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ce\nnCAGCGGCGACGAACATCGCGGCCGGCGCGGCGCGCACCCGGCGTGCAGGTCTTCGGCGTCCATGGC\nnGGGC  
TGACGGGCGTGCTCGCACCCGCGCTCGACCGCAAGCACGTGCGGCAAGCACTGCGCGCGCGCCGCA\nnCGTGGGCGACCA  
CCGGCGAGCACAGCGCGCTGCTCGTGCGCTGCGGCGATCACTGGCAAGGCGACGCATT\nnCGCGCACCATGGGCGGCGC  
GGCTCGACTACCGGCTGCTCGGCCAGTCAGGCTGGGAATACGTGGCCGCG\nnTACGATCACGACGGTCTGCTTTGGGAGC  
GAAACCTGCATGCGGAACCTCGGCTACGCGGAGCGCCTCGTCC\nnGGGTGCGCTGGGGCGGCGCCCGCATTTCGCGATCGCT  
ACCGCTGGGCTGCGTGGCGCGGCCGATCCGGAT\nnCGTCAACGGCACGATCCATCGTTTTCGGGTCCAACGGGTTTCGAGC  
ATGTGAGGAATCGGCATGGCGCAC\nnGGCGCGACCGACGTGCAATTCCACAGCGACACCTACGGCGACGCCGACAGCA  
TCGAGCTCGACATCGGCA\nnATCTCGCCGCCGCGACGATCGTGTCGAAGGCACGATCGACAGCTTCGCGAAGGTGCGCG  
ATCCGCTGCA\nnGCGTAATCCATTTCGTCCATGCGCCTGAATTCCGTTTCGAGATCAGCGGTGCGGAACCTGCTCGCGCGTG  
GT\nnACGGCGACGCATGCGCTTGGCGGAACGGAGCTGTTTCGTGCGGTGAGCGGCTGACCGACCGCGCTGC\nnCCGT  
CGATCTCTCCGGCAGCATCGAGGTGGCGCCCGGCCATGCGGAATTTCGGCTACCGCCCGTCTATTT\nnCTTTGGGCGGCA  
ACGCAACGACAGCAAGGTGTGTCGTGCGCGCAGTTTCATCACGTTTCGGCTAGGTTTTTC\nnAGCACCCCTGCCGTGCTC  
GTTGCCGCCGTGCAACTGGAAGACATTCTTAGCCGGATCGATGCCAACAA\nnCCGCAGCCATCATGACGGGAGCGCCTGT  
TCGTGGACGTGGTGACGCACAATCCTACTTCTGCACGACTG\nnCGACGTTTCGGGTTGGGGCGTCCATCCATTGCTTT  
CAGGCCCGCTAAGCTAATCGGCCCGGGTGCCC\nnGGTGCCGAGGCTGCGATCGCTGCAACGGAGCGGATCACGATGA  
CGATGCGGGAACCGAACAGGTTCAAG\nnGTGCTTCAGGACGTAGCCGACGGCAAGCTCAAGCCATGGCGTGCGGCAGAAC  
GGCTGGCGTCAACGACGC\nnCGCAGGTTTCGCCGCTGGCCGCCCGGCTACGTGAGCACGGTCAGGCAAGTTTCGGTGTCGG  
GACACCGATC\nnGAAGCTCGGCAACCGTCTGTGGACCCAGGAGCCGCGATCGTGCGCTGTCGATCATCCGCGACCGCT  
AC\nnGCCGATTTTCGGCCCGACGCTGCGTGTGCGAGAACCCGTGCGGACGCTTGACGGGACGCAGCCGCCCTGGCAA\nnCGGC  
ATGCCGGAACAGTCACGCGGTATCCGCCGTGGCGAAGTCGCCCCGGGGCAACGGTCGGGTGCGGA\nnCATCGTGCGCCT  
CGGCGCGCGTGAGGCCGAGCGTCTGCAACACCAGCGCCGCCACCCGCTCCGGCATGCC\nnGCTCGCCGACGCAGCCGTGCG  
CGTGACACTCCGCGCGGATCGCCGCGAGAATCGTTCCGGTCATCGTAGC\nnACGCTCATCAGCGGATCGTCGACGACGA  
AGCGGCGTGCGGCGATGCCGCGTTCGTGTCCCGCCGCAACC\nnGCCGACCGAGCCGTGGTTCGAGCGCGCGCACCGAGA  
AGCCCTCGCGCATCAGGAACCGGG\n>gi|142022655|gb|EQ086233.1|438 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnCGGGGTC  
CTGATCCGGGAGTCCGGGCCCGCGGTATCTTCGCGTGTGATCGGCACCGGCGCGGAGATG\nnGCCGCGTTTCGTGCT  
GCGCTGCTTGACGTGATGGCGCTCGAATAAGCGACCGCCATCCGGTCGACGATG\nnCGATGAAGCGCATGCCCGCGCCG  
TATCCGACGATGTGCTTACTTCATCGTTCGCGAACC GCATTTCCATCG\nnCGTTCGATCTTCGCGAGCACCGCTTCGCTGAG  
CGTCTGACGTGGCGCGGCACATCGGGCGTGCACAGCAG\nnTTCCTCGATTTCGCGCGGCCAGTAGCTCGGATGCCGCAC  
GCGCCCCGACCCGCCATCGACACATCGTCA\nnGCGTTTCGTCTCGTGTCCAGGAGAGCCACCATTCTCTGAATGTGAGA  
GAGCTCTCGTTCTGCATAGTCTC\nnGCGCCAGCATAATGATTGAAATTCCCCGAAATACGACAGCAGGCCCTTCGCTCAT  
CAACGCGCCATCGTG\nnTCGCGGTGCAAAGCATTACCGCACACGTTTGTGCGCCATCGTACGCAGAAGCCGTGTGAGTAA  
ACTGCGT\nnGGGCATTTGCTGATTTGATACGGAATGGGGAGCCTGTCCGAATGTTCTGACGACTTCGGCCCCGATGC\  
nCTGTAGTCAATCGTAGGCTTACTGTTGGGCGATGACATCTCCAAAAATGGAGATGCGGGGCGATATCGG\nnCGGGGCG  
GTGTTTCCGGCATTTCCAGGTCGTGCGACGTGCGCGGAAGGCATGCGGTTGCCGGCTGGAGC\nnAAATTCGCGATCAGC  
TCACTACCTCGACTCCACGCGCGTCCGGCGCATGACTGACCGGCAAGCACGAGGC\nnGGGCATTTGCCGTGTGCACATTT  
TGCTTTCCGCATTGCTGGCGCGCGCCTTACCAGCAGGCGTGTGCTG\nnCATCGGTTGTCGAAATGTTTCGTGGAACCTAC  
ATACTTGAGAGACACGGCCCGCTGAAGTCAGATACCGAC\nnATCGCCATGCTTCAGCCCGAAATGAAATCGACCTCCAC  
CGAAGCCCCTTTGGGCAATTGCAGCACACCG\nnACCGACGTACGCGTATGAACGCCTGCGTCGCCAAGCACCGCATACAA

CATGTCGGATGCACCATCGGCGA\nCTTCACTTTGCATCGTGAAGCCGGGTGCGCTTCGCACATAGACGGTGATGCGCGG  
 CACTGTGCTATCGC\nGTGCGAGCGTGCCGAGTGCTTCCTGATCAACGACAACGCACGCAGCGCCGATACGCCGGCGGC  
 CCGGCGA\nnCCTTCGTGAGCGAAACGGATTCCCCACTATGCCACGAAATGCACCACGTCGCCGATCCGCGGAATCT\  
 nGGCCCGCAACGTAGGCCGTATTCCCGTCGACCAGAATGGGCGTGACTTTCCGCCAATCCTGATTTCTC\nGTCCGGA  
 TTGAAGCCGAATTCCGTGCGGACCTGTGTGCTAGCTTATCGTCTCGGTTTCATCAGTGAATCCCC\nnGATACTTCAGAAAAA  
 TGAAAAAATAATTAAGCCTCGCCCGCTTGTCTCGAGGAAGAAATCCGCGCGA\nnACGATCAAAGAGACGTGCCGCTG  
 CCGTTAAAAGTTTCCGGTAATCCTCTACGTTTTCTCATACATGACA\nnCACTTTTCAGGTGTTGGAAGCAGTCCAGCAT  
 CGCTTTATCGATGCCAAATCTCGTAGTCGACGAGAATTT\nnGCTCAAAATTTTCCGCTGAAAACGCTATCATATGTGTAC  
 TCAACCTGAAGAGTGATATATGGCAACGTAC\nnAAGGAAGTGAAGCTCAAATGGAAGCTTTGGCCGAGAAAGCCGAAGC  
 TGCACGTGCTGCGGAATTCAGG\nnCGATCGTCGACGACATCCGCACAAAGTTGCGGAATTCGGCATTACCGAGAAAGA  
 CATTTTCGGTACGCG\nnTCGTGGTCGACCGGCCAAGCAAGCGTCCGCCCGGTGCGGCAAAGTATCGGACCCGAAGAC  
 CGGAGCA\nnACCTGGTCGGGCCGCGGTGCGCGCCCGCATGGATCAAGGACGCGAAGAATCGCAATCGATTCTGTATCC\  
 nAGGAATAAGATTCTCGACGCGACCGGCGCGCGCTCTGGCGCGCCGTCTGTCTATTTGAGCATCGATACC\nnGGCGGTA  
 TTAAAGCCGCTCCGGTGCGGTGAAGTTTGTGCCATCCGGCGCGTGTTGTCACGCTACGCGTC\nnGGATTGCGTGCTGCC  
 CGATCTTTCCCGATTCAACGTTCTGAATAGATCGCAAGCCAGACGTTTTCGTTT\nnCTTCGTGCGTCCACGCGACGCGA  
 TGACGGCAATGCGGCTCGATCAGCACGTAGTCGCGGGATGCATTT\nnATGACGCGCCGGGCTGCTTGAATTCGAGT  
 ACGGCCGCGCCCGATAGCAGCAGACCCATTTCGGCGCGT\nnGCGTCGTACATACCAGAAGCCGTGCGGGCTCGTGTCGCC  
 ATCGATACGATTGCTCGACGTTTCAGACGCT\nnGCCCGGTGACAAGGAGGTGATTTGTTTCATCGTCGCCGCGCTGGCCT  
 TCCGTGGTGAACAGGTTGCCGGT\nnTTGAAGGTGCATGGGCGCGCTCGTCGATGATCGTCGGTGGAAGGAATCGCGCGT  
 TGAGTGACGCGCAAT\nnATTACCGTAGTCTAACGATCACCGTCGACGCTGACTCGTTTCGTGCGAAAAACGCATCGTAGC  
 GTGTCG\nnGTGCCTGCAATCGGGCATCAATCCGTTCAAGCGAGATCCCGTACTGTTTCGGGATCTCGCCACCGAACGT\  
 nGTTGCGATTGATGCCGTCAACATCAGCATGCTTTTCATACCGCTCCCGTTCACACTCGATCCAGCGGAAC\nnGGTAAGC  
 AGGTGATCACCGCTTGCGTCTCGGGGCGCAGCGCGGCCACCACGCAAAGGCTTCGCGCGC\nnTGCTCCACCAGCATG  
 CCGACACCATCGGCTACGCCGGCGACCCCGCATCGCGCGCCAGTCGCAAGAAGC\nnGCGTGAGGCGCTTGCCATAGGCC  
 AGTTCATAAGCCGTACCATGCGGGCTGAAAACGCCGGGCGGCACGGG\nnCGGCAAATCGCCTGTGAGCTCGCGGACGTC  
 GCGTTGACCACAGATCGAAGCGGCCATGCGGTGAGC\nnTCCCGTAAGCGCACGCCGTTATCGGGCGTACGAAGCC  
 AATTGCAATGCCAGCGCTTGCGCTTCTCGA\nnCGTGACGATTCCGCATCACAGTTGCGCGGGGCCGCGCCAGGAAC  
 GGAAGCAACGCGCCGCGCGCCG\nnGCCCGCGGCACCGAGCAGCAACACACGCTTGCCCGGAGCGCGGTGTGAAGATTC  
 ACCTCGATATCGCGC\nnAGCAGACCGATGCCGTGCAAGTTTTCCGCGAGAATCCGATCCCCCTCGAACTTGAACGCATTTC  
 GCCGCGC\nnCTGCGAGCGTCGCGCGCTCGCTGCGCTCATCGCCATGTCGAACGCCGCGAGCTTGAACGGTGC\nn>gi|1  
 42022655|gb|EQ086233.1|210 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
 scaffold, whole genome shotgun sequence\nGAGCGACGGGAAAAGACGTACGGGAACTCACGCCGTGCG  
 CAGGCGCGCCGGGCGAAGGAGCTCTGCTCG\nnAGCATCTCGACGGCAACCTCGGGCTGGCCGACGTGGCGAGCGCCTGC  
 GGTCTGTCTCGCAGCTATTTCTG\nnGAAGGCGTTTCATCAGACGACGGGCTGCCGCCGATCGCTGGCTCATCGTCAG  
 CGCGTGAACGGGCT\nnAAGGAGTGATGCGCAATTCGGGCTGCCGTTGAGCATGATCGCCGACGCGTGCGGTTTCGCC  
 GATCAGA\nnGCCACTTCTCGCGACCTTCACGCGCCTGACCGGCGTCAGCCCTCGCCGCTGGCGCGCCGACAATACCTG\  
 nATCGTCGTGATGCGGGGCCGGTAGGTGAGGTTTCGTTTCGTGTTTCGGTATCGGAGCGTATGCATCCTGA\nnCCGGCGC  
 GATTGGGCGGAGTGCAGGCGGGCGCGGAAGCGGCGTGCGCAGGTTGCTTATGCGTAGAAGCC\nnGATCAGCATCTCGAG  
 ATTGCGTTGCGCGTGCTGCTCGAGCGGCGCAATTCGCCGATCGAAAACAACGGG\nnGCCTGCGCGCGCAGATGCTCGAG  
 CGCCTGCGCTGCGGCCGCGCGGAAGGCTGCCGACGGTGCATCGGGAT\nnCCTGCGCCCACTCGCGCGGATCGCGCGGTC  
 GTATTGCGGAGCCGGTCCGGTGCCCGCGAGGATCGC\nnGAGATCGGCGTCGAGGAAAATTTGCGCGCGCGCTGCAA  
 TTCCGCATCGTCGGCGAACCCGTCAGGCAGC\nnCGGTGCGACTTCGTGCGCAGCACCAGGTCGCGGGCAACGGACACGTG  
 CGATGCATGCGCGTGCAACCAGG\nnCTGCGTCGCAATGTTCTGTCGCGACCTGCGCGAGCCATTGCGCGCTGAGCGCTTC  
 GTTGTCCGCATAATC\nnCGGCAGTGCTGTCGATAGACGACATCGTGCGCCACACGGCGAGCTCGATGGCCGGCCATAG  
 CCGGTCC\nnGCACGATACGGCGGAGGTGCGCGAACAGTTCCGCCAGATGCGCGAGCGTGTGATAGAAGCGCTGCGGCT\  
 nCGCCGTACGCGCGCTCGACGAGCGGCCACACGGTCGTGTCGCGAAGCACGCGTCGAGGCGGCGCGCAG\nnCGTGCGC  
 GCGATATCGGAATCGAGGGCGTTTCATGGCGGCGATCGGCTCAGGCGGGCGATGCGCGAAGCAT\nnAGCAAACGTGGCCGC

GAGCAGGCATGCCGCTTGCGCACGCGTACCGATCGCACCGTATCGGAGGTTGCCA\nTGAAACGAATCCGGCCGCGCTGC  
CCGCTGCTCGCGATCTGCGCGTTTCGGGGCGCTCGCGCTGCACGCAGA\nCTTCGCATCGGCGGCGCCGACGTATCGATC  
AGCCGCGCGGATGACGGAACGACCGCGCCGGGCGAGCCG\nGTGAGCCGGTACACGCTGGCGAACAAGCGCGCGTGACG  
CTGAAAATCATC\n>gi|142022655|gb|EQ086233.1|237 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nGGACCGG  
ATAGCCCTGGTCGCGCGCTTCGCTGAGTGCCGTGTACGCATCCGAGCCGAGCACGACCGAGTA\nCGGGCCGTCGACGCC  
GGCAAGCCGACGCGCTTCGAGCGCGTCGCTGATCGCGTCGGATATGCGCTGACG\nTCGGCCGGCAGCGTGAGCTTGCG  
GTTTCGACGTGCCTTCGCGGATACCGACGATGCCGGCGCGCGGTAGC\nCGTCGAAAATCGCGCCGTCTCGGCGAACGC  
GAGCCGCTGCGCGCGCTCCTTCGCCGGTTGCCAGTCGGC\nGTCTAACTCGAACGGCACCGTCAGCTCGACG\nATCGTGCGGACTTCGCGCAGCCGCGCATTACCAGCTCGCGCGGGGC  
GGCGACGTGACGAGGTGCCCG\nTGCCGACGGCCGACAGCTCGGGCCCTCGGGCCGTCGACATCGACGACGCGGCG  
GCCGGCCACGGATCG\nTTTGAACGTGCGTGCCACTTCCTCCTCGATTGCTCCAGGCGGAGGACGAGATCGGGGCGAG  
TTCGCGG\nTGCAGGTTATTCATGGGTTTGGGTTGCCCTACGTCAGCGGAAACACTTCATGGTGGGCCGATCGGCTGAC  
nGGACTGCGCTGAGCCCGGCGGCCGATCGCGGACGAAAATCAATATTAGCAGGCTCTATTGACACCGCTG\nCAGCACA  
AATCCGACGACATCTGCGGCGGCGCAACTGTTCTGCAATTCGCGTGTGGGGCGACGAGCG\nTATGCGCTGAACGGT  
GCGACCGGATGACGATACCCCGGGGCGCGCCGGCGCTCGCGTCGCGCGGGCCGG\nGCCCGATGGCGCGCCGAACCTGT  
TGGCGGCCGACGCCCCATCGCGTGTGCTGCATTCTTGCGAACATTCA\nTCCGGCAGGGTTCCCGACGCGCGGGGCTCGG  
CGGTATTAGTTGCACGGACCGTGCCCGGATGAATCATA\nGACGCAGCCTATTGAAGCTTCACGACAAGTGAGGGGCGT  
TCATCGCATGGCCGCGCGTCTCGCCCGGCGC\nGCGGTGGGCGCGCCGGACTTTGGCGATGGGCGTTACGACGGCACGCC  
TTTCAGCGAGCCGTGCGGATGCG\nGTGCGAATGACGCCGGTTCGGCGCCGCGGCATCGGCGGCCGCGCACGAAGAACAG  
CGAACCGGTAACCGC\nGCGGCTGTAGTCGAGCAGGCGGTGCTAATTGCCGAATTCGGCAGCTATATCACTGGATATGCG  
CGCTCGC\nCGGCACTGATCAAACAGATACTCGAGAACGTGTTCTGTGGCCACCTGCGTCGGTAGGTCCAACGTAGCAAC  
nTCTTTTTGCGGTTGAAAAGGTTGGCAGCCTGTGCCTAACACCGGCTGCCAACCGGATGATGAAACGAATA\nnAATTGCT  
GTGTGATTTAGTCGAGCGGACCTCTCCGTTTCATCCTAAGATCAAAAAAGACTTGCCAGTTT\nCGGTATCCACCGTCG  
CAGCGTAAGACAGCAGTTCGTTCCCGTCGACATTGAAGCCTTCACTGAACACGTG\nCAGTACCGAAGCGAGCGGCACTT  
GAATTCGTCGATGTTCCGCATCTGCGCCGTAGGTGTGACGTGCCCG\nACAGATTGGGTCTGATACTTCGTATGGTGA  
TATCTTGAGAAAGTAAATCTGCTCAGACTGCGCCTTCTG\nGAAAGTCGTTCCCGGAGGCATTGATTATCTGAGGGGTTG  
AAACGAAGAGATCGTCAAATTCGGACGCTTC\nGACTGTGGCCCATCTTGCGAAGTTCGACCGTGCTTTGTGCGGAGAGAT  
GGCTCCCGATCTTCTCTATCACA\nTCTGCGTTTTCGATAACAAGTTGTTTTGTAGTGTCTATAAAGTTCGACGCGTAGTA  
ATTTTACCGCTCGAT\nCGCGGGCTGCCGTTTCACATAATGCAGGATTTGATGCACATTGCAGTTGACCAGAAGCTCAA  
GATGCTC\nAGCTGCTTTAGCAAAATCTGATACTGCGGTTTCAATTTACGCCCTTCGAGCGATCTCGATTTGAACTTGA  
nCCGTTTGAGAGCGATTCTGCGAGTACCCCGCGAATAGCACTTGACCCGATCTCTGAGATATCCAGAAA\nnTTGCATC  
TGTCGTGCCGACGGGACGATGAAGTATGACTGTGGCGCTTCGTCGTGCGGTACATCTTCCGT\nCTCATGCCCGCTGCC  
CTGGAACGAAGTTCGAGCGGGCCCATCATGAAGCCCAGCCTTCGAGCAACGACG\nTGGTGCCCGGCTTCATGTTGGCC  
GATTCGTTGCATTGAGATTCTTCTTTTCTTTCCATTCTGGAAGGG\nTGTCGTAATTCATGGGCTTATCTTCTGTCCG  
CCTGTGTTTGTACGTAGCGCGATCTCATACTTGAGGTGG\nTGACATCTTTGGGTGCTAACGATTGGAACCAAGTCAC  
TACGCTCCAGCCGACGTAGCGGCGCGGATTTC\nTCATTCTGACGTTTCGGATCCGAGCAGGATGCTCACGTCTTCTAGCA  
GACGGGGAATTGGGACGGGACTG\nCGTAGTCAAGTAAGCAGTCCCGGTCTCTCGAATTCAATGCGACATTTTCGGGGGC  
GGCCGAAAAGCGCAA\nCTTTACCCGGCTACCCTTTGATTGACCAAACTTGCGTGAATTTGCGCCATCGCAGCTTGAC  
GTTTTTC\nnAACCGCATCCGTGTTCCAAGTGCTAAGCTTGCTTACCCCGTTCGTAATAATGAAATGCTGAAGTTGAGAAA  
nCTGGATTTTTTCTTTGAGAAATCCTCGTTCCCAATGGTGCTGTTGATTTTTGCTCCCATGATCGCCATGT\nnTGCCTAA  
CCGCTTTTGATAAGCTTTGACATCGTCAAGTGACCATGTCTTCATCCATTTTTAGACGGATT\nGATGGGGAGAATGTG  
CTCTAAATTAACCTTGACAGCGTCCGGTGCTTGGCACAAGCTCGCCGAGGCGCCA\nGGCACACCACTCTCAAGCATCCG  
GAGGTAGTATCTGGCAAGTGAAGTCTCGATACTGTAGCTACGAAA\nACGCTGCAATGAATGTGGGGTCAGTCGGTAG  
AGTGGTAAATCCTTTCAATATATCTTGATTGAATTGAG\nCTTTCCCTCTGTAACCCCTTTTGCTGCATCGCTGTAGAT  
TCGCTCTAAGGTGCCACCGCCTGCGCCACCA\nACGATTTGGAATCTTACGGCGACCGCAACAAGCTTTTTAAATGCTGC  
TGATACTTTTTTGGGCTCGAATT\nTATCTAGGATTGCAAGCAAAAGCGGGCGAACTTGAACCATCCAAGAAGGTTTCAT

T\n>gi|142022655|gb|EQ086233.1|507 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCAGGCAGCGTTGCGCGGTGACATGAATCAT CGGGCGGCTATCGAGGTACAACCAGTAGCCGTCGACCCT\nnGAACGGCGGTTCGAGGCCCGGTATCGAGACCGACGACCTC ATCGAAAAAGCGGCGTGTACCGTCAATGTCA\nnTCGGTAACGATGGTGGCGTGGTTCGAAATGCATGCTTCGTGTCCATGC TGAAAGCCTGGTACGCATTCTG\nnTCGCACGTCGGTTCGGTCTTGATAATTGGCTTCGCCATTGAATGATCTTCAATGGC GACAAGAAGGTGGGG\nnGGCCGCGACGGGCGCCGATCCTACAGCCCTGCGGCGCATCTCGCAGCACGGCGCGTAGCGC GGCATCG\nnAACGGCGTCGCGTGGTCGATCCCGGCGGCACGCACGCGCGATGCGTCCAGGTGGCAGTCATATGGGCGCG\ nGTGTCGCGTCGGTTCGGTGGTCGATTGCTTGAGCGCGGCTTCGATGCCGAGCGCGGCCGCGATCCGCTG\nnCGCAATG TCGTACTTCGTTCATCGGCTCCTCGCCCGACCAAGTACGAAGGCCGTCGATCGACGCGCCCGCG\nnAGATGGCGCAGCGTG AGGTCGCGGATGACGTCCGCGACATCCGGCGTATACGTCCGGTAACGGATCGCCC\nnATGCGTCCATGCCGACCGCGTCG GCGCCCGCGGAGCGGCAGCGACGATCGCCGGCAGAGGCTCGTGAC\nnGGCCGATTTCGCTCCAGTCGACGATCGGGCCG AAGAGCAGCGGCAGGCGCAGCACACGACAGCGGCGAC\nnGCGGCGAGCAGCGCCGCTCGCTTCAGCTTCGTGCGG CCGTAAATGTTTCAGCGGTTTCGGCGCGGCGT\nnCTCGCGGTACGGGCGAGCCTTGCCGTCGAAAACGTAATCGGTGGAG ATGCCGAGCGTCCATGCGCCGTA\nnGCGCGCGGCGAGCACGCCGATGCGGCGCTGGGCGTGACGTTGATGGCGCGTGCG GCCGCGGATCGCGT\nnTCGCAGACATCCGGGCGGCGCTCGGCCGCGCAGATGATCACGGCGCGCGTCCGCGAGGTGGCG AAGAGGT\nnGTTTCGAGGGCTGGCTCGTCGAGGACGTGAGTTTCGGCGAGGTTTTCCGGGGCAGCGCCAGGCGTTTTGC\ nGCCTGCGCTCTGCGGTTCCGGATGGTTCGCGACGAGTTTCAGCGACGATTTCAGGGCCAGGGATGCGGCC\nnACCGCGC GGCCGAGCAGGCCGCGCGCCGATGAGGAGGACGGTCGGGAGGTTCAGTCGACGAGGTTCGGT\nnTTGGCATGGGCGGTC ATTCGGGTGACTTGGAAGTGACGAGGTTCGCGTTTCGACGCTCGGGCCCGGTGGC\nnGGGAGCATAGCCGAGTATGACA GTCCGTACGGCGTGATGACAGATCGGCAACCCCGCGACGACGTGCG\nnTCATGTGCGCCGCGCGCGCGCTCGGCACGG CGCCCCCGTCCACCGGACGACGCCGCTCTGCCGGCGA\nnCGCCGACACGCCGCGCGCGCCCGCTGCGCGCGAGCAAG CCGCGCGCGCGCGCCCTCACGAAGATCGCCC\nnAAAAGCAAAGCCCGTTCGCTTGACGACCGGGCCGTCACCTCTGC GGAACACCGCGCCCCGCGTTACT\nnTCCGGTCGAGCGAATAGCGCCCCGCGCGTTCAGTCGAGCAACAGCAGCCCCGC GATGATGCTAACGTT\nnCTTGTAAGTTGATCATCGCCATGTATTGCTCCATCCCTTGCAACGCCAGTAGCGATGGCC GATCAAC\nnGCGGTTCGCGAGCGTATAGGCGGCGAACACCAGTTCGAGCGGACGCGTGTAGAAGCCGATCGCGATCAGCA\nnGGCCACCGGCCAGCTCGACCGCCACCGGATCGCGGCCCGCAGCTCGGGCGCGGGTTCCCCGTCGACGC\nnCATGTAC GCGACCGTGCCCGAGAAGCCGGTCAGCTTCTGCCAGCCGAACAGCACGAACAGGATCATCAAC\nnAGCACGCGAGCCGCC AGCAGCAGCTCGTCCTTCTCGACTCCAGTGAACGTAACGCATGACAATCAACC\nnTTGAATTGACGGTTTGATGCGTC CGGTGCCGAGATGCGCCGAATCGGCGCGCCAGGTTCGGCAACGAACCA\nnCTCCTCGCACGATCGGACAGCAAG\n>gi|142022655|gb|EQ086233.1|350 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nnGCAATGGGTGTCAGGAAGCGGCGCGGACGACGCCCCGTA TTTTCGAATCTTCAATCCTGTTTTCAGAGGG\nnGAAAAGTTCGACTCGAGCGGCGAATACACGCGCGGTGGGTGCCCGA ACTGTCTGCACTCCCCGCCAAAA\nnCACTCCACAGCCCATGGACCGCTTCGCTGAAGCGCTGCGGGAAGCATCCATCGC GCTCGGCGCAACTTA\nnTCCCTTCCCGATCGTGGATCACCAGTTGACGCGCGGAGCCCTCAGTTCGGTAGAGATCCT GAATACA\nnCAGAGCCAAATGAAGTGACCTTCGATCGCTAGCCGGGCTCGACCGTGCGAGGGGCGCCCCGACGACATA\nnCCGGGGGCGACGTGATGTGCAAGCGCGCACGGCCTGCTTCCGCACATCACGCCTTCCTGCCGATACAGGC\nnTGAAGCA TCAGGGCATCAGCACGGAGTCGACCACGTGAATCACGCCATTGGACTGCATGACATCGCCGAT\nnCGTGACGTGCGCGAC GTTGCCCTTGTCGTGCGTCACGGCCAGCCGTTTCGCGCCACGGCTGACGATCAGG\nnGAATCGCCTTCGACCGTCTTCAG CGTTGCCTTGCCGCCACCTTGCTCGACCGCCCTGGCAAGTTCGAATG\nnCCGTGAGTCGTCCGGATACGACGTGATACGT CAACACCTTGACGAGCATCGCCTTGTCTCCGGCTTCAA\nnCAACGTCTGCACCGTGCCGGCCGGAACCGGGCGAATGC GTCGTTGGTTCGGTGCAAATACCGTGAAGGGC\nnCCCTTGCTCGACAATGTGTGACCGAGCCACCCGCTTGACCGCAGC GACCAGCGTCGTGTGATCGTGAG\nnAATTCACCGCGTTTTTCGACGATGTTCTCGACGGGTACATCGCAGCGCCACCCAC CTCGACCGTCTGCTC\nnCATCGACATGGATCCTGCCTCGGTGGCAGGCGCGTACACGCACGTTGCGAATGCAAGGGCGAG CGCCGCA\nnCCACCCAGTTTAACTTTGCTGCAAGATTTTCAATTTCTCGTTCGAAAGATTGCCGCCGTTGACTGGAAGC\nnAATCGGAGATACGAGTCGAACCGCGCAGCGGATGCGCGATGAGTACTTCATCGCGACAGATCCGCGATC\nnCAGGGAA TGCCGAGGGCGTCGACAACCGTTCGCATCTTCAATCGCCACTGCATCCGTTCTCGCTCGGCCA\nnCCGTATCTGCAAATG TCATCCAGCCAAGGTGCTTGGCGTCCCACCTCGGGAAATGGCGTCGCGATGGCAT\nnTCAGAGGAGACAACTCATGAAGC

TTTTCATCGGTATCGTCGTTGCCGACGCGGCCGTTTCGCGCTCAGC\nGCTTGCGGAGGCGGAGACGACATTGCGCCCCG  
TGGCAGCGCAACATCCGGAGAATCTGGCCATCTGGACCG\nCGCTCGGCGGTGATTCGGCGGCCCCCGCGTAGTGGCCG  
GCGTGGTCAATGCGGCTGTCCGAGGCCTTTT\nGGCGGACCCGGTCGAAGCGCCCTATTTTTCGGGTGTAACAGGTAATG  
CCACCCACAGCAGTGCAGCCCAT\nGACACGCCAGCACGGCTTGAAGCCTGTCTACGCTGCAGTTCGAAGCGCTCCTCG  
GCGGTCCGTACACCT\nACCCTGGCCCTGTCAAGGTGGCCGAGACATCGACGCCAGCCGGAGATGTGCCAGGACATGA  
CGGTGC\nnCCATAACGACGTCGGCGTTCCAGGCTGCGTGT\n>gi|142022655|gb|EQ086233.1|245  
marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequence\nCCAGCCGAGCCCGAGTGCCAGCACGGCGCAGGCGGCGAGCAGCGGCACCGTGACGATCATCGCGAGCAGC  
\nCCGAGGCGGCACGGCACGAGCCACAGCCGGTCGGTAAGCGGTACGCGCAGCCGAGCCCGATCGCGAACA\nnGCGAGA  
CGAGCAGCGGATCTCGACGATTTACGCAGCAGCCGCGCGTTCGCTTCGAGGTCGAGGTTTCAG\nnCAGGCCGGCGCCGG  
CCGGGCCGAGCGGATGCCACCCGAGATAGATCATCGCAGTGTGACGGGCAGG\nnTGGCGCAGCGCGGACGTTCGCGA  
CGCCCATGCCAGCAGCACCGTGCCGACGATCAGGTACCAGAGGGTTT\nnCGTGATGTGTGCGGTAACAGATCGATCAC  
GTGGCCGCGTGCCGTACGGCTTGCGGCCGAACGCCTCGC\nnGCAGCTTCGCTTCGCGCGCTCGAGCGTTGCGCGGCGC  
GCTTTCGGCAGGTTGCGGCCCGCGGTTTCAC\nnGTAGAAGTTCAGCATCGACATCGCCGACTGGAACGGCGACGCCTTG  
CGGCGGCGGCTGCGTTCGGCCGAA\nnTGCTTGAGCGACGCGCGCATCGCGGCCGGTTCATCGGACTTGAAGATGTGCGGT  
TCGATGTGAGCGCAT\nnCGCTTTTTTGATCACGTTCGGCGACACGCGCTTGCGGTGGTTCGGATCGGTGCCCGGGCG  
CGGCCCGG\nnGCGCCGAGTCAGGCGCTTGCCGTGCGCGTGGGTGGCGGTAGCGGAAGGCATTACGTGGGCTCCATTTGGG  
\nGGACAGGGGTGTAGTGACGCGTGTGGCGGCGCAAGCCCCGTGCCGCGCGGCGCGTTCATCCGTACGAC\nnGAACCC  
GTGCGGGGCGGCACGGGTGTGTGCTGGAAGACCGGCAGGACATGCGATCCCGCCATCACGCAG\nnGAGGACACCATGCA  
ATCGCATTCGCGGATCTGGCTGGAACGACGAGATCGACAACGATGCCGATGGCG\nnACACGATCGACCGGACGATCT  
CGATACGACGGTCCATGTGACGTCGAAACGGGCAGGATCACGTTTCA\nnGGCGGCGTGGGCGGAGGCGGCCGATCCGCC  
GCCCCGAGACCGCACGGCATTCGTCGAGCTCGTGCTCGAC\nnTGCGACACGATGGAGCGCTATGCGGATCTCGACGAAAG  
CACGCGGATGCGCGTGACGCGATCCTGCACG\nnCCACCGTGGAGGATCTGGTCGACCGCCTGCCGACGAACACGCGAC  
GCTGACGATCGAGCTGACCGATGC\nnAATGCTCGATGCCGCGAGCCGTCTGCAGTAGCCGCGCTGATCAGCCTTTTCGATA  
CGCGATTTGCG\n>gi|142022655|gb|EQ086233.1|279 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nAGGCTTG  
ACCTTGTGCGGCAGCCGGCAGCCGCCGGCCGGCAGGTTGCCGTACGGATCCTCGTCGGCCGCC\nnGTGCGGGGGGGCGCC  
GGCATCACGGGGGCGGTGACGCGACACGATCGGGGGGAACGGCAACGCATGCA\nnCCCCTCGCCGTCAGCACGAAC  
GCACCGTTGGCCACCGCCGGCGGATCGGCGGCACGCCGGCGTCGCT\nnCAACCCGGTCGGCTGGGCGTCCGACGGCACG  
AAGAAGACATCCACCGCGCGCTTCTGTCATGCGTATC\nnGGCGAATAGTCGGCGAAACCGGCGTTGCGGACCGCGCCA  
TGGTCGACGTCGATCGCGAAGCCGGGCTTCG\nnTCGTGCGGAGACCGAAGAGCGCGCCGCTGGATCTGCGCTTGCGCG  
CCGTGCGGTTGACGATGCGGCC\nnCGCATACACGCCGGCCGTACGCGATGACGCGCGGTTGTTGCGCTTCGATCGAC  
ACTTCCGTACGTAC\nnGCGACGACCGAGCCGGCCGTTTCTGTCATCGCGACGCCCCACGCGTCCCCGGCCGGCAGCGTG  
CGCGCGC\nnCGTAGCCGACTTGTGACGGCCAGCGCGAGCGCTGCCGATGCGCGCGTGTCTGGGCGCGCCAGCCG  
\nCGTCATCCGGTAGCGACCGGATCCTGCCGCGCCGAGTGCGCGAGCTCGTCGACAGCGTTTCCATCACG\nnAACGCCG  
TATGCGAGTTGCCGCCCCGAGCGCCACGTCTGGACCGGCACGTGCGCCTCGGTCTGATGAACCG\nnATACCTGCATCGGGA  
AGCCGTACGGGCTGTTCTGTCACGCCTTCGGTCAGGCTCGGATCGGTGCCGCGCTT\nnGAGCATCGTCGTGCGCTCGAGCG  
GCGAGCCCTTCAGCACAGACTGGCCGACGACACGTGCTGCCAGTCG\nnCGCACGGCGCCGCTGCCGTCCACGCCGATGT  
CGACCGGATGACGACCATCGGGCGGTAATAGCCGCCGC\nnGCAGATCGTCCTCGCGCGTCCAGATCGTCTTGACGGGGC  
CGAGATGGCCGGCCGCGAGGTACGCGGCGGA\nnCACGTGGGCGGCTTCGACCACGTAGTCCGACGTGCGCGTCGAGCGCC  
GGCCATAGTCGCCGCCCGAGGTC\nnAGCGTGAAGATCTGGACTTCTCCGGGGCGACGCCGAGCGCTTCGCGACCGCCG  
CGCGGTGCGTCGTC\n>gi|142022655|gb|EQ086233.1|378 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequence\nCATGGCG  
CGCTGCTGCACGCGCAACGCCTGTGCGATCCAGAGCGCGGCGAGTCGACGCGATCACCGCGAT\nnGCCGATCGACGCGGC  
GAACAGCGCGGCATCGTAGCGGATGCCGGTTCCATGCGCATCGCGGCCATCCCC\nnGTGTAATGCATGCCGGGATCCC  
GGCGCCCATCAGTGCGCGCCACGAGCAGCCGCCAGCCAGCC\nnGTGCGCGCGTGACGACGGTCAGCGCGAAGTA  
CGACACGAGCACGGCAATCGCCAGCGACGCGCCCGTATC\nnCGGCCATGCGTAGCCGAGCGGAATCGGCAGCGAAAACGC

GAGCATGCCGACGAAATGCATCGACCAGATG\nCCGGTGCCCATGGCCACCGCGCCGCCGCCCAGCCACGCACGCTTGAG  
CTTCGGATTGTTCGAGCAGCGAAA\nTGAAGGCGGCCAGGTTCGAGTGTCTATAGGAGGCCAGCGTCGCGATCGCGAGCGA  
CAGCAGGACGAGCGG\nnGAGATTGTAGGTGCCGTACATGATCGAGCGCCCGATCGAGGTGAGTACGTTGCCGCGGCGCG  
CA\n>gi|142022655|gb|EQ086233.1|101 marine metagenome JCVI\_SCAF\_1096627390048  
genomic scaffold, whole genome shotgun sequence\nCACATCGACACGAAGATCACCGCGCATGCGT  
TGCTGATCACGCTCGTCAGCGCGCTGCCTCGGACATGA\nnAGCGATCGATGCCGACCAGCAGTGCACACCGGCGACGG  
GCAGGTCGGGCATGACGACGAGCGTGCGAC\nnAGCGCAACCAGCCCGCTTCCGGAACGCCGCGCGCCCTTGACG  
TGAGCAGCATGATGGCGAGCATC\nnACGGCGATCTGCGACGCGGAAAGGGGCACGTGCGACGCCTGCGCGATGAACAACG  
CGGCGAGCGTCAGAT\nnAGATCGCGGTACCGTCCAGATTGAACGAATAACCCGCCGCGCAGCAGAGCCCCACGACGCCCT  
TGTCGCA\nnCCCGAGCGATTCCAGCTTGACGATCAGGCGTGGCAGAACGGGCTCCGAAGAGGACGTGCGGAGGACGATG\  
nAGCAACTCTTCGCGCAGGTAGCGCAAGAGCCGCCACAGCGCGAAGCCGTGCAGCCGCGGAGCGGGGCGA\nnGCACCAG  
TGCGACGAACAGCCCGCAGGCCACGTAGAAGGACAGCATCAGCTTCGCGAGCGAGCCGATCGA\nnGCCGATTCCGAAGCG  
GCCACCGTGAAGGCCATCGCGCGAATGCGCCGAGCGGCGCGAGCCGATGATC\nnATCGCGAGCAGCGAAAGACGAC  
CTGGGCGACGCCGTGATCAGTGCAAGAACGGGCGCCCGGCCCGCG\nnGGTGTGCGTTTCAGCGAGAAGCCGAACAACAG  
CGACAGCAGCAGCACC GGCAACCTCGCCTTTCTCGAA\nnCGCGCCGAGCATCGTATCGGGATCAGCTCAGCCCGAA  
CGCGACGAGCCCGTTTCGTTGCGCGTCCCTC\nnACGTACGGCGGAGGATGCGCGAATCGAGATGGTGACGTGATGTG  
CATGCCGGCGCCCGTTGCAACA\nnCGAACGCGGTGATGAGCCCGATGGCGAGCGCGACGGCGGTGAGCAGATAGAAGAG  
CGCGAGCGCCCGCAC\nnGATCGTGGCGCGATTGCCTGCCCGTTCCGACGCGACGTGATGCCCGAGACGATCGTGACAGAA  
CACGATC\nnGGCGCGATGGTCATCCTCACGAGGCCGACGAACGCGTCGCTGAGCGGTTTGAACATCGCGCTGCGTCCG\  
nGCCATACATGACCGACGGTCACGCCGAGCAGCATCGCGAGCAGGACCTGCACGTAGAGCGATTTGAGCAG\nnCCTGGCC  
AACCTCACGGAACGGTCTTCCTTCGTGTCTCTTGTGTCTGTCACACCGATGCATCGCTGTT\nnCGCGATGCCGTGCGG  
CTTCGCGTCGACGATCATCCGGTGAAATGCTTCGGCGGCCGGCGTGAGCGGCCGT\nnCCGCGCCGCTTGACGATGCCAC  
GCGGCGTTTGACGACCGGCTCGACGAGCGGCACGCTCGTGAGAATCG\nnGGTGGTCTGTCCCGGGCATCGCCATCGACGG  
CACCGCGGCGACGCCGAGCCCGCCTCGATCAAGCCGAG\nnCAGGGTCGTACGTGGCGCGCTTCGCATACGCTCGGCCC  
CCGCGGCGCCACGGCGGCCAGCGCCTGGTTCG\nnAGCAACAGGCGGTTCCCGGAAGTCTTGTCCACCGACACGTAATCGTG  
CTCGTACAGTTTCGTTCCAGGTAA\nnCGCGCTTCTTGCGCGGAGCGGATGGTTCGCGCGGCAGGCAGCGACGAACCGCTC  
CTGGAGCAACATCTT\nnGAACTCGATGTGCGATTCTTGGCTGCCATGAAGCTCACGCCGAAATCGGCTTCGCCGCTGAT  
GACGGCG\nnCCCAGCACCTCGTTTCGCGCTCGCGTCCAGCAGCTTGACCCGGATGCGCGGAAAGCGCTGATGATAGCGCG\  
nCGATGATGGCCGGCAGAAAGTAGTAGGCGACCGAGGGCACGCACGCGATGGTCACATGGCCCAGGCGGCT\nnCGACGAC  
ACGTGCGGAATGCCGAGCAGCGCCGATCGAGATCGTCGAGCAGCTGTTTCGGCGCTCTGGGCG\nnAACACGCGGCCGACC  
GTGGTGAGCGCGACGCGACGCGTGGTGGCTCGAACAGGCGCACGCCGAGCGCTT\nnCTCGAGCTTGTGATCCGGCGA  
CTCAACGCGGGCTGGGAAATGCTGACCGATTCCGCGGCCCTTGCGGAA\nnACTGCCCGTTTCCACGACCGCGCGAAACGCC  
TGCAAGTCGTTCAAGTCGAAGTTGATCCCCACGGGCGCG\nnTCTCCCATCTCAGATGGGGCGTATTTGCATGATTTTCG  
CCGGGCGGCCGATCGGCGCGGCACGCATTC\nnCGGCCACCTCGATCGCAACCGCGTGCGTGAGCGTGCGGCCTGGCCG  
TCGATCCGCCGCGCTCGATGTC\nnGATGGGGGGCGTCGAGCCTGCGGGCCTTACGGCGCAAGCAGGCCGCGTGCGGCCA  
GATTGGCATAACGCG\nnCGCGCACGCCGAAGGTCCAAGGCGCGATCTCGGTGCACAGCCGCACGGTGTGACGAGTGCGC  
CGAGCG\nn'

```
[18]: c = b.replace("\n", "")
      c
```

```
[18]: 'TCGGGCGAAGGCGGCAGCAAGTCGTCCACGCGCAGCGCGGCACCGGGGCTCTGCCGTGCGCTGCTTGGCCATGGCCT
CCAGCGCACCGATCGGATCAAAGCCGCTGAAGCCTTCGCGCATCAGGCGGCCATAGTTGGCGCCAGTGACCGTACCAACC
GCCTTGATGCGGCGCTCGGTCATCGCTGCATTGATCGAGTAGCCACCGCCGCCGCAAATGCCAGCAGCCCAATGCGTTC
TTCATCCACATAGGGGAGCGTTACGAGGTAGTCGACAGCCACGCGAAATCCTCGACGCGCAGTGTGCGGTCTTCGGTAA
AACGTGGTTTCGCGCGCGCTGGCACCCCTGGAAGCTGGCGTCGAAGGCGATGACGACGAAACCTTCCTTGGCCAGCGCCTCG
CCATACACGTTCCCCGATGTTTGCTCCTTGACGCTGCCGATCGGATGCGCGCTGATGATGGCGGGATATTTCTTGCTTC
```



GTCGAAGTTCGGCGGGAAGTGGATGTCGGCTGCGATATCCCAATACACATTCTTGATCTTGACGCTTTTCATGACAGCTC  
 CGTTCAGGGGGAGGGGGTAAGTTCGCCAGGCCGAATCGTTGGTAGCCAAGCGGCAACGACTCGAATATAGAGAGCCGATT  
 GGAATTCGGTAAGATCGCAATCTGGACTACAGTGGTATCTTCAAATTGACAATGGCACCTACATGGATCCCTCACTGCTT  
 CCGTCTCTCGCGTGGTTCGCCACGTCGCACATCATCGTAGCTTCACGAAAGCGGCTGCGGAAATGGGCGTTTCTCGAGC  
 AAACCTGTGCGCAGAACGTGAAGGCGCTCGAACGCCGGTTGAACGTCAAGCTGCTGTATCGAACGACTCGCGACATGTGCGC  
 TGACCGAGGAGGGGCGAGCGGCTCTACGAGGTGTGGTATCCCGCGCTGGTTCGGGTCGAGCGGACGGTCGACGCGCTGCAC  
 GAGGAGCGCGACGAGCCGTCGGGGTTGATTG>gi|142022655|gb|EQ086233.1|160 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequenceATTGGGGAG  
 GAGGCGAGTTGAGCGGCGGCAGTTGCGCTGCGTGCCTGCGCGGCTCGACATCTGATCCGCACCATGGAAATCCCCGCT  
 CAATCTTTGGAGCAGGGATGCGGGGCGATCAAGATGGGGATGCGGGATGGGGGCGACGGTGTATTTCCGCCAGAAGATTT  
 CGCCGCGGGAGCTCGCGGTGCGTACGTGCATGTTCAAACGCACGGTGC CGCATGGCAGTGGCAGACTGATCAACGCAGC  
 TGAAGCATCCGAAGCGCGCGGGCACGCGTGTCTCGACGCGTGGCCTCACATGCTGTGCGGTGCGTTCAAGACCGAAAG  
 CCACCGACCGACGCGCGAGCAATGCGCTACGCGGATCGCGTTCGACACGAGCCGCGCGGAGGCAAGGCCGACGTATTTCG  
 ATCTTCCAGAGGAAGCCTATTGGCTCGAGTCGTAGTGCTCGATATGGTAGAGCAACATGAATCCCGGGCTAAGTACAAGA  
 AGTAACCCGGCAACGAGTGAGATTGCGACGAATAAACGCTTCACCATGATCGCGCTCCTGAGTTGGTTGAGGTGAATTGG  
 AAAGTCGATTCTGGGGGATCATTCCCCGCAAGGCGCGCAATCCCCGATTGTTCTCAAGATCGCAACGCGATTTCGTGAG  
 GCCGATCTTCATGGGGTGTCTCGCTGGTAGTGATTCCGTGCTGGCCGCGCATGTGCATGACGGCATCCGGGGAG>gi|1  
 42022655|gb|EQ086233.1|41 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
 scaffold, whole genome shotgun sequenceGACCTTGATCGGCTGATCGCCGAGCGTGCCGACGTATTCA  
 TTCACAACCTGCGGCCGGGGGCGGCTGCGGATTATGGGCTCGACGCAACAGCCTGCGAATTACGAAGCCGGAGTTGATC  
 TGCTGCGAGATCGGTGCATTTCGGCCATGTGCGACCGATGAATACGCTTCCCGGTTACGACCCGCTGATGACGGCATTTC  
 CGGGATCATGAGCATCACGGGCGAAGAAGGTCAGGCACCGGTGCGAGCAGGGGTGTCGATCGTCGATTTTCGGCGCGGGCA  
 TGTGGGCTGCCATGGGCATTCTGGCAGCGCTGTACCGCGCCAGATGAAGCACTGCGGGGCGACCGTCAATGCCTCCCTG  
 CTGGAGACCGCACTGGCGTGGATGTCAGTCGGCATCGCGAACTACAACGCCGACGGCGAACC GGTTGGCCGGCATGGGTG  
 GGGCGTCGCGTTTATCGTGCCGACCGTGCGTTCGCGACCGCGGATGGCTTCCTCATCGTGAGTTGCGCGAACGATCCGC  
 TCTTCGCGCGGTTGTGTGCCGCGCTCGACCATCCGGAATGGGCGACGGACGAGCGGTTTCGCGACGAACGCGGGGCGCCTG  
 AAGAACCGCGCCGAGATCGACCGCTGATCGGCGACCGGTTGTGCGACGGGAACGCGGGAGTTCTGGACGGAAACGGTTGCA  
 GGCCGCCGGAATTCCGGTTGCGCCGATTTCAGACGACCGCGGAGGTGCTGGCGCACGAGCAGACGCAGGCGCTCGGCATCA  
 TCGAGAAGCCGTCGGACGACGAGATCGGTATCGTCGGCTGCCGCTTTCGTTTCGACGGCAAGCGTTCCTCCCGCGCTGCAT  
 CTGGCGAAGGATACCGGCGCTGACAATTCTCTCCTCGACAGTCTTCTGAAGGTCAGTCATTAGCCGCTTTTTCGAAGCA  
 GACGACGATGCGGCATGGCGGATACAGTGCCGTTGAGTCGTCAAGCGCCACTTGGTGATCCCGCCAGCCGCGCAATTTG  
 GGTGCGTTGGCGGGTCATTTATTGGCGTACCGGACAGATGACGTTGGTGTACGATCAGGGTCGACGTCAGTCGACGAGGA  
 AGCACATTTGCTGCCATCGATGCGTTATCGTGGTGAGGCGTTTGAGAGTGGAAGTGCACGCCGTGTGACATGACGCCTG  
 ACAACGTCATGAAGACGTGCACGTGAGGAGACATGAGTACATTCAACCGATTGCACCTGATTTCGTGAGTTGATTTGTTT  
 ACGCTGAAGCTGTTCTGTGCGCGGTTGAGGAGCAGCAGATCGGCTGGCCGCGATCCGGGAGAACATCGCCGCATCCAC  
 TGCAACCAAGCGTATTTCAGGATCTCGAAGATATTGCCGGCGTAAACTTCTCGAACGAGGCCCGAAAGCGTCTCGCCCA  
 GCCCGCGCGGTGCAGTACTCCTGCGCTATATCCGCAAGATTTTCGACGATCTGGACGATATGCGATCCGAGATCGCGTCG  
 TTCACCGACGGCATGCAGGGCGAGCTGACGATCGCATCCGCACGGTCGATCATCGTTCCGTTTCTATCGCGCGAAGTAGG  
 CGAGTTCAGTCGGGAATACCCGCTTGTGCGAACTTGTGATGCAGGAGCTCGAAAACGCTGAAATCGTGACAGGCCGTTGTTT  
 GGGGCGAAGCCGATGTGCGGGTGTTCGCCGCGACACACGAGCTGGATCTCGGCGGTGTGATGTACGCGGTACCGGGAA  
 GACAGGCTGATTGCCGTGCTCCGAAAGGGGCATCGCCTGAGCGAGCAGGCCAGCGTGACCTTCAAGGATCTCCTGCCCGA  
 GAATGTCATCGCCGTACGGCGATGCTGGGGGCGGTACGTGCGGCCACGCGGCGGCTGGGCGAAGAGTACAAGCCGCGCT  
 ACAGCGTCCGAGCGCGGGTGTGCGAGTCAGCCTGGTGCAGGCCGGGTTAGGCGTCACCATTCACCCGAATGCCTCGTC  
 AGTCATGAACTGTTTCAGCCGGGTTGCCGTGCTGGAATTGGCCGAACCATGGGCGTTGCGACGCATTATATCGCCACGGC  
 GCGCGGCCGAGACCAGAATCCGATTGTGCGTGCATGATCAAGCAACTGCTGGACCGGCCTCGCGACTGAGTGGCCGCGT  
 CGGCGCTGGATGGCCACGTTCCCGCACTTTTCGACATCGAAATCACAACGACAGACAAGGAGAGGCCAGATGCAGCCCATT  
 GTGCTTGGCAGGTTGGAGATTTCAGAAGGTTGTGAGATGGAACGAGCCTTCCGATCGAAGTGCTGTTTCCGGGCGTCAC

TGCAGGCGACCTTTCCAGCTGAAGGGCTGGTACTGGGACGATGTGCTTTCGCTCGATCCGGCTCAAGCGGGGTGCGGCC  
TGAGCATGCACAGTTACCTCGTACGTTTCAGCGGAAAGACGATCCTGATCGACAGCTGCAACGGGAACGACAAGGATCGC  
TCGACCGAAGCCTTTATCGGCTGAATACGCCATGGCTCGACAGGCTGCGGAAACTCGGCATTGCGCCGGAAGATGTCGA  
CATGGTCATGTGCACGCACCTGCACGTGGATCATGTGGCTGGAACACGCGCCTCGACAACGGGCGCTGGGTACCGACTT  
TCCCGAATGCGCGCTATGTATGAACAGGCGCGATGTGACCATATCAGTTTCGCGGAGCACCCATGCGATGATTGCGGAG  
CCGTGGCTCGATTCCATCTCGCCGTGATCGATGCCGGGCGAGCCGATCTCGTCGACGGAGACGAGGTCAATCTCGGGAA  
GGTGGGTGACGGCATGTGGCTGCGGCCGGCGTATGGCCACTCGCCCGGAGCTGCATGATCCACGCTTGCTGCGGTGGTT  
CGCCGGCGGTATTTCTGGCGACGTGGTGACCATCCAGTTCAGCTGATCCGCGCGGATCTTCATTTCTGTACGACTGG  
AACGGCGATCTGGCGGAACAGGTGCGCATCGCGCTGCTCGAGGAAATTGCCGGAACCGATACGATGCTGTTTCCGGCCCA  
TTTCAGATATGGATCGGGCGGCATGTGAGCGGGACGGCGACAAGTATCGCTTTCGCTTCACCCAGGTCTGACGAGTC>  
gi|142022655|gb|EQ086233.1|221 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
scaffold, whole genome shotgun sequenceGCCGGCGATTGACGTCTGCAGCCTCACGTCCAAACTCAA  
TTTGAGATTCTTGACAGACAGATGGGCTAATTAAGTAGGTGATCGCATTATGGTGCAACATACGATCACCTGCCGATGC  
CAGTCGAGTCAAGTGTATAGGGTAAAAATAGTCAAAAATATGTAATTAAGGGGGCGTGATATGTTTCAAGCCGCCAACCC  
ACACCGCCACCCAGCATCCACCTAGCCAGCGGCCGCCAATTCCACCTGTGGCCAAGCTTGTGAGGGCCGATCCAAT  
ACCACCATAAACAGCACCAGATGCAATTGCATAAACCAATCCGCTTACGCCCCCGCATTCTGGTTTGGTTGGCCTTCT  
TGCTCCTCTCCCCCGATACTAGCAACAGTTCATAAGCGTTAATTCTTGAATGCAGCCAATCTCCAATGAATTTTCAAT  
TGCCGAAGAGTGATTATCATTTTTTCTTCCGTGTATTTGCATAGTTAAAAAGAACTTCCGCCAAAACAACGGATATAGC  
ACCTGCAAAAACATCTCTCAAAATCGATCCTTCTAGAATTTCAATTTTCCCAAAAATAAACCATATCGCAACGCCAAGCA  
AAAAAATCAGAAGGAGTTTTCGGCAACCGCGTTAATGTCCATATCGATAATTTTTTATTTCTCCAGATCACGCCAGAAG  
GATGAGCCGACACTGTCATACGATATTCATGTGAAAATACAAAAATGTAAGCTCAGATCGTCGATTATAATCGCGCCA  
CGAACTGCAATACACGTTTCTACCGTAGCACTCCGTATTAATGCTCGAACCTCCTTTTGGAGGCAATCCTATCAAGTTA  
AGCCGGAACCCCTTGTAACAAAGGCTCGCTCATTGACGATGCGCCTCAACGCCATCTTCAATATTTTCGCGATATGAAATT  
GGCTTGACCCCATCTTGCCAAAGCTCCTCGAAGGCCTTCATACCGCGGATCTGTGCACTCGCCATGTGACGATCCGCGT  
CGCTTCAGCCGACGCCGACGCTCGATTTCTGTC AACCTGAGCTGTTGGTTTCTAGCGCCCCGCGCGGGCCATACAGACC  
GAACTCGACTGTTTCTTCGCCATGCTGATCGATCAGGCTGAAGTATCGCGCGGTGAGTGCAGCAAGCCGTGTCCACGGC  
GCGACAGCACTTGCCGTTCTCCGTATTCACCGACCTCAATGATCAACTGCTCACCAGCATCGAACAGTACTTGTTCACGC  
TCCGATGGCCCGCGCTGCGGGTCTGCGGGCCAAAGGCTCCGACCTGCGCCTGACCTCGCGGGATGCCACGCGACGCATG  
ATCACCGCCTGATCCGCGTCCGCGCGTGTCTACGACGAATCTGCTCGACGCGCACGCCTATTCGGCCCGGACATTGCTG  
CGCTTTAACGCCGCTCCAAGGTACGAAGTGCCTGGCATCGTTTCTCGATCTGTTTAACCAGCGAGTTGCACGTCCGTTG  
CCCTTTTCGACAAGGCTGACTTGCCGTGATTGTGCGCAATATTGACGGACCGGATGTCTTGGTCTGGGAGGTTTTCCGC  
AGAAAGGCGAAGCTAACGATGCTCCGTATTTTTGCTCGATGGGCGGCCTGTGATCGATTTTCTGGTGGGAAAAAGAG  
AAGTTTTGCGTGCCTCTATCGCTAACCTTCATTAGCGATAGGGTTACGTTTTGGACCAAACTGCCATTTTGCAAAACGC  
GTCACATCCACCTGTACACCGCTGTATCCCGAAGCCTATAACTTTTGGATGCCTGTTCTTCCGACTCGCTCTCCTTT  
CCCGACGCCGCTTCCCTGGCCGTCTTGCGTACCTGGTACGCGAGCGTCCGCTCGCGCGAAGCCGTGAGCGTTATTGTCC  
CGACCAGCTCGGAGAAGGAAAGTCCGCCCCGTGGTGTGACTGGTCTGATCCGACGGAAACTCGTCGAGTTTGCCATCAGTC  
GTCATCGTGACGCTCTGGCGAACTGTTTCAGTGCGAAGCAGGCGAACGGATTGTCACCGCAAAGCGGTGGCCCCGGCC  
GTCGCCGCGTTGCATACGAACGGCATCAGGACGCTGGCCGAGCTGATTGTGCGAATTTCCCGCGCCGTCGCTGGTGGCT  
CGGCATTCCAGTCTCGGGGAGCGTAGCGCGCACCGAGTCGAAGCCTTCTTTGCGGCTTGTTCCGCGCTGACGGAGCGGG  
CTCGTTTCGCTGGTCTGGCGACACCCCGGAATCCGTGACGCCGTGGGAAAAACATCCGCGTCCCCCATGATGTGACGGT  
TCGCGCGGCATCTTCCGTGCGCCGAAGGCGATGTGCGCACGATCCCAATCAGGCGGAACGGTCCGGCGTACGCCAGGGTC  
ATAACTGTTTTCCCGGGAGTAGCGAATGGCATCGAGCCAAGGCACAGTCGATTTTCATCGTCGAGCAGATGGCGGCGGCC  
GGTACGGTATCGGCGCGCAAGATGTTCCGCGAATACGGCATCTATTGCGACGGCAAGATGGTCGCGCTCGTCTGCGACGA  
TCGGTTGTTCTGTC AAGCCGACGGCCGAAGGCCGCGCGTTTCTCGGCAGTGCGAAGAAGGCCCGCCGTATCCGGCTGCGA  
AGCCGACCTCGTCATCGCCGGCGAGCGCTGGGACGATCGCGAATGGCTGTGCGCGCTGATCCGGATCACCGCCGCGCAG  
TTGCCGGTGCCGATGAAGCGGCGCGGATAACGGCTCATTGCGGTGCCATGTTGCGCTGCGGCCAGGCGTATTCTGTGACA  
CATGTGCGCGACCATCGCGCATCCGGCGGGAGCACAGAATGAAGACATCCGGACATTGCGCGGGCACGTGGCTTGCTCG

CCGCGCCCGGCGCGCTCGTC>gi|142022655|gb|EQ086233.1|294 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequenceGATCAGCCC  
CGCATACGCGTACGCGCGTGCAGCGCCGAAGAGCGTCTGCTGCCGAGGATCTTGCGCAACTCGAAGCCAATGCCTGCCA  
TCAGCGGCCCTCGTTGCGCGAGCCGGATGCAGCCACGCCCCGACCCGCTTGATGCACCGGGCAACCGGCTTCGAACGGAC  
GCCCCGCGACGCATCCCGTCGGGGGCGATCGCGAGCGCATCGTACAAACGCCCCGATACCGGTGATCATCTGTGCGCGCGTG  
TAGAAGCGTCGGACCCGGGAGAGCCGGCTTGCTGCGCGGCCTGCCAGCGCTGCGCGTCGCCGAGCAGTTCGAGCGACGC  
CTGCGCGAACGCGGCAGGATTCGCGATCGGCACGACCGCACCGCGGCCCGCAACGCGCGGTCTGCTCTCCGAGCCCTT  
CGATCAGTTGCCGGCACGACCCGACGTCCGTCGTAATCGACGGCACGCCGGCCGCGAAGCCTTCGAGCACGACGAGCGGG  
AGCGCCTCGCTGATCGACGTCAGCGCGAGCACGTGCACTTCGGCAAAATGTCTGAATCCGCTGAAAGCCGAGAAAGCG  
CACCGTCGTGCGAGCCCGAGGCTCTGCGCGAGCGCCCGGCATTCCAGCGCATACGCGGGATCCTCTTCTCCGGCCCGA  
CGATCCAGCCTTCGATTTCCGGCATCGTCCGCGATGCGATGAACATCGCGCGACGAACGTCTTGATGTCCTTGATCGGC  
ACCACACGGCCGATCAGCGCGACGATGCGGCGCGGGCCCGCTCCCGCTTGTCAGCAGCGGTGCGAAGCCGTCCACGTC  
GACGCCGTTCCGAATGTTGCGGGTCCGGGCGGCATCGGCCCGCTACGCGATCTGACGGTGTGATTTGGCCTCATAGAGCG  
CGACGATCTCGTCGGCCGCGTCTGATGCGAGCTTGCCGAGCGCTTCAAAGAAACGCACCCACAGCGCGCGAAATAGCTG  
ATCGTCGATACGTCCCGCTCGAACGCGCCCCGGTTGTGCGGATCCACTCGCTCTGCAGCAGATCGATCTTGCGTTCTT  
GGTATAGATGCCGTGCTCGGAAACGAGCAGCGGTGCGCCCGTTCGGTAATGCAACAGCGCACCGAGAAAGCCCGCATAAC  
CCGTGACACGGTGTGATAGACCCGCGCCGGCGGCAGCCCGCTGCGATCCGCGCAAGCTGCCACAGCGGCTTGTCATG  
ATCCGAACCGTCCAGAAGTAGTCGGTGAAGGAAGGGTCGGTGCACTAGCGGTGATAACGCTCGACGATGAAGTCCACGC  
CGCGCGGTGGTGGAAACTGCTGCTCGTTACGCGTCCGTCATCGCCGATCAGTTGGACCAGGTCGGCCATCAGTGCAG  
CCGGGTTGCGCGCGACGCCGTCCGCGCTCAGCCCTCGCCACGCATCGTGCAGCGCAGCCGAACGCTTGAACGCTTCGGGG  
TCGCGGGGAATCTGCCGTGCGCCGCCCGTATCGGCAGGCACGTGCGCATAACAGATAATGGGATTCTGAAGTGAACGACGTT  
GTCGGGCAACGCGTATGCGGCGCCATGATAGTCGTCTTCCCGGTGCGCATGAACACGATCGCGAAACGCTTGTCGGAT  
ATGCACGGATCATCTCGTTGACCCAACTCGATACGCCGCCTCGCACATACGGGAACGTGCCTTCGAGCAGCAGGCAGACG  
TCGATCGCGGTGACGACCGGCACGGCGGTGAATGGACGGATTTCATGACGACAGTAGCGAACGACGGGATTGAGAACG  
GGCAGCGCGGCCGCGTTGCCGAGCGACGCGAGCAGCTCGGCCGACGCGAATACTCGCCGCGCAGGAACGCGGCCCTCCGC  
AAGCCACGGCACGAGACGCTCGCGCGAAAGCCCCATTTCGAGCGCGGCTCATGTATTGCACTGCTGCATCCGCGTCGC  
CGGTGCGGAGTGCGAGGCGGCCGGAACGAGCCACATCGCCGCTCGCCTTCGTGATCGCGAGCGCGGCTGCGCGTGG  
CGGTGCGCCTGTTGCGCGGTGTGCCCGTGACCGCGCCGAGCACGAGATTCTGATAGACCAGCTCGAAGTACAGCTCGGC  
AAGCCGGCGATGGGTGCGCTGGCGCGTTGCGTCGTCTGCGCGCGCTCGAGCGTCTGCTGCGCGCGGAAGATGTTCTGCG  
CGATCTCGTTCTCGGCGCGGTGAGCGTGCCGATGCAATGAGCCGACGTCTTCGAGAGGATCGGCGAGCAGATCGCGC  
AGCAGCGTGCCGGTCTGCGCGTCCGCATGCTCTGGATCGCCACCAGCGCGGAAAGACGATCCGACGCGGCCACCCGCGT  
ATTCACGAGACGCGCCTGCAGACGCGCACCGCCCGCTGCGAGACGCGGACACGAGATAGGTGACGAATTCGGGCAAAG  
GCACGTGCGCGTATGCGTCGTCTGCTTTCGCGGGCCAGCAAGCCGCCACACCACTCGCCAGCACGACAAGGCCG  
CCGCCGACGGGCACGAACGTGCATGCGAGCCAGAGCGCCGAGCGTCCCGTTGCGCGCGGTTCCGTTACGTTGCGGCAG  
CAACTGCCGTTACAGCACCGCCTGCGCCGCGCCGCTGCCACCATCGCGCCGAACACGGGCCACAGGGGTCAACCGGCG  
GTGCGGGAACGGGCATCGCCGCCCGCCTGGGCGAACCCGCCGAACCGCGCGAGCACGGCATACTGCAGCGCCGCGACG  
ATCGCAATGCCGACGGACGACGCGAGTGCGGCCAGCATCGACCGCGCGGACCCGTAACGCGTGGCGCCCCGACTGGG  
CGCCACGAACCGGATCGCAGCTACACATGAACGCCGCTTCCCTTCAGCAGGCGCTGCAATGCCGCCCCGGTTTCGAGG  
CTGCCCAGATGCAGCGTATGCACGCCGATCCGCGCGCCTTCGAGATCGAGACCGAACTGCGCATCGAGGCTCGCCTCGAT  
GCGTGCGAGATAGCCGTGACGCGCGCGTATCGGTTGCCGGCATCAGGTTACGAGCACCGATTGCGTCGGCGTCTCTCA  
GCGGCCATATGAGATCGAGCGAACGGCGCCCGCGCACCATGCTCGAACAGCGAATCGTCCGCCTCGTCGTGCTGGAAC  
ACGAGCGCAACACGACGACGCGATCCCCGACTCGCGCTGCAGGCGCGTCAGACGCGCGATGTCGAGCGCGAATTCGTA  
CGGGCATGCCGTTACTGCCGCCACACGTGCGGCACGACGCGCGATGCTCGACGCCGTCCGCGTAATAGCCGAGCAGCA  
CCAGCAGCAGTTGCAGATTGTCAAGTTTCAGCGACAGAAACGGCATGCGCTTGATCGCGAGCAACCGGATCAAGCGGCCG  
TCGGCGCACACGACCGGCGCGGCGACAGCAGTGCCGTATTTCGGCGCCACGTGCTGCTGCTTTTTCGGATGGGCGACGCA  
TTTCGTCTCGAGCGCTCGGGCCACGAGGCCGTGCGCGGATCGAAATCGAATGCATCGCCGAGCCGCGCAACCGGTTGCG  
ACACGATTCTGGGCGAGTTCTGTCGCACCTTCGACCGGATACAGCGCCGCCACCTCGATCTGACATGCCTGTGCGACGAAT

TCGAGCAGCGCTGCGCACCGGGTAGCCGCACCTTGCTCGCCGCGCCCTGCGGGCGCGCATCGAATCCGTGCGCGAGCGA  
CAGGCGGCGCAATTCGGTGATCGAATCCCGCAGCGTGGTTCGGCTTCGACAGCAGATCCTTTTCCAGCCGCTCGTGCGACA  
GGCGCATCAGGTAGTGGCTCTTCGTGATCGCGACGAGCCGGTCGTTTCAGATAGTCGTTGAGCG>gi|142022655|gb|  
EQ086233.1|323 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole  
genome shotgun sequenceACGCCCCGGCGCACCGCGAGTACCGCGCCGCGGGGCACTCCTTGACCCCGCATGATCG  
ATTCGGATGAAACCCGAAAACCTCGTCGCCTGCCACGAATGCGACCTGCTGTTTTGGCGGCCGCGCGCTTGCGCGCGC  
TGGCTGCGCACTGCCCAGGTGCCGTGCCCGGTGGGCGGCAGCGCGCACGGCCGTCCGGCGCTCGACCGGCGGTGCGCG  
ATCGCGCTCGCCGCGCTGTTACGCTCTTCATCGCGCAGGCCCTTCCCATCGTCGCGCTCGACGCGCGCGCATCGCATC  
GCACGCGACGCTGGCCGACGCGGTGGCCGCGTTGCGCTTGAACGGGAACCGGCGGTGGCGGCGATCGTGTTCTGCACGA  
CGATGTTGTTCCCGCTGCTGGAACCTCGCCGCGTGGCTGTACGTGCTCGTACCGTTGCGCGCGGGCCGCTACCGCCCCGC  
TTCGAGCCGGTCTGCGCAACATGCAGCGGTGCGCCCGTGGAGCATGGTCGAGGTGTTCTGCTCGGCATCCTGGTTCAC  
GATCGTCAAGATGACGAGCCTCGCGCACGTGATACCGGGCCCCGCGCTGTTTGCGTTCCGGCGCCCTACCGTGTTGCTCG  
GCTTCTCGCGTCATTCGACCCGGGCGGCCTGTGGGAAGCGCGCAGCAAATCATCGCGCTGCGCGCGCGGGTACGTCC  
GCCGCGGTATCGCGCCGGCGGCACACGCCGCGACGCGCTGCACCGGTGACGCCGACACAGCGGACGCAACGAACGCGAC  
CGGCGCGACCGGGCGCCACCGCTCGGCAAGCGTCACGGCCCGCGCCGCGGGGCTGGTCGCATGTCATACCTGCGGACGCG  
TGCAACCGCATACCGAAGCCGCCGACGCGCGCTGCACGCGCTGCGGTTTCGACGCTGCACGAACGCGCCCCGCGCAGCGCC  
GCACGCACGGGCGCGCTCGTCATCGCGCCGCGCTTCTCTACATCCCGGCGAACCTGCTGCCCATCATGCATGCGACCTC  
GCTCGGCCGCGCCGAGGACGACACCATTTCTCGCCGGCGTCGCGTACTTCTGGACGTCCGGCGACTGGCCGCTCGCCGTCG  
TCGTGTTGCTCGCGAGCGTGCTCGTGCCGATGCTCAAGCTCGCGATTCTCGCGTTGCGAGCCATTGCCGCGCATCGCGGC  
ACGCCATGGCGCCCGCTGGAACGCGCGCGGTCCATCGACTGGTGGAGCGCGTGGGCCGCTGGTCGATGCTCGACGTGTT  
CGTCGTGCGCACTGACGATCGCGCTCGTGCAATTCGTTTCGTTTCGCCGAGATCACGCCCGGCCCGCGCGCACTGGCATTCTG  
GCGCAGTGGTCGTGCTACCATGTGTGCGTCGATGCAGTTCGATCCGCGACTCATCTGGGACGGCGCGCACCGCTGCGCC  
GCCAGCCCTCGATCCTGAACGGGCGCAAGCCCGTCACAACGCTACGTCCATTCAAGATGCTTCCCTCGATTTGCCTG  
AACCCGAGATCCGGCCGCGCAGCCGCTGGATCCCGTCCCTCGTCTGGATCGTGCCGCTCGTGTGCGCGCTGATCGGGCTG  
GCGCTCGTGTACCGCGGCATTGCCGCGACCGGCCGACGATCACCGTCACGTTTCGGAACGCAGAAGGCCTCGAAGCCGG  
AAAGACCAAGGTCCGTTACAAGGACGTGATATCGGCTCGGTGCGAGCCATCACGCTCACGCCGACTTCAAGCGCGTG  
TCGTGCGCATCCAGCTCACCAAGGACGCCGCGCAGTTCGCGAACCGCGATACGCGATTCTGGTTCGTTCCGGCCGCGTATC  
GGCGCCACCGGTATCTCCGGGCTCGGCACGCTGCTCTCCGGCGCGTACATCGGTGCGGACCTCGGCCGCTCGCCGACGA  
ACAGACCGACTTCGCCGACTCGAGACACCGCCCGCGCTCACCGCGGACCAGCAGGGCCGCGCGCTACCTGCTGCACGGCG  
ATTGCTCGGCTCGATCGACATCGGCTCGCCGCTGTTCTATCGGCACCTGCAGGTGCGCCAGGTGGTTCGCTTCTCGCTC  
GACAAGGATGGCACGGGCGTCGACGTGCAAGCGTTCTGTAAGGCGCCGTACGACAAATACGTGCGCACCCAAACGCGCTG  
GTGGCACGCAAGCGGCTTGACCTGCGGCTCGATTGCAACGGCCTGAAGCTGAACACGCAATCGCTCGCGACGGTGGTTCG  
TCGGCGGCCTCGCGTTTTCAGTCGCCGCCCCGGCAAGCCGATGCACCGCCGCGCAGCCGACAACGCGTTCGTTCCGGCTCGCG  
CAGGACGAAGCCGATGCGATGCGCGACCCGACGGCACGCCGCTGATGGTCGTATGCGCTTCGACCAGTCGCTGCGCGG  
ACTGTCCGTGCGCGCGCCGCTCGATTTGCGCGGCATCGCGCTCGGTACAGTACGAACATCGGGATCGAGTACAACGAGC  
ACTCGCGCACGTTTCAGCATGAAGGTCACGATGGCGCTGTATCCGTGAGACTGAGCCGGCACAGCGACACCGCGCTGCCG  
GCCCCGATACGGCCGGCGGCCACGACCTCCTCGAGCACCTGGTGCTGCAGGGGCTGCGCGGGCAGTTGCGCACGGGCAG  
CCTGCTCACGGGCCAGTTGTACGTGGCGCTCGACATGTTCCCGAAAGCGCCGCGCGAGCGTGGACGTGCACCGCACGC  
CGGTGAGCTGCCGACCGTTCCGAACACGCTCGACGAACGAGGTTTCAGCTCGCCGACATCGCTCGGAAGCTCGACAAG  
GTTCCATTTCGACCAGATCGGCAACAGCCTGAACGGTGCCTCGAGCATGCGAACCAGTTGTTTCGACACCTCGACGAGCA  
AGTCGTGCCGAGGCGCGGACACGCTGGCGGGCGCAACGCACGTTTCGACGCCGCGCAGGCGACGCTGCGGCAGGATT  
CGCCGATGCAATCGACGTTTCATGACGCGATGCAATCGCTCACGCAGACGCTCCAGTCGCTCAATACGCTGGCCGACTAT  
CTCGAGCGGCATCCGGAAGCGTGCTCTTCGGCAAGAAAGGAGAACCGAAATGACGCCGATTCGTTTCGACGTTCCGG  
ATGCCGATCCGCATCGCGACGTGCATCGCGCTGGCCGTGCTCGGCGCTGTACGTGCGCGCCCGTACGGTTCATACGCT  
CGGGATGGCGGATGGCGCGGGCGGCACACCGACGCTCGCGTCCCGCATGGCTGATCGACATGCAGCGCTGCACGTGG  
CGGCGCCGGCGGACGGCAACCGGCTCGCGGTGCAGCGCGGCCCGAACGGGTGACATCCTGGAACAGGAGCGCTGGGTT  
GCGCCGCTCGGCGACGAGATGCGCGACGGAAGTGTGACGCGCGTCACGTCCCGGCTGAACACGTTTCGACGTTACCGCGT

CGCTCATCCGGATGGCACGCCGGTCTATCGGGTCGCCGTGGACGTCCAGCGTGTGGAATCGTGGCCGGCGTCTCACGTGC  
 TGCTCGATGCGACGTGGACGGTGGACGCCGGCTCAGGACAGCCGGCACTGACTTGCCGCAGCATCGTTCGGGCCGGTGGC  
 TCGGCGGGCTACGACGCGCTCGTCGACGCGCATCGCCATGCGCTCGACACGCTCGCGCTCGGCATCGCCGCCGGCATTCCG  
 CGCGGGGCCACGCACGCGCCGCCGCCCGCCGGATGCCGCGCGGGAACACCCGACCGACGGATCGCGCCAGCCACCGCCG  
 AACCGGGCGCGCCGGATCCGCTCATCGATTGAATCGGGCAGGAAAAACCGCGCGGCCGGGCCGGGTCTGGGTCTGGGTTGGGC  
 GATCGGCCTCGCGCCGGTGCCTGCCGACCCAGCGACATCGTCAGGCGAAACCGTCCCGACGAACCACGAATCGGCGCG  
 GCGGCGTTCCCGCGTCGATCCGGTATTTGCGCGGGCGCGGTGTTGCGCGCTTCGATCATTTCCACCTGCATGGACCACC  
 CTGGATGAAACCCGCCGACTCACGCGCGAGCAACGCAGGCTCGACACCCGGAACGGCTCATCGCTACCGCTCGCGCAC  
 TATTCATCGCGAAAGGCTTTGCCGATACGAGCGTCGAACATATCGCCGAGGGCGCGGGGTACACGCGCGCGCGTTCTAT  
 GCGCACTTCGAACACAAGCGCGAGCTGTTGATCGAGATCCTTCGCCGCGACCATGATCGAATGCTCACGAAAGGTCTGGCT  
 AGTGGCACCGGTGAGCGGCAGCGCTGCACCAGGCCGACTACGCGGAGATTGCCGCCGGCTGGGAATGTTTCCCGCTGT  
 GGGTCGAGGTTTCATCTGTACGCGCTGCGCGATGCCGGGTTTCGGCACATCGTCGACTGCCTGCACACCGAACCCGCCGCG  
 CTACCGGAGACCACAAGCGGCACGGCGGGCGCCACCGATCGGATGCCGTCCGCGGCCGGCTGGGCGGCCGCGCTCGGCGT  
 CGCACTGCTGAGCACCGGCGCGCGGATGCCCGGCGCGACGATCCGGCGTAGCAAGCAGCCACGCGAGCGCCGTTTCGCC  
 CCATCGCCCCGGCTATAATTTCAATCGTTCCCCGACGAGTATCCGCACAATGACGGACGCCAACCTTCTCTACGTGCTGG  
 TAGTCGACGACGATCCGGCAATCCGCGAACTGCTGTCCGGCTATCTGTCCGACAACGACATCCGCGTGACGGTAGCCGAT  
 TCCGGCCAGGAAATGGCGGGCGCGCTGGACACCTTTGCAATCGATCTCGTGATTCTCGACCTGCGCATGCCGGGAGAAGA  
 CGGCGTGCAGATCGCCCGGAGCTGCGC>gi|142022655|gb|EQ086233.1|564 marine metagenome  
 JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequenceCTGGTAGCC  
 ATGCAGCAAGGCTGGCGCTAGATGTACGGCCAGATTGGCGAGCTGGCCAATAACGCGGGAACGCTGTGACATCAATCGA  
 CACTCTGAGCGAAAGCGCTCAACAAGCCGACAGCGCCGCCATCCAGATGGCGAGCGGGATGCAACGGGTCTGGCGGCCAGA  
 TCCGGCATATTCCCTCGTAAGCCGCCTAAGTGGAGGACGCGGTGCGCGGAAGAACGGCCGGGGCCGGTGGTGGCGTGAT  
 CGAGCAGACCATCGCCAGCATCTACGCCATCGTCAACGTGATCAAGGACATCGCCGATCAGACCAGTCTGCTGGCGTTCA  
 ACGCCTGTATCGAGGCGCGCGCGCTGGCGATCAGGGGAGCGGCCTCGCCGTGGTGGCCGACGTAGTGCGCAAATTGGCG  
 GAGCGCACCGGCCAATCCGCCGACGACATCGCTCGCACCGTCATATGCGCCAGTTAATGGCTACGGCGTCCGACGGCATG  
 CAGTAAGACGGCACCGCGCCCGCCAGATCGGCGCCGACGCGATCGCTTCGATCCGCGCCAGTGTCTAACAGGTAAGTGC  
 CCTGGTGCACAGATCACGCAGGCCTTGCGCGAGCAGAGCGCGGCCGGCGCTGCCATCGCCGGCCAGGTGAACCGAATCGC  
 CGGTGGCAGCCGCGAAAGCAGCAAGGCTGCCCGGCGACACTGCCGAGGCGGCGCGCTTTCTGGATCGGCTAGTGATGGAA  
 ATGAAGCAGATATGGCGCGCAAACCCCTCTGCATAAACAACAAGGCCGCTTTCCGGCAAGGAAAGCGGGCCTTATGCC  
 GATCTGGATAGGGTCATGGGCCAGCGCCTGCAGCGACTCGCGCATGATCTGGCGGTGCTCGTCCAATCTCGATCAAATCG  
 GCCTCGACGCCGATCTTGCGCAACTGGATGGCGGCCGACATGCCGAAAGCCACCGCCACCAACAGAACCTCATGCGCT  
 ACGGCCATGGTGGGCTGCTCTTCATTGCATTGTTAAATCCAGGCGAACAACCTCGCCTGTCTCTTGTGATTTCAGTGC  
 CACCGGCTGCAGGCTTTGCCCCAGGCAAGGCCCAAGGACGCATCGCCCGTCTTCCGGCCGGAATTGCGCCCCGTGGGCGA  
 AGCAGACGATGCGACTGCCGTGCGCGTTCAGATAAGCATGCTTGCGCCAGGCCATCGGGGCCCGGCCATATGCGGACAG  
 GCATTCCGATAGCCGTACAGCGTTCCGCCCTTCGCCACCACGAACAAGCTGTACGGCCGGCGCGCTCGGGGTGAAAACC  
 ACGGGCGGCGCCATCGGCGATTTTCATCGCGACGGCCAGTAGAACGCGGCGCGGCACGGCACTCAGTGCTTGCCACCAGG  
 CGGCGGGCCGGACGGCGCCCACTTTTCGCGATGCTGGAACAGGAACAGCTGCGAGGCATCGGCGCCCATCGGCGCTTCGC  
 GCGCGGCCCACTGCTCGTCTGAGATCCATATCGGCGTCTGATTTCGATATGGCAGCCGAGCGGGCTGTTGAAGTACCAG  
 AACCAGTTGGAGCCGAACCTTGTTGGCGGCCCGGCCCCAGAACGACTGGTAGCCCTTGTTACACAGACGCGTGCCGGCCTG  
 CATTACCTCGGTGGGGCCGCCATGTGGAAGGTGAAGTGTTCGACGCCCTGCAT>gi|142022655|gb|EQ086233.  
 1|521 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
 shotgun sequenceCGTTGTTCCGACAGGTCGTCCGCATAGCCGGCCGAGCTGAACTGCGTGACATACGGGCGAATCTT  
 CTTACGACGTGCTAGGACAGCCTGGTAGTCGGCCGGGTTCTTGCTCTCCGGGTCCTTGCCCATGTACTGCAGCGCCGCCG  
 CGAACGCGGCGTCCGGCGCGTCGAGCAGCGAGATGCCGACGCCCTTCAGCTTCGCCGCGTTTCGCCGGTTCGAACAGCAGC  
 GCCAGCTGTGCGTCTGCTGCGTGTGCGCGAGCCGCTTCCTGACCGCTCGACGTTGTAGCCGATCCCGCTCGTGCCCCA  
 CGCCACGGCACACCGTACTGGTTGCCCGGGTCGGCCTTCGCGACCATCTTCATCAGCACCGGGTCGAGGTTTCGCGAGAT  
 TCGGCAGCTTCGACTTGTGAGCTTCTGGTACACGCCGGCCTGGATCTGCTGCGCGAGATAGTTCGACGTCGGCACGACG

ATGTCGTAGCCCGAGCTGCCCGCAAGCAGCTTCGTCTGCAGCGTATCG>gi|142022655|gb|EQ086233.1|455  
marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun  
sequenceTTCGAGTCGCTCGACGCGCCTCGACGGGCAACTAGCCGCCGTTGGCCAGCCTTGAACGGATCTCGTCGGTTCG  
AGAGGTCGCGCAGATGCGTGGCAATCTGCCAAAGATTCCCCACGGATCCTTCACCATCGCCCGCCTGTCGCCGTACGGC  
ATGTCGGCCGGGGCTTCGAGCGGCGTTCGCGTGTGCTGCAAGCGCCCGCGATAGGTCAGGTCGGCATCCTCGACATACAC  
GTACAGAAATGCCGGCATCGGCTCACGCAGGCCGTCTTCCCCGCTGATCATCAACACGGAATCACCGATCCGGATTTCCG  
CAGGCAAGCCCGGGCGGTATTCCGCCCTGCGCATGAAAAACCGTCCGGATGAATCCGATCAGATTGTGCGGATCCGGCACG  
ACGATCCGGGGCGTGACGGTGTGCCATCCATCGGGCACGAACCTGGTCATGTCGGGCAGCCTCCGAAGAGATTTCGGGG  
GTATTGTGCGCCGGGCGCCCTCTGTCGGCAAAGGGGGCGACGCAACGCGGCAAGCGCCGTCAATCTCCGGGCGATCCACG  
CACCGCGGGCGGAGTCGCCCATCCCCCGCATCGAGCGACCGCGCCACGCGAAACCCGAGATCGTCGATTGCAAAGGTTCG  
GATGACTGCGACGGCGATTCTGTCGAAGGCAGCCCCGTTTCGACATCGGCCAGCCCCCG>gi|142022655|gb|EQ08  
6233.1|229 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole  
genome shotgun sequenceGGGTTCCGGCGTCTACTGGGGCGGCTACTGCAACCACGGCAACGTGCTGTTCCCCAC  
GTGGCACCGGGCTTACCTGCTGCGGCTGGAACAGGCGCTGCAGAGCATCCCCGTTGCGAGCAGGTGATGCTGCCGTTCT  
GGGACGAAACCAGCGAGGAGTCGCTGACGAAAGGCATTCCGTGGGCGCTGACCGCCCCGATTTCGAGCTCGACGGCGAA  
AAGATTCCGAACCCGCTGGCCTCCTTCACGTTCAACCGGGCGATCACCGACAACATCAATGGCGACAATCCGAACACAG  
CAAGCCGCTGCCTTACGTCACGGTTCGCTACCCGTTGTGCGGGCTCGTGGGGACCGACAAGGACAAGGCCGCCACCGAGA  
AGCACAACGCCAGGTTCCAGGACGCCAAGGAGAACGTCGTCATACTGAACCGCAACATCGTGGACTGGCTCACGTCGTAC  
ATCGTCGTGAAGGGCGACGTGGTGCCGACCAACGTGAAGAGCAAGTTTCGAGCAGTGCTGACGCGCCCAACTACACGGT  
GTTTTCGAACACGTCGTCCGCGGCCAGTGGAACGAGAATCTTCGGACGGCGCGATCCCCGTGGTGCCGCTCGAGTCGC  
CGCACAACGATATTACCTTGCCGTGCGCGGATTTCGACGTGCCGCACGGACCGAACAAGGGCGATTTCGCGCCGATTCCG  
GGCGCAACGGCGACATGGGCGAAAAACGACACGGCGGGCCTCGACCCGATCTTCTATTTCCACCACTGCTTCATCGATCG  
CGTGTTCTGGCTGTGGCAGAAGCGGCATGGTTACACGCAGCACCTCGACGTCATCGCCGAGTATCCGGGCACCAACTCCG  
TCGATGCGCAGGGGCGGACGCGGGCGTCTGCCCCAATCGTGGCTCACGCTCGAAACGCCGCTCGATCCGTTCAAGAAG  
AGCGAGAACGGCAAGGAGCGCGCTACACGTCGCTCGACTGCATCAACATCGAGGAACAGCTCGGCTACACGTACGGCCC  
CGGTTGCTGGAACCTGCCGAAGTTGTGCTCGCTCGCGGCACCGACCGTGCCGGCCGGCAACAGCCGAAGGTGCTGCGCG  
TGTCGGGGCTGAATCGTGCGCGGATCGCCGGGTCTTCTCGTTTCCGCGTACGTCAACGTGGACGGCGAACGGCAACTG  
CTCGGCACGGAGGCCGTGCTGAGCCGCTGGAGCGTGCAGTCTCGCGGAACTGCCAGACGCACCTCGAAGTGAAGGCGTT  
CTTCCCGCTGAATCACTTCGCCGAAAGCGCGGTGCAGGATGCGCAGTACGACGTCGAAATCCATACGCGCGACGGCGTGC  
GCCTGCAGACGCCACCCGCGTCGCCGCCGACTCGCAAGCCGTCGTCGGCGCGGCGCCGCAACCCGTACGGAACTGTTT  
CGCCTCGAGGTCCGCTAGCGGCGCGCGTGCAGAACCGGCACCGTCGCCCCGGCGCGGTTTACCTCGGTACTTCGACAAC  
GTGACATGACGGGGAGGCGCGGGCATGAGCTACACACCCGAAACCGGCAGCCTGGTGGGGCTGTGGACCTATCGCAGTGT  
CCTGAACGAGCCCGACCCGGCCACCGCATTCAACGACCTTGAATTCGGCCTCGGCACGATCGAAATCGCGCAGGCGCCGG  
CGGGCATTTTCAAGGTCGCATTTTCGGTCCGGGATGGGAGCTGCAGCTCAACGGCTGGATCAGTACGGGAATCCGGGG  
ACGGTGCGATTCCAGGGGCGCGGTGTGGTCAGCGGTGAGGAATGGGTATACGACTATGTGCGGTACGTGAGCGCGCCCTG  
GCCAACCGGCATCGACCAGCGCCCGGCGCTGACCGGTTTCGATCGTCCGTACGGTGCCGCATGCGAGCGGCAGCGGCGGGC  
TCGCGCCGGCCGGGGTGGTGTGCTCGTGGTATGCGGTCATGCGAGATCCCGCCTGAGCGGCCATGACTGCGCCGACTGCA  
CCGCCGGCCGACACGGCCAGGCGTCCGGCGACGGTGACGGACGCGATTGCGCGAGGCGTTGGGATCGGGCCTGCTCGAAT  
GGCTGTTATTTTCGCTATCCGAACGGATATGCGGCGCGATGGCATTTCGCACTCCTCGGCTACCTGGCTTGCGACTTGCCGA  
CCCGGCTGCACGGTGCTCCTTCCAGCTGGATATGGAACGGCATCGGCGCAATGGTTTTCGGCATCGCCGGCTATTCGATG  
CAAGGCGGCCTGTCCGGCGGGTGCGCGCGCAGGGCAATGCTCCTGTGCACGGGGGCCATCGGCGCGGGATTGACGCTCGA  
TGCGATGCGGTCCCCGTCGATGCCATTCTCGACATCTGCGCGGAAACGCTCGATGCCCCGCGCATGTGGAACACACTCG  
TGCTGCATCTCCAATGGTTCCCGATGTGATGCTGGCCATGCTGGCGCTGCTGATCTTGCGCGAGGCCGACCGCAAGGAC  
CGTCGCCGACGCGGCGCGGCAAGCATGCTCGTTGTGGTGCCGTTGCGGCTCGCGCTCGAATTCGCGGGGATGCAGCTCGT  
CATGGCGCTCGGCATGACCGCGATCAGGGCATCGGCGTTTCGCTGCCGGCTGCGTTGGGATACGAGCGGCGTGGCAGTGT  
CGATGCTCGCCAGCATGCTGGCGTTTCGACGCATTGAGCGATCGGGCGGCACGCGCAGGGGGCGGTTGAATGGTCGATTGC  
GAAATAAACACGGCCATATTGAAGAGGCGCGGCCATCCGTTGTGACGAGGCGCGGTTCAAGCTGCTCTGGAAAAGTGT

GATGACGCGGTTTTCGCTCGGCTGCCGGCGCGCTTTGGCAGCGGCAGCCGAGTTTTTAACCCGCGTCACACGGACGCTGCA  
CGGACTCCAGGCGGCGCTCGTCACTATTCACGATGCTCGTGCTCGTGCTCGTGCTCGCGCCACGGATCACGATCGCGCGG  
ATGACGATGCCTCCACTCGTCGCGTTCCCAGTAGCGATGGCCGTCCCAGTAGCGATCGCCGTGCATGCCGATCTCGACCG  
ATACACCGGCCCCCTGCGGCGGGCTCTCGCTCTTCGTATCTCGGTTGCGGTGAACC>gi|142022655|gb|EQ086233  
.1|422 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome  
shotgun sequenceAGCGCGTAATAGAACTGCCGGTAGGTGCGGTTGTAATAGCGGATGCGCGGCCGATCGCGCCGT  
GCAGGTCGTGCAGCCGCTCGATCGGCAGGTTCCGGCATCATGTGGTGCTCCACGTGATAGTTGTTGCCGTTCTGTAACAC  
GTCATGAACGCATTGCTCTCGATCGTCCGCGTGTTCGGAACGGGTGCGGTGCTCGTCAGGTGCGAGCGGTAATGCTCGGG  
CATCTCGACCAGCGCATGCACGGGCGCGACACGAACACGAGCGGCACGACCCACACCCACACGGCGAACACGATATGCA  
GCGCGATCGACACGCCGGCCAGCAGCGCGATCGCCGCGACCATCAGCAGGTGATCGCGGCGGATCGCACGCGACACGAGC  
GGGTTGTGCGCGAAGCTTGCGCCCGGCACGACGAGCTTCGCGACGTTCTTCGCGAACTGGATATAGTGCGCGGCCATCGA  
CAGCCGCCAGGCCACAGGCCGAGATTGACGACAGGGCTCGCGCCGTAAGTGGTCGCCGTAGTCGAAGAACTCCTTGTTCT  
CGGGCGTGCCGAGCAGCCGTTGGTGGCGCAGGTGGCTGTCTGTAACGATGGAACGACACGAGCATCGGCATGCCGAGC  
AGCAGCGCGAACACCATGTTGAGCCGTTGCTCCTGAATCCCTGTAATGCAGCGCCTGGTGTTCAGCTCGACCCCATG  
CGCGTACATCGCGCCGATCAGCAGGATGCCGACGATCTTCAGGACGACATACGTCGACAGCGCGAACAGCATGCCGTGCG  
CGATGAGCGCGACGTAGACCAGAGCTTCAGCGAGAAGACGACGCCGCTCGGGCGGCTCTTCATGTCTGAAACCGGTTTC  
TGACGGGCGAGAAGCGATGAGGTGACGGACAAAATGCTCTCCTTGCCCGGACCCGTGCCCGGCACGATGACGCGGCGACAC  
TGCGCGCTGCCGGGTTGCAATTCTTCTGTGAGTCGATGCTTCGGACAAACGAATTTTCCGAACGCATTGCATGACGCCC  
GCGCATGAGCACACCACATCGTGAAGTAGACGACCGGTGCAATGCGTACCGACCATTCACCTCGCGCGGGCTCGCACGTA  
TACGCACATGGCAGTGCCGAGTGGATGCGTGCTGCTCATGCATAGCCTGCAAATATTCCGTTTGTGCGGTCATGTGATC  
CGGAAGAATCCCCCTCACTCGACGAACGGCGCCGCGGCACCGTGGCGCGAGCGGCCGTGTTCTCAAGGGGAGAATCC  
TCGTGACGAAGTGCGCATCCACCATTCATCGCCTCATCGAATCGCTCGAGGACGTGCGGGCGCCATGCATCGCATGACG  
TTCGTGACGACGCGCGGAACGAGTCGAGCATCACGTACCGGCGCTTCGCGGAAGAGGTGTTTCGCCAGGCCGGCGCACT  
GCGCGAACTCGATGTGCGCGAGAACGACCTGGTGATGCTGGCGTTGCCCGCATCGGTGAGCAGCGGCGCGGATGATGG  
CCTGCGTGATGACGGGCGCACTGCCCTGCACCGTGCCCGTTCCCGCGCGCCGCGCCACCACCGGGCGGCAGGTGATCGAC  
GTCGCGATGCGAGCTGTACCGGCCGCGCCTCGTCATCGCCGCGGATGCGCAGTCCGCGCGTGGCGCGACGACGCGTTTCC  
CGCGTCGTCCACGCGCGTCTGTCGATCTCGACACGCTGTGCTCGTCGCGGATGCCGGTGCGCGCGCGCTGATCAGCGCCA  
AGCATGGCCGCGATCCGCATCAGTCCAGCTCAGTCCGGCTCGACGTGCGATCCGA>gi|142022655|gb|EQ0862  
33.1|384 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole  
genome shotgun sequenceCTTCGATCGCATGTCACGCCGCGCGACATTTTCGACGAACGCGTCGAGCGACACGT  
GAAACATGGGCACGTCCGAGCGCTGCTGGAGCGCCCTCGCCAGGCTGCTCTTCCCTGCACTGGTCGGGCCGTGCAATACG  
ACAGCAACGGGTTTCATGGCGATGGTCCGCGGGAGAGATTGCCGTTGTCCGGGTGCGAGCGCTAGATCAGCTCGGCGACG  
CGCAGATGACGAGCGCGGCCGGAACGAGCACCGCCTGCGCACACACCGTGCCGAGCACGCGTGCGCCGGCGAGTGTGCT  
GATGGCCCCGCGAAAGCTGTTCTCCGAGATCCGCCCTTCGATGACGTGCTCGGTGATGACGGAACCTGCGGGTGCATGA  
CGACCGCCAGCACGAGGTGCGGAAACCGTTGATGACGGACGACAGGTTGCGCGATGTCACTCTCAGCTCGGGCTTCAGG  
TAGCCCCGCTACAGCGACGCGAATACGCCGACCGTCCAGATCGCCATCGCGGCGACGTTCAACGCAATGACGCGCCACGA  
CACGCCGTGTTTCGCGGTGAGTTGCGTGACGTTCTGGCTCGAAGGAGTGCAGCGCGGACGCGCAGATAGTTGATCCCGG  
CGGACTGCACGCATGCAGCAGGCGGGGAATCGAGCGGTTACCTGGAAGTGTTCGACGGCGGACTGAAATAGCGC  
TGAAATGTGCGGATCAGGAACGACCCGACGATGGTGCACCGTGGCGGACAGCAGGAACACCGGAAGTGGGCCAGCAA  
CCCGCTCCCCATCCGTTGCGGATGTCCATTTCAACGCGTTTCGCGATGAACGGCCCCCTGAAAGGAATTGGCCGTTTCGGG  
ACACCAGCGCGATGATCCCGAACAAATGAAAACGACAACGCGATGCGACGCGTGCGCACGCCGGCGATCCGCACGGCATA  
GCCAGTGTGCGGATCATATGGATGACGAAGGTAAGCCCGCAGATGATCCAAAGCTGACTGTCCATGCTGGTGCCGCGTGC  
GTGACGGGATCACGAGTGTACCCGAGCGCGGCGCGCGCATCGGCTGGCGAAGCTGCCTTCGCAAGCGTCACGGTTGGG  
CGACGAGCGTCAGGCGCTACCGTGCGGCGCCAATCCCCGTGCTTCATCCAGCATCCATGTGCGGAAGGCTGCAACGGCC  
CGTTGGCCAACTCGATCGGTTCCGGCAGCAGAGGCGAGAAGTTTTTCGTCACCGTCTTTCCCGCCGGCCACGGTGCGACC  
AGCCGGCCCTGCTCGAGCTCCGCATCGATGTACAGGCGCGGCACCAGCGGACACCGAGGCCGGCCAGCGCGCTTCGAT  
CAGCATCGAATGAAGTCTGAGCGCGCGCCACCGCCGAGTTGGTCAGCACGATGCCGACGCGCTGCGCATAAGCCTGCC

ATGCGTCCGGGTTCTGTGCGCCGGTGAAGACGCGGCAATGCATCCAGCGACGGGTTCTTCCCGGCGCCGGCGAGCAGCGCC  
GGACTGCAGACGGGCACCAGCACTTCTCCAGCAGCGGATGCAGATGCATCCCCACCCAGGCCGGATGCTCGAAATGAAT  
CGCTGCGTCGAAACCGCTCCCGGCCAGCAGGAACGGCTCCATCCGCTCGGCGATATGCACGGTGATGTTCCGGATGCCGGT  
CCTGAAAATGTTTCAGCCGGGGGATCAACCAGCGGGTGGCGAAGGTCGGGATCGCGGCGATATCGAGGCTCGCGCCTTCG  
ATCGGCTGCCCCATCAGGTACAGGCTGTCCCCTTCGAGCCGATCCAGAATCTCGCGAACCTGCACCCGATAGCGCGCACC  
GTTCCGGCGCCAGCCGCACCCGGTTGCCGATTGCTCGAAGAGCGTGACGCCCAGGAACGCCTCCAGCCGGCCGATCTGAC  
GGCTCACGGCACCCCTCGGTGCGTGCCAGCTCCTCGGCGGCTCGGGCAAAGCTGCCGTGTCTGGCCGCGGCTTCGAACACC  
TGGAGTGCGGAATTGCTTGAATCTTGCGGGACATGGAAGTCTCGTGGGTGGGATATGTTGCCATCAGCTTGAGTAAA  
TGTAAGTGAAGGCTGACTTTATTTTCGTTTTATGGCGCGGAATGGCGAGCTTAGCATTACGTTAGCGCACTGAACATCACA  
TGTTCCGGGATGGATGCGCACCCCTGCCCAAAGGACAAACCCATGATCTACACCGTGGAATGCAGCTTCGCCGACCTCGAC  
AGCGAAGCCGAATGGAACGACTTCTACAGCCTTGAAAAGCTGCCCGCCCTGATCTCGGTACAGGGGTTTCACACGTCGCA  
GCGGTTCAAGGCGATCACCCGCGGTTGCCCGGTCTATCTGGCGATCCACACGATCGACGGCCTCGATGTCTGACCGGCG  
ACGAATATCGCCGAAAGGCGGCGGCAATTTTCGCGAAATGGCAACGGCACATCACCGACTGGCACCCGAATCTCTACAGC  
GATATCGGCTTCGCGCCGGCGGTAAAGGACGGTGAACATCTCGCACTGAGCGCCGATGGCCCGGATTGCTGCTTCGATT  
GGGCTCGAACCGCTGGCCATGCAGGCCGTTGCGCTGGAGAAGTTTCCGGCGCGCCGTTGGCTCGCCGTCGTGCCACGGA  
ATAGCGCGCGGCTTGTCGAGGCGCTCCCCGAAGGTATTACCTTTACGCACCGATGACGGAGCAGTTGACGAGCACACAC  
GCCCTTTTCGATTGCGCAGGAGTAGACCGCGATGCCGAACGTCACGATTTCCGGTCGACGCGAAGCGGATGCCGGCTGATGA  
AAGCCTTGCCGAGCTGTGCGCGACTGCGTCGAGCTCTGCACGCAGGTGCTTGAGGCTGAACCTCAAGAACGTCACGTC  
TTTTCTGGCGGTTTCGGCACGGGCATGGACGTCCCGTCTTCGCCGACATCCGGTATCGCGTCGGCACGGCTCGCACGCCG  
GAGGTCATGAACCGGTTTCATGGAAGCGCTGGATCAGGCGATCGTTCCGCCGACCGGCCTCACGGCGCGCATTCCGGTGTT  
CGGTTACACCGCGTCGAACCTTCACGCCCCGAATTAGCCGATTCGGAGCAACACGATGTCCACCCTTACCTTTCCCGTTC  
AGCAAGTCGGCGACTTCACGATTACCGCGATCAGCGATGGATACCTCACTGCGAGCCTCGACTTCCTGTGCAATATCGAT  
TCGGATGACGCATCAAAAATGCAGCGTGATGCGGGGAGAAAGAGCCGGCGGCGGTTTCATATCAATTGCTACGTGCTACG  
CGGAGCGGGCCACACCGTGCTCATCGACGGCGGGGCTGGCGGTTTCAAGCAATGGGGCGGCCAACTCCGGACCAACCTGG  
CGCTTGCCGGTATCGAACCCGCTGCGATCGACACCATCCTGCTTACCCACGCGCATCCCGATCACGTCCGGCGGGCTGGTG  
AATGATGCAGGACAGGTGCGGTTTCACATGCCGAGCTGGTCGTCCATCAGCAAGAGGTCAAGTTCTGGCAGGACGACGG  
AAATCTCGGTGCGGCCAGCGAGCGGGCCCGCGGCAACTTTGCGAAGGCGCGTCAGGTGTTTCGACGCTTATGTGACAGGC  
TTCGATGTTTCGACGATGGACAAGTG>gi|142022655|gb|EQ086233.1|280 marine metagenome

JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequenceGCAGCGACG  
AGCCGCCCCAGATCTTGATCCCGCACGCGCAGCTCGCCGGCCCGACGTGACCGTGACGCTGAGCGGCTCCATCGCCGCA  
TGCGCGAGATACGGAATTCGTAATCCTGCCGAATGCGCGTGCCGCGCGGCGGCGCGCTGCCGGCGTCGGCCGTGCGCAC  
CACGGTCCCCGGCGTGCCGGCGAGCTGCTTGTACGCCTGCATCTGTTCCGGATGACGTGACGGTCGAGCCGGCATCCTTCC  
ACGTGATCTTCAGCGGTGCGGCGGAGCTTGCCGGGCCAGTAGCCGTTCCGGATCACGGCCACGCCGTGCCCGCGCA  
TCGGTCCGGATCTCGACGACCTCGACACGCCCTTGACGGCACGGGCCCGCCGCGTCTGAGCTCGCGACCTTGCCGCC  
GAAGCGCGGGCGCGCGACACGGCCACCATCGTGTCCGGCAGACGCCAGTCCATCCCGTACTTGAGCGTGCCGTGCA  
GCATCCCGCGCGCTCGATGCGCGCGTCCGGCTTGCCGATGATGCGGAACTGGTCCGGCGACTTCAACGACACCTGCTGC  
GGCACCGGCATCGCCATCGCGGCCGTTGCGAGTTGCGCGTAGGTGCGCGGTAAGTTGCCGGAGGTACAGACACCGTTCCG  
CGTCCGGCACGTGGCGGATCGACGTTCCAGCGCTGCGCGGCCCGCGCAGCAGCATCGCGGGCGGATGCGCCGAGTT  
CGCGGTATTGCGCGAACGAGTGATTGACCGACGTGCAACCGGCCGTCATCTGGATGCCCCGTACCCGGATCCTTGTAAGGC  
TCGCCGGCCGGCGCGAGCACGGTGCGCACGTTGCGCCAGTCGGCATCGAGTTGTCGGCCAGCGCCATCGGCAGCGCGGT  
GCTGACGCCCTGGCCGACTCCGTCCGTTGACGGCAACGGTCACGGTGTTGTCCGGCGCGATGATCAGGAATGCTTGCG  
GAGGCGAAGCCTGAAACGCTTTGCCGGCTTCGGCCGTGGCCCGGTCGGCGCGATGCCAGCGCAAGGCAGCCGCTCGCC  
GCCATCGTGAACCTTCAGGAAGCCACGCCGGCCGAGCTTCGGAGAACCTGTTCCGGCGACCGGTGAGGCGTCGGAGTGGGG  
GGTGTCGGTGCTCGCCGGGCGCCCGGCTGTGCGAATACGCATGGTCAATGTCTCCCGGCGAGGCTCAGGAAGCGCGCTGG  
CCGCATCGTGAATCGCGGCGCGGATGCGCTGGTAGGTGCGCGAACGGCACAGGTTGCCGGCCATCGCGCGATCGATGTGCG  
GCGTCGGTCCGGCCTTGTGTCGCTGACGAGGCCGGCTGCGCTCATCATCTGGCCGCTCTGGCAGTAGCCGATTGCGC  
GACGTCGTGCCGGATCCAGGCGTCGAGGACGGCCTTGCCGACCGCATCGTCGGCGAGATGTTCCGACGTCGTGATGTGCG



CGCTCCCCACCGCCGACAGCGCCAGCGTGCAGCTTCGGGCGCGCGCGCGTGCACATGAACGGTGCAGGCGCCGAGTCG  
CCGGTGCCGCATCCGAACCTTGGTGCCGGTCAGCCCGAGGTTGTGCGCAGCACCCAGAGCACCGGTGTGAGGGATCGGC  
GTCGATCTCGGCCGCGCGGTTGTTGATGTTCAACGTCACCATCGTGAAGTTCTCCATTGGACACCTCGTCATGGTTTCGC  
AGGCCAAAGCGTAATGCGGCGGACGTGAATCAATTAAGCAGATTTCGAGAAAAACCCAGCATCTGCAAAATTCGATTTTT  
TGCTTTTCTTAATGATTTTTCGGCCGAAGATTAATAATATAAAATTTACGTTTTCCGTACAAATCATCGAATCTTTTTGAAA  
GATTGTTGAGTTTTGACAAGATCTTGCCGACCCGTGGCAAACGTTTGCCAGTCGAGCCTATATTTCACTCGCGATCAAT  
GCGGAAGGCGTGAAATCGGATAACAGAAACCGAAGATTCAATATCCCGGAAATCTGTACCTCACCGGTTGACCGGATG  
ATTGGCCTGTTCTGTTTACCGGGCGGTCATCCCCGTCCCGTCTGGAGCAACGGGACGGGCGGTCTTCGGGACGGAGACAT  
CGAGACACGGGCGGGCATGGGTGCGCCGACGGTAATTTCAATCGTCCGATACAGCCGCTCTAAATGAAAATGGGAAAA  
ATTACGGGGCAGGGGGTTCAAAAGTAAAAAATGCCAATTAATAATTTCTCGTTGATTAATTCTCGTTTGTGCGCAATTCA  
GCCATCCGCTCGTGTGTTGGTGACGGGAAGATATTTTCAAAAAATATCGAATAACGCATTCCGGTATTTTCCTTCGGCAT  
TCCGATGCGCGAAAGGGAGTTTCGGCGGCTTCGATTCTCTTTATGGGAAGGTGTGATTATGCGTATGCGCATGGGCTCGA  
ACAAAAAGCGCGCGAAAGCACGTCCGGCATTCTCGTCCCGACCGGATGCTGGTTCGCGCTGGTTCGGCG>gi|1420226  
55|gb|EQ086233.1|158 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold,  
whole genome shotgun sequenceGCCGGTTTTTCACGCGGAGATCGGGCGATTCCATCACCTGAAGGCATCGTG  
GTCGGTCGACGGGATGCTTGGGTGCGCTGCAGCGCGAGCCATTCTGAAATTGCGATTTTTCGGTTGGATGAAGCAGTCTTC  
GAAACAGTGCCGGTATGAATGCGAGCAAGAGGGCGATTTCGATGCCTCGTCGCGATCTCGTATTCTGCAAAACGACAAGGA  
GTCGATCGTGAGTGTGAAATCAAGGGAATCCGAAAAATCGCGTGCTTTATACCGGGGCGGTGGTCTGGACGCCAACCT  
ACCTGCTCCCATTCAGCATCGAGGAATCGATGGCCCGGTTCCGGCTGTCCGAAAGCACCGCTGGCTGGCTCGCATCCGCG  
GTGCTGCTCTGTCTTTTCGCTTTCCATCATCGTTTTTCGGTCGTACACGGCGACGTTGAACAAGAGAACGGGCGCGCTCGC  
TGCGGCAGCGTTCCCGGTTCTGTGACGTGTGCGATGTTTTACATAATTTCCATGTATTCTGTCGCCGGAAGCTGATTC  
TCGGCGCCGCGTTGGGCGTCAGTTGCGTGTGCGCGTATGGGGTGATTGCACACGTCAAGCACCCGGAGAAGGTGCGGGCG  
CAAGTAGCCGTTGCGATGGCCATCATATTTTCTGTGCGATGTACGTCGTACCGATCGTGAATGTGAAGATGGGGCATAT  
CGGGGTGGATGCCGTCCAGCTTGCCGTGGTGTGGCGGCGCTGTTTTTCGGCGCGTTTCATGCATCTGCCGTCAACACAA  
GCGAGGAAGCGGTGCGGAAGGACGCCATCAAATCCAGTGACGTCCGAGGGATTCTGATCAGCGCATTCTTCATCTACGTC  
TCGACAGCGGCCCTGATGGGATTGCGCGTGAGGTGCGTGCACCCGGGGCGTGAAGCGGAGCAACTCGGCATGCTGTT  
CATGCTCAATGCGGCCCTCCAGTTGCCGTGGGGGTGGCGGTCAACTGGCTGGGCGATCGTTTTGGGCTGTTCAAGCCGA  
TCGCCCTTGGAATCGCGCTGTTGATCGCATGCAGCCTGGGCATGTATTGCGTGGGCGGCAAGTGGGCGTTTCTCGTGTG  
ACGGCGCTCGTCAGTGCCGGCGCCACGCTGGTGGCTCCTTATATGGTGGCGGCACTTTTCGAAATGGATGCCAGCGGAAG  
AAGCACCGCGACCTGTGCGTCGGCAATCAATCTTGGCCTGACGGTGGGCCCCGCGATTGCCGGCACGGTGTTCAGCGTGG  
CTGGTTTGGTGGGTTGGCTGGCTCTCCGTGGGTTGCTGGTGGTGGCGATCTTCTGACCAGCGTGGCGATCCGGCAG  
TTCAAGGCAAAACATGCCGAAGCCGCTGTGCGCTGACGTGCGCGGCGAAGCGATCAACTGAGCGCGGCCGCACTGACCGC  
ACCGGTACCAGGGCGGCGCCAGCATGCCTGACATGCCATTGCGGAGCAGACATCGCTCGGAAGAGCCCGCTCGCGTCGAC  
AGGACAGCGCAAAAACTGCCGTGAACGTGGATGACGAACGCAGCGGAGACAAATATCTCGACGATCATGGATTTCTGCA  
TGAAACTCAAATTCTATCTTTATTGCTGGCTGATCT>gi|142022655|gb|EQ086233.1|59 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ceCGTGAATTCGATCTGCGGCAGCACGATCGCATGCATCACCCCCGCCAGTCGATCTGCACGAAGCCGAGTTCGAGCAC  
ATAGGCGATCAGGCCGAGCGTCATCCATTTGAGCCAGCGCGCATAGCGCTCGTACGACAGCGTCATCTGCAGCGCGAGCG  
ACGCGAGCCCGTAGAACGGCGCATACCAGTAGCTGCGCCCGCCCGCAGCAGCGCGGTGCGCTCGCCATCGCCGCGAGA  
TCGGCGCCGACGTTGACGACGTTGACGATCGCGACCATCACGATCAGCGTCTTGCGGAGCCGGTTCGAGAGAAGTGCTCGCG  
GATGGCGGCGACGAGGCCCGCGCGTTGCGGCGGCCGATGAACGCCGACGCGAGCTGCACGGCGACCATCATCGGCAGCG  
TGACGACGATCATCCACAGCATGTTGAAGCCGAACCTGGCTACCGGCTGCGCGTAGACCGCGATGTTGCCCGGATCGGCA  
TCGGACGCCCCCGCAGCAGGCCGGGGCCGAGCCGCCGTTACCAGGCGTTCGCGGTGCGCGGCGCGGCGCCCGCGGCCG  
CTGCGCGTCTGTCGCCCGCGGTGTTTCATCGGGTGTCTTCCCTTGCGCGGTGCGGGCTGGCGCGATCGCGGTTCGCGGTTCGGG  
TTCGCCGGCCAGGCCGATGCGGTGCGCAATGCGTCGACGAGCGTCCGGTACGCGTGCTCGCGCGAATCGGCGCGCGCGCG  
TGCGCAGCACGTACGACGGGTGATAGGTGGCGACGACGAGCCGGCTTCCGCCGTGCGAATCGGTGCGCGCGCGCGCGCTCC  
AGCGTCGCATCGTTGTCTTGACGACGCGCGCGCAACGCGGTGCCGCCGAGCGCGACGATGACGCGTGGCCGTACCGCGTC

GAGCTCGCGCTCGAGCCAGTAGCGGCACGCGGCCACCTCGCGTTGCGCGGGCGTCTTGTGCAGCCGCCGCTTGCCGCGCG  
GCTCCCATTTGAAATGCTTGACCGCGTTTCGTACAGTAGCACTGCGCGCGCTCGATCCCGGCTTCATCGAGTGCACGGTCC  
AGCATGCGCCCGGCCGCGCCGACGAACGGCAATCCGTTGCGGTCTCTGGTTCGCGCGGCTGCTCGCCGACAGCATGAT  
CGTCGCGTCGACGGGGCCGCGCCGGGACCGGCTGGGTTCGATGCTCCACAGCGCGCAGCGGCGGCACGCGTCGAGTG  
TCGCCGGTGCATCGTCGGCAGCGGGCGGCGTCCGGCGCGTCATGTACGACGCCCCGTGCGGCACGTGGCGGGCAGGTTCG  
CCCATGCAAGCGGGGCGGAACGGAAGGAAAGCGCAGCGTCACGCACATGAATGGCTCCGGTACGACGATTCGACCTGTG  
CAGCAACCGCCGTTCCGCATCGGCCGAGACGGGCGCGTGGTTTTCGTGGAGGCGGAGGTCAACGACAGGAGAAACGCGAT  
GCTTCCCGATTCCGGCAACGATCCGGCCGCGGACGACGTGCTGCTATGGCGCGCACCCGACAGCGTGCCGGCGCGGTTCGC  
ATCGCGTGCTGATCGTCGACGACTATCGCGACGCGGCCGATGCGCTGCGCCTGCTGCTCGAGGCGCGCGGCTTCGAATGC  
CGCGTTGCCGTCGATCCGTTCCCGTCTGCGCGATCGCGCGGACTGGCAGCCGTTTCGCGGTCTGCTCGATATCGCGAT  
GCCGGGCTCGACGGGCTGCAGCTTGGCGCGCGGTGCGTGGCGATCCGCACACCGGCGACATGCTGCTCGTGGCGTGCA  
GCGGCTTCGCGTCGCGACGCGACTGCGAGGCGGCGAAGCAAGCCGGTTTCGACGCGCACTGTGCGAAGCCGCTGACGCCG  
CACCGGCTGCTCGGCTACCTGGATGCGGCGAGCGGCCGCGCGCAGCGGACGCGGAGGTTACACGGCGCTTGTA AAAAACG  
CCGGCGGGGCGGGTGGCGCAGGCATCAACCGTCGGTTCGCGTGGCGCGCGCGCGGCGGACCTGCCCGCAGCAGCGCC  
AGGCCGCCGGCGCACGCGAGCAGTGCCGCGCTCATCGCAAGGTTCGCGCGGTAGCCGTGCAGGTCAACAGCATGCCGCC  
GATCGTCGCGCCGAGCCCGATCGCGAGTTGCACGACCGCCACCATCAGGCCGCCCGGCTTCGCGATCGTGCGGCAGCG  
TACGGGCGAGCCACGTCCACCAGCCGACCGGCGCGCGGTGCCGAGAGGCCCCATGCGCCGAGCAGGCAGGCGGTTGCC  
GCCGTGGAGCCGCCAAGAGCGTCAGCGCTACCGCAATCGCGGCCATCAGTGCCGGCGTGGCCACGAGTGTCCGGTAGAG  
CCCGTGTTCCAGGACGCCGCCGATCAGGACCGTTCCCGCAAGCCCGCAGCGCGATCAGGACGAGATCAGCGACAGTG  
TGACGGTGCTGCGTGCGTGACGGTTTCGAGGAAGGGGCGCAGGTAGGTGAACAGCGTGAAGTGGCCATGAAGAACACG  
CTGACCGCCGCCATGCCGAGCGGACCGGCGGCCGCGGCGCAGCGCTGAGCGGTTGGCCGACGTGCCGCGCGGCGCGGA  
CTGCATCGACGGCAGGCTGACAAGCTTCCAGCCGAATGCGATCGCGGCGACCGGCACGACGAGAAGAACGCCAGCGCC  
ACCCGACGATCCCGCCAGAAAGCTGCCGAGCGGAGCCGCGACGACGGTCCGACGCGATTGCCACCATTACGATCGCC  
AGCGCGCGCGCACGTGCGTTGCCGGCACGAGGCGCATCGCGTGGCGGCCGACATCGACCAGAAGCCGCCGATCGCCAC  
GCCGATCAGGGCGCGGCCGCGGATGAACGCCGCGTAGTTCGCGCGGAATGCGACGAGCGTGCCGGACACGATCGTCAGCA  
GTGTCAACACCAGCAGCAACGGCTTGCATCGCGCCGCGCGGAGCGATGCGATGAACAGGCTCGTGACCAGCGCGAAC  
GCGCCGATACGAGATGGCTGGCCGGCCTGGCCTTCGCTGACGTGCAGGTGCGCTGCAATCGCGGTGAGCAGGCTGAC  
CGCATGAAGTCCGAGGCGACACGCGGAACGCGCCGAGCGACATCGCGAGGACGGCGCCCCACGCGGCCGACCGTCAG  
GTGCGGTGCCGCGATCGGGCGTGAGACGGAGGGCATGTGCGGAACCGGATCGGAACAAGGTGTGGGTGCGCGCTCAGT  
CGCAACGCGTCGTGGCGCCGTGCACGGCGAGCTGCGTGCGCGCTCGTCGGCGCGCCCGGCTGCACGGGGCCGCGGCG  
GCCTGGGCGTCGTGACCCAGCTGCGCGGTTTCATGCGCGCGGATTGAAGCCGGCGGCTCGTAGCGCAGCATGTCCTG  
GCGGACCTGCTACGCGTCAGGCCCGGTGCGGATTGCGGAACAGTGGGCCGGATGCGGCAAGCGCGGCGACGGCGGCGG  
CAGCGGCGGCGAGAAAGCGGGTAGTCATGAGGTGGGCTCCGGAATGCGGGTAGTCCGAGCGTCCGGGGCTTTGCCGAGG  
CGCATCGGGTCGATGGCTTCGAGTCTAGCGAGCCGCGCGGATCGATTAAGTAGCGCGCACGCTTGACTCATGAGCC  
AAACTGATGAATCGCCGCCGCGGTGCTACGATTGGCGTCCGTCCCGACCGGAGCTTTTCGACATGGCGCGGAAAAATTT  
GAACGACCTGCTCGTCTTTCTCGCGGTGGCCCCGCGAGCGCAGCTTCACGCGCGCGGCCACGCGGCTCGGCGTGTGCGAGT  
CGGCGCTGAGCCATAACATCCGCGATCTCGAAGCGCGGTGCGGCTGCGGCTGCTGACGCGCACGACGCGCAGCGTGCGG  
ACGACCGACGCGGGCGAAGCGCTGTTCCAGGCGGTGGCCCCGCGCATCGACGAGATCACCGCGCAACTGGCCGCGCTGTC  
GGATTTTCAGGGACAAGCCGGCCGCGTGGTGCGATTACGGCCACCGAGCATCCGATCGATACGGTCATCTGGCCGAAAC  
TCAGGCAGGTCTGCCGACTACCCGGACATCCGCGTCGAGCTGTGCGTCGATTACGGGCTGTGGAACATCGTCGAGGAG  
CGCTACGACATCGGTGTGCGCTACGGCGACAGGTGGCGAAGGACATGATCGCCGTGCGGATCAGCCCGGATGTCCGGAT  
GGCGATGGTTCGCGGCGCGGCCCTATCTCGACGGCCGCAAGCCCGGAAGAAGCCGACAGGACCTGCTCGATCAGATTGCG  
TCACGTTGCGGCTCGCCACGGCGAAGGGCATCTATGCGTGGGAGCTGAAGAAGGGCAGGAACGACGTGCTGGCGCGCGTG  
AACGGCCGGATCACCTGCAACACGACGCCGCACATGGTCCAGGCGGCGCTCGACGGCTTCGGCATCGCGTTTCGTACCGGA  
GGACATCGTGCTGGAGCATGTCCGACGCGGCCGCTGCGGATCGTGATGCCGACTGGTGCCCGTTTTTCCCGGCTACC  
ACGCGTACTATCCGAGCCGCCCGCGCGTGCCTGCGTTCACCGTCGTGATCGACGCGCTGCGGCATCGCGGTAGCGC  
GTGTCGCGTCGGGGCCCGCTCGCGACGCCCCGCAAGAAAAACGCCCCGCGGGGCGGGCGTTTCGACGCGGCTGCGCC

GGCGAACCGGCGCCGCGTGGCGGCGCGATCAGTGACGGTCGCGTGCGCGCGGATCGTTGAGTTCCTCGGTCGGGTCTGGTC  
AGCCCGAGTTCGTGCGCGGCGCTCCAG>gi|142022655|gb|EQ086233.1|319 marine metagenome  
JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequenceCAGCGGCGA  
CGAACATCGCGGCGCGCGGCGCGGCGCACCCGGCGTGCAGGTCTTCGGCGTCCATGGCGGGCTGACGGGCGTGCTCG  
CACCCGCGCTCGACCGCAAGCACGTGCGGCGAAGCACTGCGCGCGCGCCGCACGTGGGCGACCACCGGCGAGCACAGCGCG  
CTGCTCGTGCGCTGCGGCGATCACTGGCAAGGCGACGCATTTCGCGCACCATGGGCGGCGCGGCTCGACTACCGGCTGCT  
CGGCCAGTCAGGCTGGGAATACGTGGCCGCGTACGATCAGACGGTCTGCTTTGGGAGCGCAACCTGCATGCGGAACCTCG  
GCTACGCGGAGCGCCTCGTCCGGGTGCGCTGGGGCGGCGCCCGCATTTCGCGATCGCTACCGCTGGGCTGCGTGCGCGGGC  
CGGATCCGGATCGTCAACGGCACGATCCATCGTTTTCGGGTCCAACGGGTTTCGAGCATGTGAGGAATCGGCATGGCGCAC  
CGGCGGACCGACGTCAATTCCACAGCGACACCTACGGCGACGCCGACAGCATCGAGCTCGACATCGGCAATCTCGCCG  
CCGCGACGATCGTGGTCAAGGCACGATCGACAGCTTCGCGAAGGTCGGCGATCCGCTGCAGCGTAATCCATTCTGCCAT  
GCGCCTGAATTCCGTTTTGAGATCAGCGGTGCGGAACCTGCTCGCGCGTGGTACGGCGACGCATGCGCTTGGCGGAACGGA  
GCTGTTCTGTCGCGGTGAGCGGTGACCGACCGCGCTGCCCGTCGATCTCTCCGGCAGCATCGAGGTGGCGCCCGGCC  
ATGCGGAATTCGGCTACCGCCCGGTCTATTTCTTTGGGCGGCAACGCAACGACAGCAAGGTGTGGTCTCGGCGCAGTTC  
ATCAGTTCGGCTAGGTTTTTCAGCACCAACCTGCCGTGCTCGTTGCCGCCGTGCAACTGGAAGACATTCTTAGCCGGATC  
GATGCCAACAAACCGCAGCCATCATGACGGGAGCGCCTGTTCTGTGGACGTGGTGCACGCACAATCTACTTCTGCACGACT  
GCGACGTTTTCGGGTTGGGGCGTCCATCCCATTGCTTTTCAGGCCGCGTAAGCTAATCGGCCGCCGGGTGCCCGGTGCCGGA  
GGCTGCGATCGCTGCAACGGAGCGGATCAGATGACGATGCGGGAACCGAACAGGTTCAAGGTCTTCAGGACGTAGCCG  
ACGGCAAGCTCAAGCCATGGCGTGCGGCAGAACGGCTGGCGTCAACGACGCCGAGGTTCCGCCGTGGCCGCCCGGCTA  
CGTGAGCACGGTCAGGCAAGTTCGGTGTGCGGGACACCGATCGAAGCTCGGCAACCGTCTGTCTGGACCCAGGAGCCCGCA  
TCGTGCGCTGTGATCATCCGCGACCGCTACGCCGATTTTCGGCCCGACGCTGCCTGCGAGAACCCGTCCGCGACGCTTGAC  
GGGACGCAGCCGCCCTGGCAACGGCATGCCGGAACAGTCACGCGGTATCCGCCGTGGCGAAGTCGCCCGGGGGAACGG  
TCGGGTGCGGACATCGTGCGCCTCGGCGCGCGTGAGGCCGAGCGTCTGCAACACCAGCGCCGCCACCCGCTCCGGCATGC  
CGCTCGCCGACGCAGCCGTGCGGTGACACTCCGCGCGGATCGCCGCGAGAATCGTTCCGGTTCATCGCTAGCACGCTCATC  
AGCGGATCGTCGACGACGAAGCGGCGTGCGGCGATGCCGCGTTCGCTGTCCCGCCGCAACCGCCGACCGAGCCCGTGCTC  
GAGCGCGCGCACCGAGAAGCCCTCGCGCATCAGGAACCGGG>gi|142022655|gb|EQ086233.1|438 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ceCGGGGTCTGATCCGGGAGTCCGGGCGCGCGGCGTATCTTCGCGTGTGATCGGCACCGGCGCGGAGATGGCCGCGTT  
TCGTGCTGCGCTGCTTGACGTGATGGCGCTCGAATAAGCGACCGGCCATCCGGTCGACGATGCGATGAAGCGCATGCCCG  
CGCCGTATCCGACGATGTGCTTACTTCATCGTCGGAACCGCATTTCATCGCGTCGATCTTCGCGAGCACCGCTTCGCT  
GAGCGTCTGACGTGGCGCGGCACATCGGGCGTCGACAGCAGTTCCTCGATTGCGCCGCGCCAGTAGCTCGGATGCCGCA  
CGCGCCCCGCACCCGCCATCGACACATCGTCAGCGTTCGTCTCGCTGTCCAGGAGAGCCACCATTCTCTGAATGTGAGAG  
AGCTCTCGTTCTGCATAGTCTCGCGCCAGCATAATGATTGAAATCCCCCGAAATACGACAGCAGGCCTTCGGTCATCAA  
CGCGCCATCGTGTGCGGTGCAAAGCATTACCGCACACGTTTGTGCGCATCGTACGCAGAAGCCGTGTCAGTAAACTGC  
GTGGGCATTTGCTGATTTGATACGGAATGGGGAGCCTGTCCGAATGTTCTGACGACTTCCGGCCCCGATGCCCTGTAGT  
CAATCGTAGGCTTACTGTTGGGCGATGACATCTCCAAAATGGAGATGCGGGGCGATATCGGCGGGGCGGTGTTTTCCGGC  
ATTTCCAGGTCTGCGACGTGCGCGGAAGGCATGCGGTTGCCGGCTGGAGCAAATTCGCGATCAGTCACTACCTCGAC  
TCCACGCGCGTCCGGCGCATGACTGACCGGAAGCACGAGGCGGGCATTGCGGTGTGCACATTTTGCTTCCGCATTGC  
TGGCGCGCGCCTTACCAGCAGGCGTGTGCTGCATCGGGTGTGAAATGTTTCGTGGAACCTACATACTTGAGAGACACGG  
CCCGCTGAAGTCAGATACCGACATCGCCATGCTTCAGCCCGAAAAATGAAATCGACCTCCACCGAAGCCCTTTGGGCAAT  
TGCAGCACACCGACCGACGTACGCGTATGAACGCCTGCGTCGCCAAGCACCGCATACAACATGTCGGATGCACCATCGGC  
GACTTCACTTTGCATCGTGAAGCCGGGTGCGCTTCGCACATAGACGGTGATGCGCGGCACTGTGCTATCGCGTCGAGCG  
TGCCGCGAGTGCTTCTGATCAACGACAACGCACGCAGCGCCGATACGCCGGCGGCGCCGCGACCTTCGTGAGCGAAACG  
GATTCCCCCACTATGCCACGAAATGCACCACGTGCGCGATCCGCGGAATCTGGCCCGCAACGTAGGCGGTATTCCCGTC  
GACCAGAATGGGCGTGACTTTCCGCCAATCCTGATTTCTCGTCCGATTGAAGCCGAATTCGGTCGCGACCTGTGTCA  
GCTTATCGTCTCGGTTTCATCAGTGAATCCCCGATACTTCAGAAAAATGTAATAATAATTAAGCCTCGCCCGCTTGTCT  
CGAGGAAGAAATCCCGCCGCGAACGATCAAAGAGACGTGCCGCTGCCGTAAAAAGTTTCCGGTAATCCTCTACGTTTTTC

TCATACATGACACACTTTTCAGGTGTTTGAAGCAGTCCAGCATCGCTTTATCGATGCCAAATCTCGTAGTCGACGAGAAT  
TTGCTCAAAATTTTCCGCTGAAAACGCTATCATATGTGTACTCAACCTGAAGAGTGATATATGGCAACGTACAAGGAAC  
GAAAGCTCAAATGGAAGCTTTGGCCGAGAAAGCCGAAGCTGCACGTGCTGCGGAATTCCAGGCGATCGTCGACGACATCC  
GCACCAAAGTTGCGGAATTCCGGCATTACCGAGAAAGACATTTTCGGTACGCGTCTGGTTCGACCGGCCAAGCAAGCGTCC  
GCCCCGGTGCAGGCAAAGTATCGCGACCCGAAGACCGGAGCAACCTGGTGGGGCCGCGGTGCGCGCCCCGATGGATCAA  
GGACGCGAAGAATCGCAATCGATTCTGTATCCAGGAATAAGATTCTCGACGCGCACCGGCGCGCGCTCTGGCGCGCCGTC  
CTGCTATTTGAGCATCGATACCGGCGGTATTAAGCCGCTCCGGTGGCGTGAAGTTTGTGCCATCCGGCGCGTGGTTGC  
ACGCTACGCGTCCGATTGCGTGTGCCGATCTTTCCCGATTCAACGTTCTGTAATAGATCGCAAGCCAGACGGTTTTCGT  
TTCTTCGTGCGTCCACGCGACGCGATGACGGCAATGCGGCTCGATCAGCACGTAGTCGCCGGGATGCATTTTCATGACGCG  
CCGGGCTGCTTCGAATTCGAGTACGGCCGCGCCGATAGCAGCACGACCCATTCCGGCGCGTGGTTCGTCATACCAGAAG  
CCGTGCGGGCTCGTGTGCCCATCGATACGATTGCTCGACGTTTCAGACGCTGCCCGGTGACAAGGAGTGCATTTGTTTC  
ATCGTCGCCGCGTGGCCTTCCGTGGTGAACAGGTTGCCGTTTGAAGGTGCATGGGGCGGCCTCGTCGATGATCGTCGG  
TGGAAGGAATCGCGCGTTGAGTGACGCGCAATATTACCGTAGTCTAACGATCACCGTCGACGCTGACTCGTTTCGTTCGG  
AAAAACGATCGTAGCGTGTGCGTGCCTGCAATCGGGCATCAATCCGTTCAGCGAGATCCCGTACTGTTCCGGATCTCG  
CCCACCGAACGTGTTGCGATTGATGCCGTCAACATCAGCATGCTTTTCATACCGCCTCCCGTTCACACTCGATCCAGCGGA  
ACGGTAAGCAGGTCGATACCGCTTTCGTCTCGGGGCGCACGCCGCGCCACACGCAAAGGCTTCGCCCGCCTGCTCCAC  
CAGCATGCCGACACCATCGGCTACGCCGGCGACCCCGCATCGCGCGCCAGTCGCAAGAACGGCGTGAGGCGCTTGCCAT  
AGGCCAGTTTCATAAGCCGTACCATGCGGGCTGAAAACGCCGGGCGGCACGGGCGGCAAATCGCCTGTCAGACTCGCGGAC  
GTCGCGTTGACACACGATCGAAGCGGCCATGCGGTGAGCTCCCGTAAGCGCACGCCGTTATCGGGCCGTACGAAGC  
CAATTGCAATGCCAGCGCTTGCGCTTCTCGACGTGACGATTGCCATCACAGTTGCGCGGGGCGCGCCAGGAACG  
GAAGCAACGCGCGCGCGCGCGCGCGCGCGCGCACCGAGCAGCAACACACGCTTGCCCGCGAGCGGCGGTGTAAGATTACCC  
TCGATATCGCGCAGCAGACCGATGCCGTGCAAGTTTTCCGCGAGAATCCGATCCCCCTCGAATTGAACGCATTGCGCGC  
GCCTGCGAGCGTCGCGCGCTCGCTGCGCTCATCGGCCATGTGCAACGCCGCGAGCTTGAACGGTGC>gi|142022655|  
gb|EQ086233.1|210 marine metagenome JCVI\_SCAF\_1096627390048 genomic scaffold,  
whole genome shotgun sequenceGAGCGACGGGAAAAGACGTACGGGAAACTCACGCCGTCGACGGCGCGCCGG  
GCGAAGGAGCTCCTGCTCGAGCATCTCGACGGCAACCTCGGGCTGGCCGACGTGGCGAGCGCCTGCGGTCTGTCTCGCAG  
CTATTTTCGTGAAGGCGTTTCATCAGACGACCGGGCTGCCGCCGATCGCTGGCTCATCGTGCAGCGCGTCGAACGGGCTA  
AGGAGTGGATGCGCAATTCGGGCCTGCCGTTGAGCATGATCGCCGACGCGTGGGTTTCGCCGATCAGAGCCACTTCTCG  
CGGACCTTCACGCGCCTGACCGGCTCAGCCCTCGCCGCTGGCGCGCCGACAATACCTGATCGTCTGATGCGGGGCCGG  
GTAGGTGAGGTTTCGCTTCGTTCGGTATCGGAGCGTATGCATCTGACCGGCGGATTGGGCGGAGTGCAGGGCGGGC  
GCGGAAGCGGCGTGCAGGGTGCTTATGCGTAGAAGCCGATCAGCATCTCGAGATTGCGTTGCGCGTGCTGCTCGAGCG  
GCGCGAATTCGCCGATCGAAACAACGGGGCCTGCGCGCGCAGATGCTCGAGCGCCTGCTGCGCGCGCGGGAAGGCT  
GCCGACGGTGCATCGGGATCCTGCGCCCACTCGCGCGGATCGCGCGTCTGATTTCGCGGAGCCGTCGGGTGCCGCCG  
GAGGATCGCGAGATCGCGTCGAGGAAAATTTGCGCGCGCGCTGCAATTCGCGATCGTCGGCGAACCCGTCAGGCAGCC  
GGTGCGACTTCGTGCGCAGCACCAGGTGCGGGCAACGGACACGTGCGATGCATGCGCGTGAACAGGCTGCGTCGCAA  
TGTTTCGTGCGCGACCTGCGCGAGCCATTGCGCGCTGAGCGCTTCGTTGTCCGCATAATCCGGCAGTGTCTGTCGATAGAC  
GACATCGTGCGCCACACGGCGAGCTCGATGGCCGGCCATAGCCGTTCCGCACGATACGGCGGAGGTGCGCGAACAGTT  
CCGCCAGATGCGCGAGCGTGTGATAGAAGCGCTGCGGCTCGCCGTACGCGCGCTCGACGAGCGGCCACACGGTCTGCTG  
CCGAAGCACGCGTGCAGGCGGCGCGCAGCGTGCAGCGGATATCGGAATCGAGGGCGTTTCATGGCGCGGATCGGCTCAGG  
CGGGCGATGCGCGAAGCATAGCAAACGTGGCCGCGAGCAGGCATGCCGCTTGCACGCGTACCGATCGCACCGTATCGG  
AGGTTGCCATGAAACGAATCCGGCCGCGCTGCCCGCTGCTCGCGATCTGCGCGTTCCGGGCGCTCGCGCTGCACGCAGAC  
TTCGCATCGGCGGCGCCGACGTATCGATCAGCCGCGCGGATGACGGAACGACCGCGGCGGGCAGCCGGTGAGCCGGTA  
CACGCTGGCGAACAAGCGCGCGTGACGCTGAAAATCATC>gi|142022655|gb|EQ086233.1|237 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ceGGACCGGATAGCCCTGGTCGCGCGCTTCGCTGAGTGCCGTGTACGCATCCGAGCCGAGCACGACCGAGTACGGGCCGT  
CGACGCCGGAAGCCGAGCGCTTCGAGCGCGTCTGATCGCGTGGGATATGCGCTGACGTGCGCGGCGAGCGTGAGC  
TTGCGGTTTCGACGTGCCTTCGCGGATACCGACGATGCCGCGGCGGGTAGCCGTGAAAATCGCGCCGCTCCTCGGCGAA

CGCGAGCCGCTGCGCGGCGTCCCTTCGCCGGTTGCCAGTCGGCGTCCGCGCGCCGCGCTCGACGCTGTCGATCGCATCGC  
 GGCTCAACTCGAACGGCACCGTCAGCTCGACGATCGTGCGGACTTCGCGCAGCCGCGCATTACACGCTCGCGCGGGGCG  
 GCGACGTCGACGAGGTGCCCGGTGCCGACGGCCGACAGCTCGGGCCCTCGGGGCCGTGACATCGACGACGCGGCGGCC  
 GGCCACGGATCGTTTGAACGTGCGTGCCAATTCTCTCTCGATTTGCTCCCAGGCGGAGGACGAGATCGGGGCGAGTTTCGC  
 GGTGCAGGTTATTCATGGGTTTGGGTTGCCTCACGTCAGCGGAAACACTTCATGGTGGGCCGATCGGCTGACGGAAGTGC  
 CTGAGCCCCGGCGGGCCGATCGCGGACGAAAATCAATATTAGCAGGCTCTATTGACACCGCTGCAGCACAAATCCGGACGA  
 CATCCTGCGGCGGCGCAACTGTTCTGCAATTCGCGTGTGGGGCGACGGAGCGTATGCGCTGAACGGTGCACCGGATGAC  
 GATACCCCGGGGGCGCGCCGGCGCTCGCGTCCGCGGGCCGGGCCCGATGGCGCGCCGAACCTGTTGGCGGGCGACGGCC  
 CATCGCGTGTCTGTCATTCTTGCGAACATTTCATCCGGCAGGGTTCCCGGACGCCGGGGCTCGGCGGTATTAGTTGCACG  
 GACCGTGCCCGGATGAATCATAGACGCAGCCTATTGAAGCTTCACGACAAGTGAGGGGCGTTTCATCGCATGGCCGCGCGT  
 CTCGCCCCGGCGCGCGGTGGGCGCGCCGACTTTGGCGATGGGCGTTACGACGGCACGCCCTTCAGCGAGCCGTGCGGATG  
 CGGTGCGAATGACGCCGTTCCGGCGCCGCGGCATCGGCGCGCGGCACGAAGAACAGCGAACCGGTAACCGCGCGGCTGT  
 AGTCGAGCAGGCGGTGCTAATTGCCGAATTCCGGCAGCTATATCACTGGATATGCGCGCTCGCCGGCACTGATCAAAACAGA  
 TACTCGAGAACGTGTTCTGTTGCCACCTGCGTCCGTAGGTCCAACGTAGCAACTCTTTTTGCCGTTGAAAAGGTTGGCAGC  
 CTGTGCCTAACACCGGCTGCCAACCGGATGATGAAACGAATAAATTGCTGTGTGATTTAGTCGAGCGGACCTCTCCGTTT  
 CATCCTAAGATCAAAAACGACTTGGCCAGTTTCGGTATCCACCGTCGCAGCGTAAGACAGCAGTTCTGTTCCCGTCGACAT  
 TGAAGCCTTCACTGAACACGTGCAGTACCGAAGCGAGCGGCACTTGAATTCGTGCGATGTTCCGGCATCTGCGCCGTAGGTG  
 TCGACGTGCCCCGACAGATTGGGTCTGATACTTCGTTCATGGTGATATCTTGAGAAAGTAAATCTGCTCAGACTGCGCCTTC  
 TGGAAAGTCGTTCCCGGAGGCATTGATTATCTGAGGGTTGAAACGAAGAGATCGTCAAATTCGGACGCTTCGACTGTGG  
 CCCATCTTGCGAAGTTCGACCGTGCTTTGTGCGGAGAGATGGCTCCCGATCTTCTCTATCACATCTGCGTTTTCGATAACA  
 AGTTGTTTTGTAGTGTCTATAAAGTTCGCAGCGTAGTAATTTTACCGCTCGATCGCGGGCTGCCGGTTACATAATGCAGG  
 ATTTGATGCACATTGCAGGTTGACCAGAAGCTCAAGATGCTCAGCTGCTTTAGCAAAATCTGATACTGCGGTTTCAATTT  
 CAGCCTTCCGAGCGATCTCGATTGAACTTGACCGTTTGAGAGCGATTCTGCGAGTACCCCGCGAATAGCACTTGGACC  
 CGATCTCTGAGATATCCAGAAATTGCATCTGTCTGCGGACGGGACGATGAAGTATGACTGTGGCGCTTCCGTCGTGCGG  
 TACATCTTCCGTCTCATGCCCGCTGCCCTGGAACGAAGTTCGAGCGGGCCCATCATGAAGCCCAGCCTTCGAGCAACGA  
 CGTGGTGCCCGGCTTCATGTTGGCCGATTGTTGCAATTGAGATTCTTCTTTTCTTTCCATTCTGGAAGGGTGTCTGTA  
 TTCATGGGCTTATCTTCTGTCCGCTGTGTTGTACGTAGCGCGATCTCATACTTGAGGTGGTGACATCTTTGGGTGCTA  
 ACGATTGGAACCAAGTCACTACGCTCCAGCCGACGTAGCGGCGCGGATTCTCATTTCGTACGTTTCGGATCCGAGCAGG  
 ATGCTCACGTCTTCTAGCAGACGGGAATTGGGACGGGACTGCGTAGTCAAGTAAGCAGTCCCGGTCTCTCGAATTCAAT  
 GCGACATTTTCGGGGGCGGCCGAAAAGCGCAACTTTACCGGCTACCTTTGATTGACCAAACCTTGGTGCAATTTTCGG  
 CCATCGCAGCTTGACGTTTTTCAACCGCATCCGTGTTCCAAGTGCTAAGCTTGCTTACCCCGTTCGTAATAATGAAATGCT  
 GAAGTTGAGAAACTGGATTTTTTCTTTGAGAATTCCTCGTTCCCAATGGTGCTGTTGATTTTTGCTCCCATGATCGCCAT  
 GTTGCTAACCCTTTTGTAAAGCTTTGACATCGTCAAGTGACCATGTCTTCATCCATTTTTCAGACGATTGATGGGGA  
 GAATGTGCTCTAAATTAACCTTGACCAGCGTCGGTGCTTGGCACAAGCTCGCCGAGGCGCCAGGCACACCACTCTCAAGC  
 ATCCGGAGGTAGTATCTGGCAAGTGACTGCTTCGATACTGTAGCTACGGAAAAACGCTGCAATGAATGTGGGGTCAGTCGG  
 TAGAGTGGTAAATCCTTTCAATATATCTTGGATTGAATTGAGCTTTCCTCTGTAACCCCTTTTGTGTCATCGCTGTAGA  
 TTCGCTCTAAGGTGCCACCGCTGCGCCACCAACGATTTGGAATCTTACGGCGACCGCAACAAGCTTTTAAATGCTGCT  
 GATACTTTTTTGGGCTCGAATTTATCTAGGATTGCAAGCAAAAGCGGGCGAAGTTGAACCATCCAAGAAGGTTTCATT>g  
 i|142022655|gb|EQ086233.1|507 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
 scaffold, whole genome shotgun sequenceCAGGCAGCGTTGCGCGCGTGACATGAATCATCGGGCGGCTA  
 TCGAGGTACAACCAGTAGCCGTGACCCCTGAACGGCGGTGAGGCCCGGTATCGAGACCGACGACCTCATCGAAAAAGCG  
 GCGTGTACCGTCAATGTATCGGTAACGATGGTGGCGTGGTCAAAATGCATGCTTCGTGTCCATGCTGGAAAGCCTGGTA  
 CGCATTCTGTGCGACGTGCGTGGTCTTGATAATTGGCTTCGCCATTGAATGATCTTCAATGGCGACAAGAAGGTGGGGG  
 GCCGCGACGGGCGCCCGATCCTACAGCCCTGCGGCGCATCTCGCAGCACGGCGCGTAGCGCGGCATCGAACGGCGTCGC  
 GTGGTCGATCCCGGCGGCACGCACGCGGATGCGTCCAGGTGGCAGTCATATGGGCGCGGTGTGCGTGGTGGTGGTGGT  
 CGATTGCTTGAGCGCGGCTTCGATGCCGAGCGCGGCCGCGATCCGCTGCGCAATGTGCTACTTCGTCATCGGCTCCTCG  
 CCCGACCAGTGACGAAGGCCCGTGATCGACGCGCCCGCGAGATGGCGCAGCGTGAGGTGCGGATGACGTCCGCGACATC

CGGCGTATACGTCGGGTAACGGATCGCCCATGCGTCCATGCCGACCGCGTCGGCGCCCGGCCGAGCGGCAGCGACGATCG  
 CCGGCACGAGGCTCGTGACGGCCGATTGCTCCAGTCGACGATCGGGCCGAAGAGCAGCGGCAGCGCAGCACACACGAC  
 AGCGGCGACGCGGCGAGCAGCGCCGCTCGCCTTCCAGCTTCGTGCGGCCGTAATGTTACAGCGGTTTCGGCGCGGCGTC  
 CTCGCGGTACGGGGCAGCCTTGCCGTGCAAAACGTAATCGGTGGAGATGCCGAGCGTCCATGCGCCGTAGCGCGCGGCGA  
 GCACGCCGATGCGGGCCTGGGCCGTGACGTTGATGGCGCGTGCGGCCGCGGGATCGCGTTCGCAGACATCCGGGGCGGCGC  
 TCGGCCGCGCAGATGATCACGGCGCCGCTCGGCAGGTGGCGAAGAGGTGTTGAGGGCTGGCTCGTCGAGGACGTGAG  
 TTCGGCGAGGTTTTCCGGGGCAGCGCCAGGCGTTTTGCGCCTGCGCTCTGCGGGTTCCGGATGGTCGCGACGAGGTTCA  
 GCGACGATTACGGGCCAGGGATGCGGCCACCGCGCGGCCGAGCAGGCCGCGCGCCGATGAGGAGGACGGTCGGGAGG  
 TCAGTCGACGAGGTTCGGGTTTGGCATGGGCGGTCAATCGGGTGACTTGGAAGTGACGAGGTTCGGTTCGACGCTCGGGC  
 CCGGGTGGCGGGGAGCATAGCCGAGTATGACAGTCGGTACGGCGTGATGACAGATCGGCAACCCCGCGACGACGTGCGT  
 CATGTGCGCCGCGCGCGCTCGGCACGGCGCCCCCGTCCACCGGACGACGCCCGCTCTGCCGGCGACGCCGACACGC  
 CGCGCGGCCCGCCCTGCGCGCGAGCAAGCCGCGGCGCGCCGCTCACGAAGATCGCCAAAAGCAAAGCCCGGTGCGCT  
 TGACGCAACCGGGCCGTACCTCTGCGGAAAACCGCGCCCCGCGTTACTTCCGGTCGAGCGAATAGCGCCCCGGCCCGGT  
 CAGTGCGAGCAACAGCAGCCCGCGATGATGCTAACGTTCTTGTAGAAGTTGATCATCGCCATGTATTGCTCCATCCCTT  
 GCAACGCCAGTAGCGATGGCCGATCAACGCGGTGCGAGCGTATAGCGCGCAACACCAGTGCAGCGGACGCGTGTAG  
 AAGCCGATCGCGATCAGCAGGCCACCGGCCAGCTCGACCGCCACCGCGATCGCGGCCCGCAGCTCGGGCGCGGGTTCCC  
 CGTCGACGCCATGTACGCGACCGTGCCCGAGAAGCCGGTCAGCTTCTGCCAGCCGAACAGCACGAACAGGATCATCAACA  
 GCACGCGAGCCGCCAGCAGCAGCTCGTCCTTCTTCGACTCCAGTGAAACGTAACGCATGACAATCACCTTGAATTGACG  
 GTTTGATGCGTCCGGTGCCGAGATGCGCCGAATCGGCGCGCCAGGTTCGGCAACGAACCACTCCTCGCACGATCGGACAGC  
 AAG>gi|142022655|gb|EQ086233.1|350 marine metagenome JCVI\_SCAF\_1096627390048  
 genomic scaffold, whole genome shotgun sequenceGCAATGGGTGTGAGGAAGCGGCGCCGACGCAGC  
 CCCGATTTTTCGAATCTTCAATCCTGTTTTGCAGGGGGAAAAGTTCGACTCGAGCGGCGAATACACGCGCGGTGGGTGC  
 CCGAACTGTCTGCACTCCCCGCCAAAACACTCCACAGCCCATGGACCGCTTCCGCTGAAGCGCTGCGGGAAGCATCCATC  
 GCGCTCGGCGCAACTTATCCCTTCCCGATCGTGATCACCAGGTTGACACGCGCGGAGCCCTCAGTTCGGTAGAGATCCT  
 GAATACACAGAGCCAAATGAAGTGACCTTCGATCGCTAGCCGGGCTCGACCGTGCGACGGGGCGCCCCGACGACATACCG  
 GGGGCGACGTGATGTGCAAGCGCGACGGCCTGCTTCCGCACATCACGCCTTCTGCGGATACAGGTGAAGCATCAGGG  
 CATCAGCACGGAGTCGACCACGTGAATCACGCCATTGGACTGCATGACATCGCCGATCGTGACGTGCGCGACGTTGCCCT  
 TGTCGTGCGGTACAGGCCAGCCGTTTCGCGCCACGGCTGACGATCAGGGAATCGCCTTCGACCGTCTTCAGCGTTGCCCTG  
 CCGCCACCTTGCTCGACCGCCCTGGCAAGGTGCAATGCCGTGAGTCGTCCGGATACGACGTGATACGTCAACACCTTGAC  
 GAGCATCGCCTTGTTCTCCGGCTTCAACAACGTCTGCACCGTCCGGGCCGGAACGCGGCGAATGCGTCGTGGTTCGGTG  
 CAAATACCGTGAAGGGCCCTTGCTCGACAATGTGTGACACAGCCACCCGCTTGACCGCAGCGACGAGCGTCGTGTGA  
 TCGTGAGAATTACCGCGTTTTTCGACGATGTTCTCGACGGGTACATCGCAGCGCCACCCACCTCGACCGTCTGCTCCAT  
 CGACATGGATCCTGCCTCGGTGGCAGGCGGTACACGCACGTTGCGAATGCAAGGGCGAGCGCCGACCAACCCAGTTTAA  
 CTTTGCTGCAAGATTTCAATTTTCGTTCTCCGAAAGATTGCCGCCGTTGACTGGAAGCAATCGGAGATACGAGTCGAACCG  
 CGCAGCGGATGCGCGATGAGTACTTCATCGCGGACAGATCCGCGATCCAGGGAATGCCGAGGGCGTCGACAACCGTTCGC  
 ATCTTCAATCGCCACTGCATCCGTTCTCGCTCGGCCACCGTATCTGCAATGTGATCCAGCCAAGGTGCTTGGCGTCCCA  
 CCTCGGGAATGGCGTCGCGATGGCATTGAGAGGAGACAACCTCATGAAGCTTTTCATCGGTATCGTCGTTGCCGACGCG  
 CGCGTTTCGCGCTCAGCGCTTGGGAGGCGGAGACGACATTGCGCCCGTGGCAGCGCAACATCCGGAGAATCTGGCCATC  
 TGGACCGCGCTCGGCGGTGATTCGGCGGCCCCCGGCGTAGTGGCCGGCGTGGTCAATGCGGCTGTGCGAGGCGCTTTTGGC  
 GGACCCGGTCAAGCGCCCTATTTTGGCGCTGTAACAGGTAATGCCACCCACAGCAGTGCAGCCCATGACACGCCAGCAC  
 GGCTTGAAGCCTGTCTACGCTGCAGTTGCAAGCGCTCCTCGGCGGTCCGTACACCTACCCTGGCCCTGTCAAGGTGGCC  
 GGAGACATCGACGCCAGCCGGAGATGTGCCAGGACATGACGGTGGCCATAACGACGTGCGCGTTCCAGGCTGCGTGT>  
 gi|142022655|gb|EQ086233.1|245 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
 scaffold, whole genome shotgun sequenceCCAGCCGAGCCGAGTGCCAGCACGCGCAGGCGGCGAGCA  
 GCGGCACCGTGACGATCATCGGAGCAGCCCCAGGCGGCACGGCACGACCAGCCGGTCGGTAAGCGGTACGCGCAGC  
 CGCAGCCCGATCGCGAACAGCGAGACGAGCAGCGCATCTCGACGATTTACGCAGCAGCCGCGCGTCGCTTCGAGGTC  
 GAGGTTACGAGGCCGCGCGCGGCGGCGGAGCGCGATGCCACCCGAGATAGATCATCGCAGTGCTGACGGGCAGGT

GGCGCAGCGCGGACGTCGCGACGCCATGCCAGCAGCACCCTGCCGACGATCAGGTACCAGAGGGTTTCGTGCATGTGT  
GCGGTAACAGATCGATCACGTGGCCGCGTGCCGTACGGCTTGGCGCCGAACGCCTCGCGCAGCTTCCGCTTCGCGCGCT  
CGAGCGTTGCGCGGCGCGCTTTCGGCAGGTTGCGGCCCCGCGCGGTTACGTAGAAGTTCAGCATCGACATCGCCGACTGG  
AACGGCGACGCCTTGCGGCGGCGGCTGCGTTGCGCCGAATGCTTGAGCGACGCGGCGATCGCGGCCGGGTTCATCGGACTT  
GAAGATGTGCGGGTTCGATGTGAGCGCATCGCTTTTTTGCATCACGTGCGGCGGACCAGCGCTTGGGTGGTGCGGATCGG  
TGCCGCGGGCGCGGCCCGGGCGCCGAGTCAGGCGCTTGCCGTGCGCGTGCGGTGGCGGTAGCGGAAGGCATTACGTGGGCT  
CCATTTGGGGGACAGGGGTGTAGTGACGCGTGTGGCGGCGCAAGCCCCGTGCGCGCGGCGCGCTCATCCGTACGACG  
AACCCGTGCGGGGCGGCACGGGTGTTGCTGGAAGGACCGGCAGGACATGCGATCCCGCCATCACGCAGGAGGACACCAT  
GCAATCGCATTCCCGGATCTGGCTGGAACGACGAGATCGACAACGATGCCGATGGCGACACGATCGACGCGGACGATC  
TCGATACGACGGTCCATGTGACGTCGAAACGGGCAGGATCACGTTTACGGCGCGTGCGCGGAGGCGGCCGATCCGCCG  
CCCGAGACCGCACGGCATTGGTTCGAGCTCGTGCTCGACTGCGACACGATGGAGCGCTATGCGGATCTCGACGAAAGCAC  
GCGGATGCGCGTGACGCGATCCTGCACGCCACCGTGGAGGATCTGGTTCGACCGCCTGCCGACGAACACGCGACGCTGA  
CGATCGAGCTGACCGATGCAATGCTCGATGCCGCGAGCCGTCTGCAGTAGCCGCGCTGATCAGCCTTTCGATACGCGATT  
TGCG>gi|142022655|gb|EQ086233.1|279 marine metagenome JCVI\_SCAF\_1096627390048  
genomic scaffold, whole genome shotgun sequenceAGGCTTGACCTTGTGCGGCAGCCGCAGCCGCC  
GGCCGCGAGGTTGCCGTACGGATCCTCGTCGCGCCGCGCTCGCGGGGCGCGCCGGCATCACGGGGCGGTCGACGCGACCA  
GCATCGGGGGGAACGGCAACGCATGACCCGCTCGCCGTCAGCACGAACGCACCGTTGGCCACCGCCGCGCGATCGGC  
GGCAGCGCGCGCTCGCTCAACCCGTCGGCTGGGCGTCCGACGGCACGAAGAAGACATCCACCGCGCGCTTCCTGCAT  
GCGTATCGGCGAATAGTCGGCGAAACCGGCGTTGCGGACCGCGCCATGGTCGACGTCGATCGCGAAGCCGGGCTTCGTGCG  
TCGCGAGACCGAACAGCGCGCCGCCCTGGATCTGCGCTTGCGCGCCGTCGGGTTGACGATGCGGCCCGCATACACGCCG  
GCCGTCACGCGATGCACGCGCGGTTGTTGCGCTTCGATCGACACTTCCGTACGTACGCGACGACCGAGCCGGCCGTTTC  
GTGCATCGCGACGCCCCACGCGTGCCCGCGCGGCGAGCTGCGCGCGCCGTAGCCGACTTGTGACGCGCAGCGCGAGCG  
CCTGCCGATGCGCGGCTGCTCGGGGCGGCCAGCCGCTCATCCGTTAGCGACCGGATCCTGCCGCGCGAGTGCGCG  
AGCTCGTCGACAGCGTTTCCATCACGAACGCCGTATGCGAGTTGCCGCCGAGCGCCACGTCTGGACCGGCACGTGCGC  
CTCGGTCTGATGAACCGATACCTGCATCGGGAAGCCGTACGGGCTGTTGTCACGCCTTCGGTCAGGCTCGGATCGGTGC  
CGCGCTTGAGCATCGTCGTGCGTCGAGCGGCGAGCCCTTACGACAGACTGGCCGACGACACGTGCTGCCAGTCGCGC  
ACGGCGCCGCTGCCGTCCACGCCGATGTGACGCGATGCAGCACCATCGGGCGGTAATAGCCGCCGCGCAGATCGTCCTC  
GCGCGTCCAGATCGTCTTGACGGGGCCGAGATGGCCGGCCGCGAGGTACGCGGCGGACACGTGGGCGGCTTCGACACGT  
AGTCCGACGTGCGCGTCGAGCGCCGCCATAGTCGCCGCCGAGGTACGCGTGAAGATCTGGACTTCTCCGGGGCGACG  
CCGAGCGCCTTCGCGACCGCCGCGGTCGGTCTGTC>gi|142022655|gb|EQ086233.1|378 marine  
metagenome JCVI\_SCAF\_1096627390048 genomic scaffold, whole genome shotgun sequen  
ceCATGGCGCGCCTGCTGCACGCGCAACGCCTGTGCGATCCAGAGCGCGGAGTCGACGCGATCACCGCGATGCCGATCG  
ACGCGGCGAACAGCGCGGCATCGTAGCGGATGCCGGTTCATGCGCATCGCGGCCATCCCCGTGTAATGCATGCCGGCG  
ATCCCGGCGCCCATCAGTGCGCGCCACGAGCAGCCGCCAGCCAGCCGTGCGCGCGTGACGACGGTCAGCGCGAA  
GTACGACACGAGCACGGCAATCGCCAGCGACGCGCCCGTATCCGGCCATGCGTAGCCGAGCGGAATCGGCAGCGAAAACG  
CGAGCATGCCGACGAAATGCATCGACCAGATGCCGGTGCCATGGCCACCGCGCCGCCAGCCACGCACGCTTGAGC  
TTCGGATTGTGAGCAGCGAAATGAAGGCGGCCAGGTCGAGTGTGCTATAGGAGGCCAGGTCGCGATCGCGAGCGACAG  
CAGGACGAGCGGGAGATTGTAGGTGCCGTACATGATCGAGCGCCGGATCGAGGTGAGTACGTTGCCGCGGCGCGCA>gi  
|142022655|gb|EQ086233.1|101 marine metagenome JCVI\_SCAF\_1096627390048 genomic  
scaffold, whole genome shotgun sequenceCACATCGACACGAAGATCACCGCGCATGCGTTGCTGATCAC  
GCTCGTCAGCGCGCGTGCCCTCGGACATGAAGCGATCGATGCCAGCAGCAGTGCACACCGGCGACGGGCGAGGTCGGGCA  
TGACGACGAGCGTGCGACACGCGCAACAGCCCGCTTCCGGAACGCCGCGCGCCCTTGACGTCGAGCAGCATGATG  
GCGAGCATCACGGCGATCTGCGACGCGGAAAGGGGCACGTGCGACGCTGCGCGATGAACAACGCGGCGAGCGTCAGATA  
GATCGCGGTACCGTCCAGATTGAACGAATAACCCGCCGCGCAGCAGAGCCCCACGACGCGCTTGTGCGACCCGAGCGATT  
CCAGCTTGACGATCAGGCGTGCGAGAACGGGCTCCGAAGAGGACGTCGCGAGGACGATGAGCAACTTTCGCGCAGGTAG  
CGCAAGAGCCGCCACAGCGCGAAGCCGTGCAGCCGCGGAGCGGGGCGAGCACCAGTGCAGCAACAGCCCGCAGGCCAC  
GTAGAAGGACAGCATCAGCTTCGCGAGCGAGCCGATCGAGCCGATTCCGAAGCGGCCACCGTGAAGGCCATCGCGCCGA

```

ATGCGCCGAGCGGCGGAGCCGATGATCATCGCGAGCACGCGAAAGACGACCTGGGCGACGCCGTCGATCAGTGCAAGA
ACGGGCCCGCCCGGCCGCGGGTGTGCGTTCAGCGAGAAGCCGAACAACAGCGACAGCAGCAGCACCGGCAACACCTCGCC
TTTCTCGAACGCGCCGAGCATCGTATCGGGGATCAGCTCAGCCCGAACGCGACGAGCCCGTTCCGGTTGCGCGTCCCTCA
CGTACGGCGCGAGGATGCGCGAATCGAGATGGTGGACGTCGATGTGCATGCCGGCGCCCGTTGCAACACGAACGCGGTG
ATGAGCCCCGATGGCGAGCGCGACGGCGGTGAGCAGATAGAAGAGCGCGAGCGCCCGCACGATCGTGCGGCCGATTGCCTG
CCCGTTCGCCAGCGACGTGATGCCCAGACGATCGTGCAGAACACGATCGGCGCGATGGTCATCCTCACGAGGCCGACGA
ACGCGTCGCTGAGCGGTTTGAACATCGCGCCTGCGTCCGGCCATACATGACCGACGGTCACGCCGAGCAGCATCGCGAGC
AGGACCTGCACGTAGAGCGATTTGAGCAGCCTGGCCAACCTCACGGAACGGTCTTCCTTCGTGTCTCTTGTGTGCTTGTG
ACACCGATGCATCGCTGTTTCGCGATGCCGTGCGCTTCGCGTCGACGATCATCCGGTGAAATGCTTCGGCGGCCGGCGTG
AGCGGCCGTCCGCGCCGCTTGACGATGCCCACGCGGCGTTTGACGACCGGCTCGACGAGCGGCACGCTCGTGAGAATCGG
GTGGTCGTGCCCGGGCATCGCCATCGACGGCACCGCGGCGACGCCGAGCCCCGCCTCGATCAAGCCGAGCAGGGTCGTCA
CGTGGCGCGCTTCGCATACGCTCGGCCCCCGCGGCGCCACGGCGGCCAGCGCCTGGTCGAGCAACAGGCGGTTCCCGGAA
GTCTTGTCCACCGACACGTAATCGTGCTCGTACAGTTCGTTCCAGGTAACGCGCTTCTTGCGCGGAGCGGATGGTCGCG
GCGGCAGGCAGCGACGAACCGCTCCTGGAGCAACATCTTGAACTCGATGTGCGATTCTGGCTGCCCATGAAGCTCACGC
CGAAATCGGCTTCGCCGCTGATGACGGCGCCAGCACCTCGTTCGCGCTCGCGTCCAGCAGCTTGACCCGGATGCGCGGA
AAGCGCTGATGATAGCGCGCGATGATGGCCGGCAGAAAGTAGTAGCGACCGAGGGCACGCACGCGATGGTCACATGGCC
CAGGCGGCTCGACGACACGTCGCGAATGCCGAGCAGCGCCGATCGAGATCGTCGAGCAGCTGTTTCGGCGCTCTGGGCGA
ACACGCGGCCGACCGTGGTGAGCGGACGCGACGCGTGGTGCGCTCGAACAGGCGCACGCCGAGCGCTTCTCGAGCTTG
TCGATCCGGCGACTCAACGCGGGCTGGGAAATGCTGACCGATTCGCGGCGCTTTCGCGAACTGCCCGTTCCACGACCGC
GCGAAACGCCTGCAAGTCGTTCAAGTCGAAGTTGATCCCCACGGGCGCGTCTCCCCATCTCAGATGGGGCGTATTTTGCA
TGATTTTCGCCGGGCGGCCGCATCGGCGGGCACGCATTTCGCGCCACCTCGATCGCAACCGCGTGCGTGAGCGTGCGGCC
TGGCCGTGATCCGCCGCGCGTCGATGTGATGGGGGGCGTCGAGCCTGCGGGCCTTACGGCGCAAGCAGGCCGCGTGCG
GCCAGATTGGCATACAGCGCGCGCACGCCGAAGGTCCAAGGCGCGATCTCGGTGCACAGCCGCACGGTGTGACGAGTGC
GCCGAGCG'

```

```

[1]: from Bio import SeqIO

def shortest_sequence_length(fasta_file):
    shortest_length = float('inf') # Initialize with infinity
    for record in SeqIO.parse(fasta_file, 'fasta'):
        sequence_length = len(record.seq)
        if sequence_length < shortest_length:
            shortest_length = sequence_length
    return shortest_length

# Example usage:
fasta_file = 'dna2.fasta' # Replace 'example.fasta' with the path to your FASTA file
shortest_length = shortest_sequence_length(fasta_file)
print("Shortest sequence length:", shortest_length)

```

Shortest sequence length: 115

```

[2]: from Bio import SeqIO

def longest_sequence(fasta_file):
    longest_sequence = None

```



```

max_length = 0
for record in SeqIO.parse(fasta_file, 'fasta'):
    sequence_length = len(record.seq)
    if sequence_length > max_length:
        max_length = sequence_length
        longest_sequence = record
return longest_sequence

# Example usage:
fasta_file = 'dna2.fasta' # Replace 'example.fasta' with the path to your
                           ↳ FASTA file
longest_seq = longest_sequence(fasta_file)
print("Longest sequence ID:", longest_seq.id)
print("Longest sequence length:", len(longest_seq))
print("Longest sequence:", longest_seq.seq)

```

Longest sequence ID: gi|142022655|gb|EQ086233.1|255

Longest sequence length: 4894

Longest sequence: CTCGACGCGCTCCGCGTCGAGGTCGCCCCGACGTCTCGCGCAGCAACTGATTCAAAAACAGGC  
CGCCGCTCATGCCGATCTTGCGGTGGATGCGCCACACCGACAGTTTCGATGCCTTCGGCATCGAGCGCTTCCTTCCACGCA  
AGCACGTGCTGGTAGACGCTGTGACGAGCGTGCCGTCGAGATCGAACAGAAAAGACGTTTCAATGCGCATGTGTATCTC  
CTGGCTCGAAAGGGGCGAGCGAACGGTCTGTAAGCGTGTCCGCACATTATCGGCGCGCGGGCATGTCATGACCATGTC  
CCGCGGCCCCGCGACGCGACGCCACCCCGTGCCGCGCCGTGTCTGTGCCGCTGGTACAATCGCGGGCATCGCCGGGCGG  
GGCTCTCCGCGCCGCGCGCCCCCAACCCTCGTCTCGCCGATTCCAGGTATGGCTACACCGGACGCCGTCAGTTCCAAGCA  
CTCGTGGTGGGTGTCTGGCCCTGGCACTCACCGCCTTCATCTTCAATACCACCGAATTCGTGCCGCTCGCGCTGCTCA  
GCGCGATCGGCGACAGCCTGCACATGCAGCCGACCGACGTGCGCCTGATGCTGACGATCTACGCGTGGGCGGTGGCCGTC  
GTGTCCTTGCCGCTGACGCTGGCCACGCGCCACGTGCGAGCGCCGCAAGCTGCTGACGGGGGCATTGCTGGTATTTCATCGC  
GAGCCACGTGCTGACCGGTGTGCGGTGGAATTCGCGGTGCTGATGGTCGGCCGGCTGGGCATCGCATGTGCGCATGCGG  
TGTTCTGGTCGATTTCCGTGCCGCTGGCCGTGCGGCTCGCGCCGAGCGACCGGAAAAAGCCGCGCGCTCAGCCTGCTGGCG  
ATGGGCACGGCGATCGCGATGGTGGCCGGCATTCCGCTCGGGCGCGTGGTCGGCGAGACGTTTCGGCTGGCGCGTCACGTT  
CCTGATCATTGCCGGCGCGGCCGGCGTTCGCGTTCGCTGCTGCTGCGCGGACGTTGCCGGTATCGCCGAGCCAGGGCGCCG  
GGTCGCTCGGCAGCATCGGCGTGTTCCTGCGCAAGCCCGCGCTGGTGGCGTTGTACGCGATCACCGTGCTCGTCGTGTCC  
GCGCACTTCACGTCGTACACGTACATCGAGCCTTTCGTCCAGAGCGTCAACCACGCGAGCAGCAGCCGATCACGTATGT  
GCTGATCCTGTTCCGCGTCGCCGGCATAACCGCCGCGATCTGCTTCAACCGCATCTATCCGCACCGGCCGACGACTTCC  
TGCTCGGGTCGATCGTCGCGCTGGCGGGATGCCTGCTGATCCTGTTCCCGTGCGCGCTGAACATCGTCACGCTGTCCGTG  
CATACGCTGGTGTGGGGCGGGCGATCGTCTGCTTCGGGTGGCGATGCAGGCATGGGTGCTGAAGCTGGCGCCGGAGGG  
GACCGACCTCGCGGTGTGATCTTCTCCGGGCTGTACAACGTGCGGATCGGCGCGGGCGCGCTGATCGGCAACCATATCG  
CCGGCGACTTCGGGCTGCCGTGGATCGGCACGTTCCGGCGCGTGGTCGGCGCGGTGGCCGTGGGATCGCGTGGCTGGCG  
TTGCGGCTGCACGCGAGGCAGGCGGCGGCCTAGCCAGCGCCGCTGGCGGCGTCCGGCATGGGGCGCAGGTCGCCACGGC  
CCGGCCTCGCGCTTCGCGCTTCGCGGGTCATGCCACCGAGTAGCCGCGCTCGGCCACGAGCGCATGACCTGACACATAGC  
TGGAGTCATCCGATGCGAGGAACGCGACGATCGTCGCCATCTCCTCGGCAGAGCCCCAGCGTCCGGCCGGTGTCTGATGCG  
GCAAACGCTTGCTGCGCTTCTTCCGACGGCCAGATGCCGCGATGATGGAACGGCGTTTCGATCAGCCCCGGGCACAGCGC  
GTTACCCGACGCGCGCTTGAACCATTCGATCGAAGCGGTCTTCGTCATCCCGATCACCGCGTGCTTGCTCGCGATAT  
AGACCGACGATTCTCGAAGCCGATCAGCCCCCATCGATGCGTTGTTGATGATGCTGCCCCAACCTGCCGACGATG  
ATTTACGCGGCTACTTCATCGAGTTGAATACACCGCGACGTTCCGCTCGAACACCATGTGAAACCGCTCGGCATCCTG  
CTCGAGCAGCGGCGGAACACGCCTTCCGTGCCCGCATTGTTGAACGCGACGTCGAGGCGGCGTAGGTGATACCGTGA  
AATCGAACAGCTTGCGCAGATCGTCTTCGTTGCGGACGTCGCGACGAAGGCCTTCGCTTCGCCGCGGCGGTACGATT

TCGTCGACCAGTTGATCGAGCTCGGGCTTGCGCCGCGCGGAGACGACGACCTTTGCGCCGCGACGCGCGAGTTCGATGGC  
CGACGCGCGGCCGATGCCGGAGCTGGCGCCCGTTACCAGGGCGATGCGGCCGTGAGTTGGGCCGGTTTCGTGATTGCGT  
TCATGATGAAGTTCTCCAGAGTGATGTTTGAAGCGACGCATGTGGCTGAGTCGAATCGGTTGCGATGCACGCTTTACGA  
GCATTGTGGTCGCTCGCGCGTCGTAAAAAATGGAACAATGGCGCATTGAGAATTCCCGTTTCGGGAATGAAGCGCGAAG  
CCGCATGCGCGCGCCAGGAGGGTTCTCCATGCTCAATCGACTCGACATCCTGAAGATCTTTGCCGCGGCCGCGCGCGC  
CCACGTTCCGCGAGGCGGCCGCGCGCCTCGGCGTGTCGCCGCGAGGTGGTGACACGCGCGGTGCGCGACCTGGAGGAGATG  
CTCGGCGAAACGCTGTTTACCGGACGACCCGGAGCATCCGGATCACCGCGTTCGGGCAAGCCTTCGCGCGCGACGCGCA  
GACGGCGCTGGCTGCCGTGACGGGCTGTTCCGCCCCGCGACCGGCCAGGCCGACGAGCCGGTGGGCGTCGTGCGGATCA  
CGGCGCCGTGCGGCATGGGGCGTCGTTACGTCCAGCCGATCCTGGGCGATCTCATGCAGCGGTACCGGGGGCTCGTGCCC  
GACTTGCGGCTGTCGGACGTGCCGTGCGCGGTGGTTCGACGAGCAGATCGACATCGGCGTGCGCGTCGGGGCCATCGGCGA  
CAACCGCTTCGTGCGCGCATGGTCGGGCCGTTGCCGATGTGGATCGTCGGCGCGCGTCGCTGATCCGGCGGCTGGGTG  
AGCCCAAAACCGCAAGGACCTCGAATCGATGCCGTGACGTGCCTGATCGATCGCGCGAGCGGCCGCGCGTGGCCGTGG  
ATGTTTCGGGGGGAGCAGCAGTTCTGTCCTCGTCGCCGGCGTACCTGACGGACGATACCGAGGTGGAGATCGAGGCCAT  
CTGTGCGGGGTGGGGCTGGGGCAGTGCTCGGAGTACCTCGTCCGGCCGTACGTGACAACGGGCGCCTCGTGCGGCTGC  
TGCCCGGCTCGAGCCCCGAGCCGTGGAAGCTGCACGTCTATCGGCCGCGCGCGGGCCGATGCCCCGGCGCATCCGGGTC  
GTGTACGACGAGCTGGTGGCGAAGCTCGCGGACGCGACGTGGCGGGGTAGGCGGTGCGGCCCGCGCGGGTTCGCGTTG  
GCGCGAAGCGGTGTCAGCCACGGCTCGACGACGTGCGCGCACCTGCGAACCAGTAACGTGCGTCGCGCACGGCTATCCGG  
CCTGGCGTGGTACGCGGCAACCGACATATTGCGGCATCATCATGGACGCTGACGCACCCGCATTTCGGTGCCCCGATATTC  
GACTGGAGACGTTACATGCCGATACCCCGATTGAAACCCGCGACGCGATGGGCGACCCCGATGTGGTCTATCTCAACCC  
GGATCACGTGCTGTGGATGTACAGCCGACCGCTGCGCAGCGGCGCGCGGGCGGCACGGCGATTTCGCGTGGACGACCGCG  
TGAAGGGCGGTTTCGTTGCTGATGGCCGACCATCCGCCGGCCGCCAACTGCTGCAGGATCTCGGGCCGTTTCGTGACGGTG  
ACGCTGGCGAACCCTGCTCCGACTATCCGGACGGTCTCGTCCATGTGCGTGCGCAGGCGATCGTCAAGATCGCGACGGA  
CGACCACGACAAGCCGCTCGCCGAGCGCACGCTCGGCTGGGTCCACGTCAAGGACGGCTCCGCGTTCAAGGTCAGCGATT  
ACGCCGGCGTGCCCGCGCAATGGCAGGCGTCGGTGGCGGGTGCCGGCCGCTGAACGGGGCTTCACGCCTGCGTAATGGCG  
CTGCGGCCCGTCGTGCCGAGTCGGCCGGCCGTGGTTCTCCGCGGCAAAACGGATGCCGTGCCTCGTGACCGCGCGGAA  
CGGCTCAGGTGCCGCTGGCCGGGCGAACCATGCGCCACCCGGTGCCGCATTTCGATACGCATTGCACCGATCCCCGAG  
TAACATGGCGCATCGCGACGAACCGGCGCCCGTTCCGGCCCTTCGCCGCGAACGCGCCCGTCGAAGCGCCCGATGCAGGTG  
TTCTCCACCGGCTTCCCGAGGAACCATCCGTCATGAACCTTTCCGCGTCCGATATCACACCGGAGGAACTCCGGCTGTTT  
GGCGACATCATCGCCCGCTGGAAGAACTGGGGCCGGTTCGTGACGTGAGGACGGTTCGTGCTGCCCGCGATCGCGCAACT  
GCTCCGCGCCGATTTCGCCGCGTCGTTTCGATTGCGACGAACACACCGGCCCTGTGGCGCAACGGCTTCTCGTACAACATCG  
ATCCGTTGCAGATCGCCCGGTACGAAGCGTGGTTCAGCATATCGATCCGGTCACATCGCAACTGCGCGCGCGCCCGGTG  
GCGACCTGCGTCGACGAGGTGGTGAGCCGGTGCGAACGGAACCGACCGAGTTCTACAACGACTTCCTGTGCCGCGACGG  
CATGCATCACGGCATCAACGTCTATGCGTTTCG

```
[3]: from Bio import SeqIO
from Bio.Seq import Seq

def find_longest_orf_in_frame(fasta_file):
    max_orf_length = 0

    # Iterate over each sequence in the FASTA file
    for record in SeqIO.parse(fasta_file, 'fasta'):
        sequence = record.seq
        # Extract ORFs from reading frame 2
        for frame in [2]:
            orfs = sequence[frame:].translate(table=1).split("*")
            for orf in orfs:
```

```

        orf_length = len(orf)
        if orf_length > max_orf_length:
            max_orf_length = orf_length

    return max_orf_length

# Example usage:
fasta_file = 'dna2.fasta' # Replace 'example.fasta' with the path to your
    ↳FASTA file
longest_orf_length = find_longest_orf_in_frame(fasta_file)
print("Length of the longest ORF in reading frame 2:", longest_orf_length)

```

Length of the longest ORF in reading frame 2: 657

[5]: #To find the length of the longest Open Reading Frame (ORF) in a DNA sequence, using Python, you can define a function that scans the sequence for start codons ('ATG') and then scans subsequent codons until it finds a stop codon ('TAA', 'TAG', or 'TGA'). Here's how you can do it:

```

def find_longest_orf(fasta_file):
    longest_orf_length = 0

    # Iterate over the sequence
    i = 0
    while i < len(fasta_file):
        # Find start codon
        if fasta_file[i:i+3] == 'ATG':
            j = i + 3
            while j < len(fasta_file):
                # Find stop codon
                if fasta_file[j:j+3] in ('TAA', 'TAG', 'TGA'):
                    orf_length = j - i + 3
                    if orf_length > longest_orf_length:
                        longest_orf_length = orf_length
                    break
                j += 3
            i += 1

    return longest_orf_length

# Example usage:
fasta_file = 'dna2.fasta' # Example DNA sequence
longest_orf_length = find_longest_orf(fasta_file)
print("Length of the longest ORF:", longest_orf_length)

```

Length of the longest ORF: 0

```
[6]: from Bio import SeqIO

def find_longest_orf_length_in_fasta(fasta_file):
    longest_orf_length = 0
    start_codon = "ATG"
    stop_codons = ["TAA", "TAG", "TGA"]

    # Iterate through sequences in the FASTA file
    for record in SeqIO.parse(fasta_file, "fasta"):
        sequence = str(record.seq)
        # Loop through the sequence
        i = 0
        while i < len(sequence):
            # Find start codon
            if sequence[i:i+3] == start_codon:
                j = i + 3
                # Search for stop codon
                while j < len(sequence):
                    codon = sequence[j:j+3]
                    if codon in stop_codons:
                        orf_length = j - i + 3
                        if orf_length > longest_orf_length:
                            longest_orf_length = orf_length
                        break
                    j += 3
                i += 3

        return longest_orf_length

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
longest_orf_length = find_longest_orf_length_in_fasta(fasta_file)
print("Length of the longest ORF in the FASTA file:", longest_orf_length)
```

Length of the longest ORF in the FASTA file: 2394

```
[9]: def punto6(fa):
    sequences=sequ(fa)

    n = 1
    lengths = []
    for i in sequences:
        # print "["+str(n)+"]"
        orfs = find_orf(i,1) + find_orf(i,2)
        for j in orfs:
            lengths.append(len(j))
        n += 1
```

```
print(max(lengths))

fa="dna2.fasta"
```

```
File "<ipython-input-9-e23534366a80>", line 5
lengths = []
~
```

TabError: inconsistent use of tabs and spaces in indentation

```
[10]: from Bio import SeqIO

def find_longest_orf_length_in_fasta(fasta_file):
    longest_orf_length = 0
    start_codon = "ATG"
    stop_codons = ["TAA", "TAG", "TGA"]

    # Iterate through sequences in the FASTA file
    for record in SeqIO.parse(fasta_file, "fasta"):
        sequence = str(record.seq)
        for frame in range(3): # Check both forward reading frames
            # Loop through the sequence with the current frame
            i = frame
            while i < len(sequence):
                # Find start codon
                if sequence[i:i+3] == start_codon:
                    j = i + 3
                    # Search for stop codon
                    while j < len(sequence):
                        codon = sequence[j:j+3]
                        if codon in stop_codons:
                            orf_length = j - i + 3
                            if orf_length > longest_orf_length:
                                longest_orf_length = orf_length
                            break
                        j += 3
                    i += 3

            return longest_orf_length

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
longest_orf_length = find_longest_orf_length_in_fasta(fasta_file)
print("Length of the longest ORF in the FASTA file:", longest_orf_length)
```

Length of the longest ORF in the FASTA file: 2394

```
[11]: from Bio import SeqIO

def find_longest_orf_length_in_sequence(fasta_file, sequence_id):
    longest_orf_length = 0
    start_codon = "ATG"
    stop_codons = ["TAA", "TAG", "TGA"]

    # Iterate through sequences in the FASTA file
    for record in SeqIO.parse(fasta_file, "fasta"):
        if record.id == sequence_id:
            sequence = str(record.seq)
            # Loop through both forward reading frames
            for frame in range(3):
                # Loop through the sequence with the current frame
                i = frame
                while i < len(sequence):
                    # Find start codon
                    if sequence[i:i+3] == start_codon:
                        j = i + 3
                        # Search for stop codon
                        while j < len(sequence):
                            codon = sequence[j:j+3]
                            if codon in stop_codons:
                                orf_length = j - i + 3
                                if orf_length > longest_orf_length:
                                    longest_orf_length = orf_length
                                break
                            j += 3
                        i += 3
                    break # Exit loop after finding the sequence with the given
↪ identifier

    return longest_orf_length

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
sequence_id = "gi|142022655|gb|EQ086233.1|16" # Replace with the specific
↪ sequence identifier
longest_orf_length = find_longest_orf_length_in_sequence(fasta_file,
↪ sequence_id)
print("Length of the longest forward ORF in sequence", sequence_id, ":",
↪ longest_orf_length)
```

Length of the longest forward ORF in sequence gi|142022655|gb|EQ086233.1|16 :  
1644

```
[12]: from Bio import SeqIO
from collections import defaultdict

def find_most_frequent_repeat(fasta_file):
    repeats_count = defaultdict(int)

    # Iterate through sequences in the FASTA file
    for record in SeqIO.parse(fasta_file, "fasta"):
        sequence = str(record.seq)
        # Find repeats of length 6
        for i in range(len(sequence) - 5):
            repeat = sequence[i:i+6]
            repeats_count[repeat] += 1

    # Find the most frequent repeat
    most_frequent_repeat = max(repeats_count, key=repeats_count.get)
    frequency = repeats_count[most_frequent_repeat]

    return most_frequent_repeat, frequency

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
most_frequent_repeat, frequency = find_most_frequent_repeat(fasta_file)
print("Most frequent repeat of length 6:", most_frequent_repeat)
print("Frequency:", frequency)
```

Most frequent repeat of length 6: GCGCGC  
Frequency: 153

```
[13]: from Bio import SeqIO
from collections import defaultdict

def find_max_occurrences_repeat(fasta_file, repeat_length):
    repeats_count = defaultdict(int)

    # Iterate through sequences in the FASTA file
    for record in SeqIO.parse(fasta_file, "fasta"):
        sequence = str(record.seq)
        # Find repeats of specified length
        for i in range(len(sequence) - repeat_length + 1):
            repeat = sequence[i:i+repeat_length]
            repeats_count[repeat] += 1

    # Find the repeat with the maximum number of occurrences
    max_repeat = max(repeats_count, key=repeats_count.get)
    max_occurrences = repeats_count[max_repeat]
```

```

    return max_repeat, max_occurrences

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
repeat_length = 7
max_repeat, max_occurrences = find_max_occurrences_repeat(fasta_file,
    ↪repeat_length)
print("Repeat of length", repeat_length, "with maximum occurrences:",
    ↪max_repeat)
print("Number of occurrences:", max_occurrences)

```

Repeat of length 7 with maximum occurrences: CGCGCCG  
 Number of occurrences: 63

```

[14]: from Bio import SeqIO
      from collections import defaultdict

      def find_repeats_of_length(fasta_file, repeat_length):
          repeats_count = defaultdict(int)

          # Iterate through sequences in the FASTA file
          for record in SeqIO.parse(fasta_file, "fasta"):
              sequence = str(record.seq)
              # Find repeats of specified length
              for i in range(len(sequence) - repeat_length + 1):
                  repeat = sequence[i:i+repeat_length]
                  repeats_count[repeat] += 1

          # Find the maximum number of occurrences of a repeat of length 12
          max_occurrences = max(repeats_count.values())

          # Find all 12-base sequences that occur Max times
          sequences_with_max_occurrences = [sequence for sequence, count in
    ↪repeats_count.items() if count == max_occurrences]

          return len(sequences_with_max_occurrences)

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
repeat_length = 12
Max = 3 # Specify the number of copies of the most frequent repeat of length 12
num_different_sequences = find_repeats_of_length(fasta_file, repeat_length)
print("Number of different 12-base sequences that occur", Max, "times:",
    ↪num_different_sequences)

```

Number of different 12-base sequences that occur 3 times: 4



```
[19]: from Bio import SeqIO

def find_longest_orf_start_position(fasta_file):
    longest_orf_start = None
    longest_orf_length = 0

    # Iterate through sequences in the FASTA file
    for record in SeqIO.parse(fasta_file, "fasta"):
        sequence = str(record.seq)
        # Get reading frame 3
        reading_frame_3 = sequence[2:]
        # Find start codon (ATG)
        start_codon_index = reading_frame_3.find("ATG")
        while start_codon_index != -1:
            # Search for stop codons (TAA, TAG, TGA)
            stop_codon_indices = [reading_frame_3.find(stop_codon,
→start_codon_index + 3) for stop_codon in ["TAA", "TAG", "TGA"]]
            # Get the first stop codon after the start codon
            stop_codon_index = min(idx for idx in stop_codon_indices if idx !=
→-1)

            # Calculate ORF length
            orf_length = stop_codon_index - start_codon_index
            if orf_length > longest_orf_length:
                longest_orf_length = orf_length
                longest_orf_start = start_codon_index + 3 # Account for the
→offset

            # Find next start codon
            start_codon_index = reading_frame_3.find("ATG", stop_codon_index +
→3)

        return longest_orf_start

# Example usage:
fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
start_position = find_longest_orf_start_position(fasta_file)
print("Starting position of the longest ORF in reading frame 3:",
→start_position)
```

```

↳ -----
ValueError                                Traceback (most recent call
↳last)
```

```
<ipython-input-19-1418364700bb> in <module>
    29 # Example usage:
```

```

30 fasta_file = "dna2.fasta" # Replace with the path to your FASTA file
--> 31 start_position = find_longest_orf_start_position(fasta_file)
32 print("Starting position of the longest ORF in reading frame 3:",
↳start_position)
33

```

```

<ipython-input-19-1418364700bb> in
↳find_longest_orf_start_position(fasta_file)
16         stop_codon_indices = [reading_frame_3.find(stop_codon,
↳start_codon_index + 3) for stop_codon in ["TAA", "TAG", "TGA"]]
17         # Get the first stop codon after the start codon
--> 18         stop_codon_index = min(idx for idx in stop_codon_indices
↳if idx != -1)
19         # Calculate ORF length
20         orf_length = stop_codon_index - start_codon_index

```

ValueError: min() arg is an empty sequence

[ ]: