Os seguintes exercícios foram baseados em desafios disponíveis em <http://rosalind.info> que é uma plataforma para aprendizado de bioinformática através da resolução de problemas. Você pode, se desejar, procurar mais desafios no referido site.

Alguns dos desafios propostos aqui são cópia dos desafios do site e outros foram inspirados no mesmo.

O professor irá disponibilizar a cada vez que for necessário, um arquivo com uma sequência de nucleotídeos para uso no exercício.

1. Uma forma de identificar a espécie a qual um indivíduo pertence é analisar o percentual de moléculas de Citosina e Guanina em uma sequência de DNA. Em uma dada espécie, a concentração destas moléculas tende a variar muito pouco, em torno de 0,1 %. Com base nesta informação, calcule o % de moléculas de Citosina e Guanina na sequência recebida do professor.

Grave no HD um arquivo texto com o percentual dos ácidos encontrados. O nome do arquivo deve ser o seu nome, seguido das letras “CG”.

1. A sequência de DNA que você está recebendo em um arquivo de texto diz respeito a um dos ácidos do par de nucleotídeos. O outro par pode ser deduzido a partir desta sequência, pois a base “A” sempre se ligará a uma base “T”, e uma base “C” sempre se ligará a uma base “G”.

A sequência complementar deve ser apresentada em ordem inversa da sequência principal. O seu programa deverá ler a sequência principal, efetuar as trocas das bases e apresentar a sequência complementar equivalente. Deve também gravar no HD esta sequência complementar, em formato txt. O nome do arquivo deve iniciar com compl e deve ter o mesmo nome do arquivo recebido (ex: arquivo original: **martim.txt**, arquivo complementar: **complmartim.txt**)

1. A criação de uma molécula de RNA pode ser feita simplesmente alterando o ácido Timina (T na sequência DNA recebida) pelo ácido Uracil (U na sequência RNA). Com base na sequência DNA recebida, crie o RNA equivalente. Grave no HD o arquivo RNA equivalente, o qual deve ter a no início do mesmo a sigla RNA e no final o nome do arquivo DNA recebido.
2. Mutações são ocorrências, na sua menor escala, de alterações que ocorrem no DNA. Estas mutações podem ocorrer em um simples ácido, e se benéficas podem trazer vantagens ao portador da célula em questão, se esta for propagada aos descendentes. Dadas duas sequências de DNA, pode-se calcular a distância de Hamming, que nada mais é do que a quantidade de símbolos que diferem em duas sequências de DNA. Com base nas duas sequências recebidas, calcule a diferença de Hamming entre as duas. Obs: a distância será um número inteiro, que deve ser gravado em um arquivo de texto no seguinte padrão: o nome do arquivo deverá iniciar com a sequência de letras DH e após o seu nome.