



Nome do Aluno: \_\_\_\_\_

Nro. de Matrícula: \_\_\_\_\_ - \_\_\_\_

**Peso: 3.0**

---

---

**TESTE - GRAU A**

---

---

1. Faça um **programa RECURSIVO** para achar um fragmento de DNA indicado (fragmento suspeito) dentro de um conjunto com a descrição de pedaços importantes do DNA de uma vítima. A história começa quando... você foi contratado pela ACME para desenvolver um software denominado de “ACME DNA-Buster, The Digital DNA Detective”, onde este software servirá para esclarecer um assassinato que ocorreu dentro do Shopping ACME. Uma mulher foi brutalmente assassinada e um suspeito foi identificado uma semana depois, onde a principal prova são as manchas de sangue encontradas em uma calça usada pelo suspeito no dia do crime (segundo testemunhas). Foram coletadas amostras do sangue da vítima e realizado um exame que permitiu obter uma descrição de pedaços importantes do DNA desta vítima. Esta descrição do DNA é usualmente bastante específica e serve para identificar de “modo único e individual” uma pessoa. O sangue da roupa do suspeito foi coletado, mas infelizmente o DNA extraído foi apenas uma pequena seqüência, onde espera-se que esta seqüência esteja presente no DNA da vítima, confirmando que a mancha de sangue é realmente da pessoa assassinada. Sua missão: verificar se o fragmento de DNA encontrado na roupa do suspeito é parte do DNA do sangue da vítima, confirmando que a mancha de sangue é do sangue da vítima. O programa deve seguir de maneira fiel as definições que seguem logo abaixo:

- **Ler um arquivo texto** contendo dados que representam **seqüências de DNA da vítima** em formato FASTA ([http://www.genomatix.de/online\\_help/help/sequence\\_formats.html](http://www.genomatix.de/online_help/help/sequence_formats.html)). Este arquivo contém uma seqüência de até 50 linhas de texto, onde cada linha possui uma seqüência de DNA de até 100 códigos de ácidos nucleicos. Exemplo deste arquivo (vítima.txt):

```
>AB000263 | descr=Homo sapiens mRNA
TTATTATGTTGCTCAGTTCTGACCTCTTTAGCAACGTACATGACGTACATGAAGTACACG
TATGACGTACATCGTAGCGACTGTCTGAATTTTTCTTCGATGACTAAAATTTCATCTGGGA
AAACAATGCAGACAGCTTGTCTGGCAATTTGTGTGGGTATCAACTGATGTTTCAGTTTGT
```

A primeira linha é uma descrição do arquivo (localização da seqüência no cromossomo e tipo), seguido de linhas (máximo 50 linhas) contendo até no máximo 100 códigos A, G, T, C cada.

- **Ler um outro arquivo texto** contendo dados do **fragmento que se deseja encontrar** no DNA da vítima. Este arquivo contém apenas uma linha de texto com uma seqüência de no máximo 100 códigos A, G, T e C. Pode conter o código “N” (vale-tudo). Exemplo deste arquivo (suspeito.txt):

```
AACTAGACTTCAATCTGGGCAAGCCACTCTG
```

- **Procurar a seqüência do fragmento junto às seqüências da vítima usando uma rotina recursiva.** Considerar que as seqüências contêm apenas os componentes A,G,T,C ou N (vale-tudo). Faça a recursão para a busca “dentro” de uma mesma linha.
- O programa deve gerar como resultado uma **saída na tela que indique se o DNA do fragmento foi achado ou não**, sendo que caso ele tenha sido encontrado, o programa deverá **indicar a posição deste no DNA da vítima (linha e coluna inicial)**. Só devem ser indicadas posições onde o DNA do fragmento tenha sido achado na sua íntegra (não indicar seqüências encontradas pela metade);

Exemplo de funcionamento do programa – executável disponível na Internet

<http://inf.unisinoss.br/~osorio/teste53/> => *dnamatch.exe* e *dnamatch.zip*

Exemplo da tela de saída do programa:

```
>> ACME Dna-Buster      <<
>> Digital DNA Detective <<

Arquivo DNA vitima: vitima.txt
Dados da vitima: >AB000263 | descr=Homo sapiens mRNA
Dados do DNA da vitima lidos...

Arquivo DNA suspeito: suspeito.txt
DNA: AACTAGACTTCAATCTGGGCAAGCCACTCTG
Dados do fragmento de DNA suspeito lidos...

# Processando dados - Match de DNA #

Resultado...
Casamento de DNA - CONFIRMADO
Posicao: Linha = 16 - Coluna = 16

>> ACME DNA-Buster agradece a preferencia <<
```

**Dicas:**

- ⇒ Recursão:
  - Achar uma seqüência é...
  - Termina quando... achou, não achou, sobra pedaço...
- ⇒ **Atenção:** seqüências pesquisadas sempre começam e terminam na mesma linha!
- ⇒ **Atenção:** tratar os 2 problemas - achar onde inicia a seqüência e achar toda a seqüência!
- ⇒ **Atenção:** o código “N” é usado para representar qualquer um dos 4 elementos (ACTG) da seqüência!

**Bom trabalho!**

**Atenção:**

- ⇒ Lembre-se de colocar o seu nome e o número de matrícula como comentários nas primeiras linhas do programa fonte no arquivo “TesteGA.pas”.
- ⇒ Envie o programa fonte da prova (TesteGA.pas) para o professor em *attach* numa mensagem de e-mail. Usar como título da mensagem: “*Subject:* Teste GA-53”. E-mail: fosorio@unisin.br
- ⇒ Todo outro uso do *correio eletrônico* (envio/recepção de mensagens) fica proibido durante a realização da prova. O aluno que enviar mensagens ou ler mensagens recebidas durante a prova ficará com ZERO na nota da prova. Avise o professor quando for enviar a resposta da prova.
- ⇒ Prova INDIVIDUAL e com CONSULTA AO MATERIAL PESSOAL (não é permitido emprestar material ao colega).

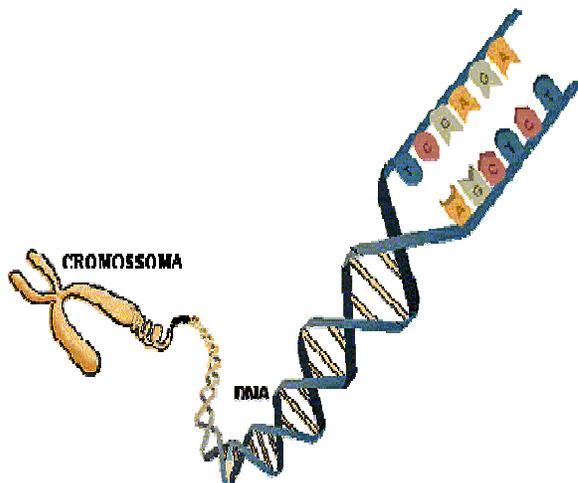


Fig: 1 – Cromossoma e ácidos nucleicos A, G, T, C (ácido desoxirribonucleico – DNA e ácido ribonucleico - RNA)

Tabela 1 - Seqüências de DNA (segundo regras da “International Union of Pure and Applied Chemistry” - IUPAC) são descritas da seguinte maneira:

A = Adenina
C = Cytosina
G = Guanina
T = Thymina
U = Uracil (substitui a Thymina no RNA)
N = A G C T (qualquer um dos 4 elementos)