Introduction to the Bioinformatics Armory

一月 26, 2013, 3:03 a.m. by Rosalind Team   **Topics**: Bioinformatics Tools

Let's Be Practical

If you are an accomplished coder, then you can write a separate program for every new task you encounter. In practice, these programs only need to be written once and posted to the web, where those of us who are not great coders can use them quickly and efficiently. In the Armory, we will familiarize ourselves with a sampling of some of the more popular bioinformatics tools taken from "out of the box" software.

To be equitable, we will focus mainly on free, internet-based software and on programs that are compatible with multiple operating systems. The "Problem" section will contain links to this software, with short descriptions about how to use it.

Problem

This initial problem is aimed at familiarizing you with Rosalind's task-solving pipeline. To solve it, you merely have to take a given DNA sequence and find its nucleotide counts; this problem is equivalent to "Counting DNA Nucleotides" in the Stronghold.

Of the many tools for DNA sequence analysis, one of the most popular is the Sequence Manipulation Suite. Commonly known as SMS 2, it comprises a collection of programs for generating, formatting, and analyzing short strands of DNA and polypeptides.

One of the simplest SMS 2 programs, called `DNA stats`, counts the number of occurrences of each nucleotide in a given strand of DNA. An online interface for `DNA stats` can be found here.

   Given: A DNA string s of length at most 1000 bp.

   Return: Four integers (separated by spaces) representing the respective number of times that the symbols 'A', 'C', 'G', and 'T' occur in s. **Note**: You must provide your answer in the format shown in the sample output below.

Sample Dataset

```
AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAAGAGTGTCTGATAGCAGC
```

Sample Output

```
20 12 17 21
```

Programming Shortcut