

# DNA

A análise genômica teve avanços significativos nessa última década, culminando com o completo mapeamento do genoma humano.

Há muitas formas de armazenar os dados gerados por esse mapeamento, mas uma das formas mais comuns de manipular os elementos é com o uso de listas duplamente encadeadas, que contém a informação das bases de DNA na ordem que parecem em um gene. Cada tipo de base é representada por um caracter.

As análises desses dados vão de contagem à simulação de crossover. A contagem precisa ser realizada para cada gene.

E nesse caso, se conta o número de vezes que cada elemento aparece no gene, gerando uma nova lista tendo como informações a base e quantas vezes ela aparece no gene. Computacionalmente falando o gene é representado por uma lista simplesmente encadeada e as informações da contagem, em uma lista ordenada (por quantidade encontrada da base, e do caracter da base para casos de empate) duplamente encadeada. Vale lembrar que um vetor não consegue armazenar a quantidade de informações de um gene e que uma base pode, em tese, não estar presente em um gene

## ENTRADA

Cada linha contém um inteiro  $N$ , de 10 a 248 956 422 pares de bases, que indica a quantidade de pares do gene a ser analisado.

Na sequência há  $N$  linhas da entrada corresponde ao mapeamento de um gene, cada linha contém dois caracteres representando as bases do DNA (cada caracter pode ser A, T, C ou G)

A leitura do 0 indica o fim da análise

## SAÍDA

Para cada gene, são produzidas 4 linhas de saída, contendo o caracter que representa a base e quantas vezes a base apareceu no gene, seguida de uma linha em branco

## EXEMPLOS

### Entrada

72

T C

G A

C T

T T

A G

A T

T T

A G

C C

G T

G G

A C

T C

C C

G G

C G

G G

G T

G A

G C

A A

G C

A T

A A

C A

A C

A G

A T

G C

G A

A C

T A

C A

A G

G A

T A

TG  
TG  
AC  
CC  
TC  
AA  
TC  
GG  
AG  
TC  
CG  
AT  
GG  
AT  
AA  
AG  
CA  
CG  
AT  
GG  
CA  
CA  
CT  
GA  
CT  
TC  
TC  
TC  
AC  
CA  
AT

G C

T G

G A

C T

T T

10

C T

G G

C G

T G

A G

G C

T C

C T

A C

A C

0

## **SAÍDA**

T 32

C 36

G 36

A 40

A 3

T 4

G 6

C 7