COMPUTING GC CONTENT & IDENTIFICATION OF DNA

Problem

The GC-content of a DNA string is given by the percentage of symbols in the string that are 'C' or 'G'. For example, the GC-content of **"AGCTATAG" is 37.5%.** Note that the reverse complement of any DNA string has the same GC-content.

Formula for calculating the GC content is:

GC_Percent = (no_of_G + no_of_C)*100/length of the sequence

Given:

A list of 5 DNA strings (Given Below)

Return:

Their corresponding GC percentage as the answer (use Python code) and the name of the gene and organism that DNA sequence belongs to (use NCBI BLAST).

DNA STRINGS:

1. ATGACCGAGTACAAGCTTGTGGTGGTGGGAGCTGGTGGTGTGGGGAAGAGTGCGCTGAC
   CATCCAGCTCATCCAGAACCACTTTGTGGACGAGTATGACCCCACCATCGAGGACTCGT
   ACAGGAAGCAGGTGGTGATCGACGGCGAGACGTGCCTGCTGGACATCCTGGACACTGCG
   GGTCAGGAGGAGTACAGCGCCATGCGGGACCAGTACATGAGGACCGGAGAGGGCTTCCT
   CTGCGTCTTCGCCATCAACAACACCAAGTCCTTCGAGGACATACACCACTACAGGGAGC
   AGATCAAGCGGGTGAAGGACTCCGAGGATGTTCCCATGGTGCTCGTGGGGAACAAGTGT
   GACCTGCCCTCGCGCACGGTGGACACCAAGCAAGCTCAGGACCTGGCCCGCAACTACGG
   CATCCCCTTCATCGAGACCTCTGCTAAAACCAGACAGGGTGTGGACGACGCCTTCTACA
   CGTTAGTGCGGGAAATCCGCAAGCACAAGGAGAAGACCAGCAAAGAGGGGAGGAAGAAG
   AAGAAGAAGAAGTCGAAGGCCAAGTGCGTGGTCATGTGA

ANSWER:

2. GGACTGGGGACAGGGGTCCTGGGGACAGGGGTCCGGGGACAGGGTCCTGGGGACAGGGG
TGTGGGGACAGGGGTCTGGGGACAGGGGTGTGGGGACAGGGGTGTGGGGACAGGGGTCT
GGGGACAGGGGTGTGGGGACAGGGGTCCGGGGACAGGGGTGTGGGGACAGGGGTCTGGG
GACAGGGGTGTGGGGACAGGGGTGTGGGGACAGGGGTCTGGGGACAGGGGTGTGGGGAC
AGGGGTCCTGGGGACAGGGGTGTGGGGACAGGGGTGTGGGGACAGGGGTGTGGGGACAG
GGGTGTGGGGACAGGGGTCCTGGGGATAGGGGTGTGGGGACAGGGGTGTGGGGACAGGG
GTCCCGGGGACAGGGGTGTGGGGACAGGGGTGTGGGGACAGGGGTCCTGGGGACAGGGG
TCTGAGGACAGGGGTGTGGGCACAGGGGTCCTGGGGACAGGGGTCCTGGGGACAGGGGT
CCTGGGGACAGGGGTCTGGGGACAGCAGCGCAAAGAGCCCCGCCCTGCAGCCTCCAGCT
CTCCTGGTCTAATGTGGAAAGTGGCCCAGGTGAGGGCTTTGCTCTCCTGGAGACATTTG
CCCCCAGCTGTGAGCAGGGACAGGTCTGGCCACCGGGCCCCTGGTTAAGACTCTAATGA
CCCGCTGGTCCTGAGGAAGAGGTGCTGACGACCAAGGAGATCTTCCCACAGACCCAGCA
CCAGGGAAATGGTCCGGAAATTGCAGCCTCAGCCCCCAGCCATCTGCCGACCCCCCCAC
CCCGCCCTAATGGGCCAGGCGGCAGGGGTTGACAGGTAGGGGAGATGGGCTCTGAGACT
ATAAAGCCAGCGGGGGCCCAGCAGCCCTCAGCCCTCCAGGACAGGCTGCATCAGAAGAG
GCCATCAAGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGGTTCCAGGGT
GGCTGGACCCCAGGCCCCAGCTCTGCAGCAGGGAGGACGTGGCTGGGCTCGTGAAGCAT
GTGGGGGTGAGCCCAGGGGCCCCAAGGCAGGGCACCTGGCCTTCAGCCTGCCTCAGCCC
TGCCTGTCTCCCAGATCACTGTCCTTCTGCCATGGCCCTGTGGATGCGCCTCCTGCCCC
TGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCCGCAGCCTTTGTGAACCAACAC
CTGTGCGGCTCACACCTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTT
CTACACACCCAAGACCCGCCGGGAGGCAGAGGACCTGCAGGGTGAGCCAACCGCCCATT
GCTGCCCCTGGCCGCCCCCAGCCACCCCCTGCTCCTGGCGCTCCCACCCAGCATGGGCA
GAAGGGGGCAGGAGGCTGCCACCCAGCAGGGGGTCAGGTGCACTTTTTTAAAAAGAAGT
TCTCTTGGTCACGTCCTAAAAGTGACCAGCTCCCTGTGGCCCAGTCAGAATCTCAGCCT
GAGGACGGTGTTGGCTTCGGCAGCCCCGAGATACATCAGAGGGTGGGCACGCTCCTCCC
TCCACTCGCCCCTCAAACAAATGCCCCGCAGCCCATTTCTCCACCCTCATTTGATGACC
GCAGATTCAAGTGTTTTGTTAAGTAAAGTCCTGGGTGACCTGGGGTCACAGGGTGCCCC
ACGCTGCCTGCCTCTGGGCGAACACCCCATCACGCCCGGAGGAGGGCGTGGCTGCCTGC
CTGAGTGGGCCAGACCCCTGTCGCCAGCCTCACGGCAGCTCCATAGTCAGGAGATGGGG
AAGATGCTGGGGACAGGCCCTGGGGAGAAGTACTGGGATCACCTGTTCAGGCTCCCACT
GTGACGCTGCCCCGGGGCGGGGGAAGGAGGTGGGACATGTGGGCGTTGGGGCCTGTAGG
TCCACACCCAGTGTGGGTGACCCTCCCTCTAACCTGGGTCCAGCCCGGCTGGAGATGGG
TGGGAGTGCGACCTAGGGCTGGCGGGCAGGCGGGCACTGTGTCTCCCTGACTGTGTCCT
CCTGTGTCCCTCTGCCTCGCCGCTGTTCCGGAACCTGCTCTGCGCGGCACGTCCTGGCA
GTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGCCCTTGGCCCT
GGAGGGGTCCCTGCAGAAGCGTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCC
TCTACCAGCTGGAGAACTACTGCAACTAGACGCAGCCTGCAGGCAGCCCCACACCCGCC
GCCTCCTGCACCGAGAGAGATGGAATAAAGCCCTTGAACCAGCCCTGCTGTGCCGTCTG
TGTGTCTTGGGGGCCCTGGGCCAAGCCCCACTTCCCGGCACTGTTGTGAGCCCCTCCCA
GCTCTCTCCACGCTCTCTGGGTGCCCACAGGTGCCAACGCCGGCCAGGCCCAGCATGCA
GTGGCTCTCCCCAAAGCGGCCATGCCTGTTGGCTGCCTGCTGCCCCCACCCTGTGGCTC
AGGGTCCAGTATGGGAGCTTCGGGGGTCTCTGAGGGGCCAGG

ANSWER:

3. GTGAAACCCTAAGGAAACTTCGGCAAGCTCCAGAGAAACGAGAGTTCCGTAATCATTTT
CTCTTGTTGTTCCTGATCGCGTAGCTCAAGCGAAAAAAATGGCGTCGTTTGATGAAGCA
CCACCAGGAAACGCCAAGGCCGGTGAGAAGATCTTCAGGACCAAGTGTGCTCAGTGTCA
CACCGTCGAAGCAGGCGCCGGTCACAAACAAGGACCCAATCTAAACGGTCTATTTGGAA
GACAATCTGGTACAACTGCTGGTTACTCTTACTCTGCTGCTAACAAGAACAAAGCTGTG
GAATGGGAAGAGAAGGCCTTGTACGATTACTTGCTCAACCCCAAGAAGTACATACCAGG
TACCAAGATGGTGTTCCCTGGGCTAAAGAAGCCGCAAGACCGTGCTGATCTCATCGCCT
ACTTGAAGGAATCTACTGCGCCTAAGTGAATCTGATGGTGGATTGATTCTGAGTTGTTA
AGTCTCTCTCAGATGAGGCTTTTTTACTTTGCTATATTCTTTTGCCAAATAAAATCTCA
AACTTTTTTTGTTTATCATCAATTACGTTCTTGGTGGGAATTTGGCTGTAATGTGTGCT
GAGGCTCACTTTATCGAGTACATTTGATTTTCTGAGGCTCATTTTCCTTTTTTTTGTCT
AGAAAGCATTTTCTCACTTATTATTGGTGCTTAAATTTGAAGGACAAAAAAGATTCGAA
TTATATATCAACTATTAACAAT

ANSWER:

4. TGTGTTGTGTCAGTGTGTTGTGTGTATAAGTGTGCGTGTGTGTGTGAGTGCTGGTAAGA
AAACTGAGAGAAAAGTGAAAAGTAACGACCGGCAAAAGCGCGAAAAACGAAAAACAAAA
AAAACTTTCCCCGCAGGAAACTCGAACTGCAAGCTGACGTTACAAGCGGTCAAAATTGG
ATTATGCTTAGGCACAAACGCAACTGCCACGCCATTCAGACGCCCAGCGACGCCGAGCA
TGTGAAGTTAAAGCCGCATTTTCATCCGCCCCAGTGGCTCCTGCACCGCGTCACCTTCT
CTTTGGAGCTGTATCACAAGAATATCATCAAGAAACTGCCCAGTCTTGGCCAATTTCAC
GTTTACCGGCTGACCAAGAACGTCTGCCAGACTTAAGTAATACCCATAACAAATATCCC
CATATAGTCCGCAATAAACTCCACAAAAAAGACAAAGTGCGTTTAATAACCATTACAAA
AAAAGTGTAACTGATTGGGAAGAAGTGTGCAACAAATATCCACAAAATCAAATACTCAA
GTGCAATAAAAAAAGTAAAAGTGATTAAAGTGCGAAAAGAAAATGTGGCCAACCATTAA
GGCCAAGATCAAGTTCTAAGTTCAAATAAACAGACGTAAAAAAAGACTAGAAAATCTAG
CGTGTTTTAAAGCGCTGGAAACTAAAATAACCCCCCCTGAAAGTAGTATAACCGAACGT
ACAAAAATGTCCACCGGCCGCCGTTTGGCCAAGCGCTCCATCATCGGCACCAAGGTGTG
CGCCAAGGGTCCGGATGGCCTCTGGTACTCCGGCACCATATCCGACGTGAAAACGCCGC
CCTCGTACAGCGGACCGCTCTCGCCGCCGCCGCCGCCCACTTTTGTGGTGCCCGGCGAG
GCACCGATTAATGCCGATACGCGCTACCTGGTCCGTTTCGATTTCAAGACCGCCGTCGA
GTCCCCAACCGCTACCACGTCGTCCGCGGCCTCGACCTCGTCCACATCCTCCACAGATC
CGTCGGTCATCGTGGAAACGCGTCGCGCTGCCAACGTACACATCAGTCCCGCTCAGGCA
CTGCGTCGCAGCGCCATGATCAAGGAGTTCCGCGAGTCGGATCTTATTGGACCCGGATT
CCGGTCCATCATGGACACCGAACTGCAGCCTGGCCAGCGGGTCTACTTCACCTACAATG
GACGCGAGCAGAGCGGCGATGTCGTCAAACACGACGCTACCAAGGATGAGGTGATTGTC
AAGATCACAACAGTTGGAAATGAGGAACCCATTGAGCTGAAGAAGCGACTGGAGGAAGT
GCGTCTGCTGGAATCGCGACGCTCCGCCCGTCTGGCAGACCAGGATCGCGACACGGACT
TTGCCAGACTCGCCGACATGAGTGGCGAACGCCGCAGAACCACCACACACAGTATTGAG
GTGCCATCGCAGCTGACGGCGCAGCACAATTCCCGGAAACGTCCGCCCAGCGATCACCA
GGACTACGGCAACTATCTGGAAACATGCCGTGCCGCCGAGATTCTGTCATCGATGAAGT
TGCAGAGTCCGCATGGCTGTAGGTCATCCTCCTTTTCCGACCCTGTCACACCCCCTTTT
CGATCACGCCTACATATAAGCTAATTAATTCTATCCATTTCACCCCAGCAATGGCCGAC
AAGTGCTCGAGTCCCGGCAGCAGCTCTTCGGCTTCCTGGAGCTCCGGTTCCCCGTCGCC

GCCATTGAGTGACGACGGCCACGCCCACCACAGCCCACACAATATCATGTCGCCCCACG
ATGCGGACAACGCACGCACACGCACAGCATCCGTGTCCACGTCGGACGAGGGCATTGTC
ATCGACTACAAGGAGGAGCGCAAGAAAAAGGTGGGTGATTCGATCTAAGGCCCCAACAT
CTAGATATAAGAAACATATAGAGATTGGTGTACAAGAATGCGGTCAGAAGAATCTGGGC
TAACGGCGGTGGAGCGTGACATTTTGTTGCATTTTAATTGCAACTTTGTGTATAGATTT
TAATTGCTATAATTATGTAAATGTCGGTCTGTTCTTGGCACGGCCACACGGCGTATGCT
TGATGCCCAATACTCATACGCAATGTAGCGCTGCTGCTGCCGGCGAAAGAATACAGAAT
ATATAGAATATCAAATAAATGTCGACAATGATTCAAGCACAGGAGA

ANSWER:

5. AAACATCGAGGGATTGGATATTGGCGTGTTAGTGAACAATGTCGGGATTCTGCCCAGCC
AAATACCCTGCAAGCTCCTTGAAACATCTGACTTGGAAGAAAGAATATATGACATTGTC
AACTGCAATGTAAAGTCCATGGTTAAGATGTGCAGAATTGTACTACCAGGAATGCAGCA
GAGAAGAAGAGGAGTCATTCTGAATGTGTCTTCTGGAATAGCCAAAATACCATGTCCCA
TTTACACCTTGTATGCAGCATCAAAGGTTTTTGTTGAGAGATTTTCACAAGGTCTTCAA
GCTGAATATATATCCAAGGGTATTATTATTCAGACAGTGGCTCCATTTGGGGTTTCAAC
CGCAATGACAGGACATCAGAAGCCAGATATGGTCACATTCACGGCTGAGGAGTTTGTGA
GAAGTTCGCTGAAGTACCTGAAGACTGGTGACCAAACGTATGGCAGCATCACTCATACT
TTACTGGGCAGGATCGTGCAGTCCATTCCTACCTGGGTCCTGCAGAGTGAAACATTTCA
GCATCACTTTCAGGAATATGTGAAGAACAGGGACAGAAGATGAGAGATGGCATTCTCCG
ACTTTATACTGTATATAGTATTGCACATTTGATATGTGTGTTTCTTTGCACTAATTAAA
ACTGTGTGTAAAAAAAAAACGTAAGACTGGAAAAGAAAAATGACAGGCCTCTGTTTTTC
CATGGTCCTTCAAAATATGCTAAATCAGTGTTGATAATGGAATCATTATTAATGGTAAT
CATATCAGCAGACTGGAGATAGAGGAGTAGTACTGACCTTGATAACATTAACGGAAGGT
CAGTTTCACAAAGGCCAAATCACAGAGTCAGACATGTGACCTTGTTGTTGTTTTTATTA
TGTTTTCTTCCAGATTAGAACGTGTTAAGGCTTATATATTTTCAAACCATGGTTGCATT
GCAAACAACATCCAAAAAACACATATTTGGCGCCACGCAGTGGTTAAATTTGGTTAAAT
CATCTTTAAAAGCTGATGAAATGAAATGTAATAATTTTGTCATTACACTGTAAAAACAA
AAACAAAACAAAAAAATAATAAAAATAAATAAATGGGTTCCACACAATTAATTTGTGTT
GGGGCAAAATGAAAAAAAAATTAAGCTAAGTTATTATTATTATTATTATTTTTATTTTA
TTTTTTTACAAATTTAAGTTGATTGAATATGAAACAATTAAGTAGCCAACCCCTCAAAA
ATTGTGTTGTTTAAGTTCAATTTAAATAGGTAGTATGAACAAACAGCAAAGTGAAAGCC
AGTATTTTACTTTCATGTGATTAAATTCTCTGTAGCCACTTGAGGAAAATGCAAACCCT
TTTTGTGTAGTTAATTATATTATTATTGTATACCAACATATCATAAAGGAAAAAAGGAT
TTGAAGAATGACATTAGAAAAAAGAAATTCTAAATCACTTGAAATTTTCAATGC

ANSWER: