

**Instituto de  
Computação**

UNIVERSIDADE ESTADUAL DE CAMPINAS



# Alinhamento de Sequências

MO640 - Biologia Computacional / MC668 - Bioinformática

---

Zanoni Dias

2024

Instituto de Computação

Alinhamento de Sequências

Alinhamento Global

Alinhamento Semi-Global

Alinhamento Local

Pontuação para Alinhamento de Caracteres

Penalidades para Blocos de Buracos

Alinhamento Global usando Espaço Linear

Alinhamento Global de Sequências Similares

Algoritmo Sub-Quadrático para Alinhamento Global

# Alinhamento de Sequências

---

# O que é Alinhamento de Sequências?

- Um alinhamento de duas sequências de caracteres  $\alpha$  e  $\beta$  é obtido inserindo espaços nas sequências, e então colocando uma sobre a outra, de modo que cada caracter ou espaço esteja emparelhado a um único caracter (ou a um espaço) da outra sequência.
- Não podem haver espaços alinhados com espaços.
- Exemplo:
  - Sequências:
    - $\alpha = \text{AAACTGCACAATCTTAATGCCCTTTTAT}$
    - $\beta = \text{GCGGATCAACTTATTCCATCTCTT}$
  - Alinhamento:
    - $\alpha' = \text{AAACTGCA-CAACTCTTCTTAATGCAGTC--CTTTTAT}$
    - $\beta' = \text{--GC-GGATCAACTCT-CTT-ATTCAGTCATCTCTT--}$



# Por que Alinhar Sequências?

- Comparar sequências
- Localizar trechos homólogos
- Predição de estrutura secundária
- Predição de função
- Inferência filogenética

# Como Comparar Alinhamentos?

- Alinhamento 1:
  - $\alpha' = \text{CAGGATGCTAGCAAAAACCATCGCGGGCGATAA}$
  - $\beta' = \text{--GGCATGTAGCACACACGACGCTGGGAGAAT--}$
- Alinhamento 2:
  - $\alpha'' = \text{CAGGATGCTAGCAAAAACCATCGC-GGGCGATAA--}$
  - $\beta'' = \text{-GGCATG-TAGCACACACGA-CGCTGGGAG--AAT}$
- Qual é o melhor?
- Critério objetivo: função de pontuação para alinhamentos.
- Exemplo 1: gap = -1, match = 2, mismatch = -4
  - Alinhamento 1: 3 gaps + 16 matches + 14 mismatches = -27
  - Alinhamento 2: 7 gaps + 22 matches + 6 mismatches = 13
- Exemplo 2: gap = -10, match = 1, mismatch = -1
  - Alinhamento 1: 3 gaps + 16 matches + 14 mismatches = -28
  - Alinhamento 2: 7 gaps + 22 matches + 6 mismatches = -54

## Problema

*Dadas duas sequências  $\alpha$  e  $\beta$  com, respectivamente,  $m$  e  $n$  caracteres, e um critério de pontuação de alinhamentos, deseja-se obter o alinhamento de pontuação máxima entre estas duas sequências.*

- Dado um esquema de pontuação e um alinhamento, é fácil determinar se este é o melhor alinhamento possível entre as duas sequências?
- Como ter certeza que um dado alinhamento é o melhor possível?
  - Calcular todos os alinhamentos possíveis!

# Tipos de Alinhamentos

- *Global*: alinhamento de pontuação máxima envolvendo as duas sequências completas. Desejável em situações onde as sequências são similares, por exemplo, ao se alinhar genes ou proteínas homólogas.
- *Semi-Global (ou Semi-Local)*: não penaliza buracos criados nas pontas das sequências. Desejável, por exemplo, no caso de montagem de genomas, onde busca-se um alinhamento de pontuação máxima entre o prefixo de uma sequência e o sufixo da outra (ou vice-versa), ou quando as sequências possuem tamanhos muito diferentes entre si.
- *Local*: alinhamento de pontuação máxima entre qualquer par de subsequências (das sequências originais). Desejável, por exemplo, para se identificar trechos altamente conservados entre dois genomas.

# Tipos de Alinhamentos

## Global

```
ACTGTGTAACCC - GCAAACCTGTAGCGCCG
A - TCTGTACAACCCTGCAAA - TGTAACGGC - -
```

## Semi-Global

```
ACTGTGTACCCG - GC
                CCCCTGCAAAATTGTACGGCCG
```

```
CGCCGACGACCCCTGCAAACCTGTTACACAT
                CGCCC - TGCAAA
```

## Local

```
ACTGTGTAAACGC - GCAAACCTGTAGCGC
  GAACGAAACCTGCAAAAGGAACCTGCCA
```

# **Alinhamento Global**

---

# Pontuação do Alinhamento Ótimo

- Inicialmente vamos considerar uma simplificação do problema de encontrar o alinhamento de pontuação ótima, onde estamos apenas interessados na pontuação desse alinhamento, e não no alinhamento em si.
- Posteriormente vamos mostrar como podemos construir o alinhamento desejado, com base na solução para obtenção da pontuação do alinhamento ótimo.
- A princípio vamos considerar abordagens de força bruta para obter a pontuação do alinhamento ótimo.

- Onde a primeira base de uma das sequências pode se alinhar?
  - Opção 1:
    - C[CAGCCGAATCGATCGCATG]
    - -[CATCAGCGATCGATCTTT]
  - Opção 2:
    - C[CAGCCGAATCGATCGCATG]
    - C[ATCAGCGATCGATCTTT]
  - Opção 3:
    - -C[CAGCCGAATCGATCGCATG]
    - C-[ATCAGCGATCGATCTTT]
  - Opção 4:
    - -C[CAGCCGAATCGATCGCATG]
    - CA[TCAGCGATCGATCTTT]
  - ...



# Alinhamento Global - Força Bruta - 1ª Abordagem

- Onde a primeira base de uma das sequências pode se alinhar?
  - ...
  - Opção  $2n - 2$ :
    - -----C [CAGCCGAATCGATCGCATG]
    - CATCAGCGATCGATCTT [T]
  - Opção  $2n - 1$ :
    - -----C [CAGCCGAATCGATCGCATG]
    - CATCAGCGATCGATCTT- [T]
  - Opção  $2n$ :
    - -----C [CAGCCGAATCGATCGCATG]
    - CATCAGCGATCGATCTTT []
  - Opção  $2n + 1$ :
    - -----C [CAGCCGAATCGATCGCATG]
    - CATCAGCGATCGATCTTT- []

---

## Algoritmo 1: Align

---

**Input:**  $\alpha, a_i, a_f, \beta, b_i, b_f$

**if**  $a_i > a_f$  **then**

    | **return**  $gap \times (b_f - b_i + 1)$

**end**

**if**  $b_i > b_f$  **then**

    | **return**  $gap \times (a_f - a_i + 1)$

**end**

$max \leftarrow gap + Align(\alpha, a_i + 1, a_f, \beta, b_i, b_f)$

**for** all  $b_k \in [b_i..b_f]$  **do**

$aux \leftarrow Align(\alpha, a_i + 1, a_f, \beta, b_k + 1, b_f)$

$aux1 \leftarrow gap \times (b_k - b_i) + \sigma(\alpha[a_i], \beta[b_k]) + aux$

$aux2 \leftarrow gap \times (b_k - b_i + 2) + aux$

**if**  $aux1 > max$  **then**

        |  $max \leftarrow aux1$

**end**

**if**  $aux2 > max$  **then**

        |  $max \leftarrow aux2$

**end**

**end**

**return**  $max$

---

# Complexidade e Tempo de Execução

- Complexidade:
  - $T(m, n) = \sum_{i=0}^n T(m-1, i) + (n+1)$
  - $T(m, n) \geq \sum_{i=0}^{n-1} T(m-1, i) + (n+1)$
  - $T(m, n) \geq \sum_{i=0}^{\min\{m,n\}-1} T(\min\{m,n\}-1, i) + (\min\{m,n\} + 1)$
  - $T(m, n) \geq \sum_{i=0}^{\min\{m,n\}-1} 2^i \geq 2^{\min\{m,n\}} - 1$
  - $T(m, n) = \Omega(2^{\min\{m,n\}})$
- Exemplo: computador de 3GHz (1 instrução por ciclo)
  - $m = n = 25$ :
    - Tempo:  $2^{25} / (3 \times 2^{30}) = 1$  centésimos de segundo
  - $m = n = 50$ :
    - Tempo:  $2^{50} / (3 \times 2^{30}) = 4$  dias
  - $m = n = 75$ :
    - Tempo:  $2^{75} / (3 \times 2^{30}) = 370$  mil anos
  - $m = n = 100$ :
    - Tempo:  $2^{100} / (3 \times 2^{30}) = 12$  trilhões de anos

- Quais as possíveis opções para a última coluna do alinhamento?
  - Opção 1:
    - [CCAGCCGAATCGATCGCAT] **G**
    - [   CATCAGCGATCGATC] **T**
  - Opção 2:
    - [CCAGCCGAATCGATCGCAT] **G**
    - [   CATCAGCGATCGATCT] **-**
  - Opção 3:
    - [CCAGCCGAATCGATCGCATG] **-**
    - [   CATCAGCGATCGATC] **T**

---

## Algoritmo 2: Align

---

**Input:**  $\alpha, m, \beta, n$

**if**  $m = 0$  **then**

    | **return**  $gap \times n$

**end**

**if**  $n = 0$  **then**

    | **return**  $gap \times m$

**end**

**return**  $max \left\{ \begin{array}{l} Align(\alpha, m - 1, \beta, n - 1) + \sigma(\alpha[m], \beta[n]), \\ Align(\alpha, m - 1, \beta, n) + gap, \\ Align(\alpha, m, \beta, n - 1) + gap \end{array} \right\}$

---

# Complexidade e Tempo de Execução

- Complexidade:
  - $T(m, n) = T(m - 1, n - 1) + T(m, n - 1) + T(m - 1, n) + \Theta(1)$
  - $T(m, n) \geq 3T(m - 1, n - 1) + \Theta(1)$
  - $T(m, n) \geq 3T(\min\{m, n\} - 1, \min\{m, n\} - 1) + \Theta(1)$
  - $T(m, n) = \Omega(3^{\min\{m, n\}})$
- Exemplo: computador de 3GHz (1 instrução por ciclo)
  - $m = n = 10$ :
    - Tempo:  $3^{10} / (3 \times 2^{30}) = 18$  milionésimos de segundo
  - $m = n = 20$ :
    - Tempo:  $3^{20} / (3 \times 2^{30}) = 1$  segundo
  - $m = n = 30$ :
    - Tempo:  $3^{30} / (3 \times 2^{30}) = 18$  horas
  - $m = n = 40$ :
    - Tempo:  $3^{40} / (3 \times 2^{30}) = 120$  anos
  - $m = n = 50$ :
    - Tempo:  $3^{50} / (3 \times 2^{30}) = 7$  milhões de anos

# Contagem de Subproblemas

- O grande problema do método anterior é que o alinhamento de dois prefixos de  $\alpha$  e de  $\beta$  podem ser recalculado muitas vezes.
- Existem quantos subproblemas distintos, envolvendo alinhamentos de prefixos não vazios de  $\alpha$  e de  $\beta$ ?
  - Apenas  $m \times n$  subproblemas!
- Memorização (*memoization*): evita o recálculo de subproblemas, armazenando os valores previamente calculados (usando um vetor, uma matriz ou um *hash*).
- No nosso caso, podemos usar uma matriz para armazenar os valores dos alinhamentos ótimos entre todos prefixos de  $\alpha$  e todos os prefixos de  $\beta$ .

---

## Algoritmo 3: Memoization

---

**Input:**  $\alpha, m, \beta, n$

for all  $i \in [1..n]$  do

    for all  $j \in [1..m]$  do

$M[i, j] \leftarrow -\infty$

    end

end

return *AlignMemoization*( $\alpha, m, \beta, n$ )

---

---

## Algoritmo 4: AlignMemoization

---

**Input:**  $\alpha, m, \beta, n$

if  $m = 0$  then return  $gap \times n$ ;

if  $n = 0$  then return  $gap \times m$ ;

if  $M[n - 1, m - 1] = -\infty$  then  $M[n - 1, m - 1] \leftarrow$  *AlignMemoization*( $\alpha, m - 1, \beta, n - 1$ );

if  $M[n - 1, m] = -\infty$  then  $M[n - 1, m] \leftarrow$  *AlignMemoization*( $\alpha, m, \beta, n - 1$ );

if  $M[n, m - 1] = -\infty$  then  $M[n, m - 1] \leftarrow$  *AlignMemoization*( $\alpha, m - 1, \beta, n$ );

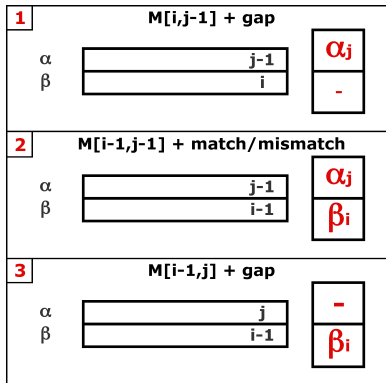
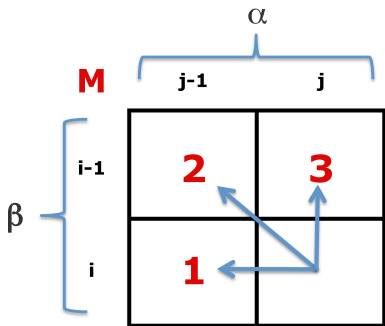
return  $\max \left\{ \begin{array}{l} M[n - 1, m - 1] + \sigma(\alpha[m], \beta[n]), \\ M[n - 1, m] + gap, \\ M[n, m - 1] + gap \end{array} \right\}$

---



- Ideias básicas:
  - Matriz  $M$  armazena os valores dos alinhamentos ótimos entre todos prefixos de  $\alpha$  e todos os prefixos de  $\beta$ .
  - O valor do alinhamento ótimo entre as duas sequências deve estar armazenado na posição  $M[n,m]$ .
  - A matriz  $M$  pode ser preenchida, numa ordem adequada, sem a necessidade de nenhuma chamada recursiva.
  - Para simplificar o algoritmo, podemos armazenar na matriz  $M$  os valores dos alinhamentos ótimos de prefixos de  $\alpha$  ou de  $\beta$  com a sequência vazia.
- Em 1970, Saul Needleman e Christian Wunsch propuseram o algoritmo de programação dinâmica para alinhamento de sequências.

# Matriz de Programação Dinâmica



---

## Algoritmo 5: Global

---

**Input:**  $\alpha, m, \beta, n$

**for** all  $i \in [0..n]$  **do**

$M[i, 0] \leftarrow \text{gap} \times i$

**end**

**for** all  $j \in [1..m]$  **do**

$M[0, j] \leftarrow \text{gap} \times j$

**end**

**for** all  $i \in [1..n]$  **do**

**for** all  $j \in [1..m]$  **do**

$M[i, j] \leftarrow \max \left\{ \begin{array}{l} M[i, j - 1] + \text{gap}, \\ M[i - 1, j - 1] + \sigma(\alpha[j], \beta[i]), \\ M[i - 1, j] + \text{gap} \end{array} \right\}$

**end**

**end**

**return**  $M[n, m]$

---

# Complexidade e Tempo de Execução

- Complexidade:
  - $T(m, n) = \Theta(mn)$
- Exemplo: computador de 3GHz (1 instrução por ciclo)
  - $m = n = 10$ :
    - Tempo:  $(10 \times 10)/(3 \times 2^{30}) = 31$  bilionésimos de segundo
  - $m = n = 100$ :
    - Tempo:  $(100 \times 100)/(3 \times 2^{30}) = 3$  milionésimos de segundo
  - $m = n = 1000$ :
    - Tempo:  $(1000 \times 1000)/(3 \times 2^{30}) = 310$  milionésimos de segundo
  - $m = n = 10000$ :
    - Tempo:  $(10000 \times 10000)/(3 \times 2^{30}) = 31$  milésimos de segundo
  - $m = n = 100000$ :
    - Tempo:  $(100000 \times 100000)/(3 \times 2^{30}) = 3$  segundos
  - $m = n = 1000000$ :
    - Tempo:  $(1000000 \times 1000000)/(3 \times 2^{30}) = 5$  minutos

# Alinhamento Global

$\times$ $\beta$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
	<b>0</b>	<b>-5</b>	<b>-10</b>	<b>-15</b>	<b>-20</b>	<b>-25</b>	<b>-30</b>	<b>-35</b>	<b>-40</b>	<b>-45</b>	<b>-50</b>	<b>-55</b>
<b>A</b>	<b>-5</b>											
<b>T</b>	<b>-10</b>											
<b>T</b>	<b>-15</b>											
<b>G</b>	<b>-20</b>											
<b>G</b>	<b>-25</b>											
<b>C</b>	<b>-30</b>											
<b>C</b>	<b>-35</b>											
<b>A</b>	<b>-40</b>											
<b>C</b>	<b>-45</b>											

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3										
<b>T</b>	-10											
<b>T</b>	-15											
<b>G</b>	-20											
<b>G</b>	-25											
<b>C</b>	-30											
<b>C</b>	-35											
<b>A</b>	-40											
<b>C</b>	-45											

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2									
<b>T</b>	-10											
<b>T</b>	-15											
<b>G</b>	-20											
<b>G</b>	-25											
<b>C</b>	-30											
<b>C</b>	-35											
<b>A</b>	-40											
<b>C</b>	-45											

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7								
<b>T</b>	-10											
<b>T</b>	-15											
<b>G</b>	-20											
<b>G</b>	-25											
<b>C</b>	-30											
<b>C</b>	-35											
<b>A</b>	-40											
<b>C</b>	-45											

Gap = -5

Match = +3

Mismatch = -2



# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	<b>-47</b>
<b>T</b>	-10											
<b>T</b>	-15											
<b>G</b>	-20											
<b>G</b>	-25											
<b>C</b>	-30											
<b>C</b>	-35											
<b>A</b>	-40											
<b>C</b>	-45											

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5					
<b>C</b>	-30											
<b>C</b>	-35											
<b>A</b>	-40											
<b>C</b>	-45											

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	<b>7</b>

$\alpha$  =

**C**

$\beta$  =

**C**

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7

$\alpha$  =

**A** **C**

$\beta$  =

**A** **C**

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7

$\alpha$  =

**A** **A** **C**

$\beta$  =

- **A** **C**

# Alinhamento Global

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7

$\alpha$  = C A A C  
 $\beta$  = C - A C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7

$\alpha$  = T C A A C  
 $\beta$  = - C - A C



# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=						G	T	C	A	A	C
$\beta$	=						C	-	C	-	A	C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=					G	G	T	C	A	A	C
$\beta$	=					G	C	-	C	-	A	C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
<b>A</b>	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
<b>C</b>	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=				G	G	G	T	C	A	A	C
$\beta$	=				G	G	C	-	C	-	A	C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=			T	G	G	G	T	C	A	A	C
$\beta$	=			T	G	G	C	-	C	-	A	C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=		C	T	G	G	G	T	C	A	A	C
$\beta$	=		T	T	G	G	C	-	C	-	A	C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=	A	C	T	G	G	G	T	C	A	A	C
$\beta$	=	A	T	T	G	G	C	-	C	-	A	C

# Alinhamento Global

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7
$\alpha$	=	A	C	T	G	G	G	T	C	A	A	C
$\beta$	=	A	T	T	G	G	C	-	C	-	A	C

- Como vimos no exemplo anterior,  $M[8, 10] = M[8, 9] - 5 = M[7, 9] + 3 = 4$ .
- Logo, na penúltima coluna do alinhamento podemos ter tanto um alinhamento de dois caracteres  $A$  ( $M[7, 9] + 3$ ), quanto um alinhamento de um  $A$  em  $\alpha$  com um buraco em  $\beta$  ( $M[8, 9] - 5$ ).
- Sendo assim, podem existir múltiplos alinhamentos ótimos.
- Dois alinhamentos ótimos tem nomes especiais, considerando a ordem de desempate, caso o valor de um célula  $M[i, j]$  possa ser obtido de duas ou mais formas diferentes:
  - *Downmost*: alinhamento obtido considerando a ordem de desempate  $M[i, j - 1]$ ,  $M[i - 1, j - 1]$  e  $M[i - 1, j]$ .
  - *Upmost*: alinhamento obtido considerando a ordem de desempate  $M[i - 1, j]$ ,  $M[i - 1, j - 1]$  e  $M[i, j - 1]$ .



# Alinhamento Global

$\times$		$\alpha$																			
		G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G	
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95	
	G	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
	G	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
	G	-15	-7	-4	4	-1	-1	-6	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
	T	-20	-12	-9	-1	2	-3	2	-3	-8	-13	-18	-23	-28	-33	-38	-43	-48	-53	-58	-63
	T	-25	-17	-14	-6	-3	0	0	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
	G	-30	-22	-19	-11	-8	0	-2	0	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
	C	-35	-27	-19	-16	-8	-5	-2	-4	-2	1	1	1	-4	-9	-14	-19	-24	-29	-34	-39
	A	-40	-32	-24	-21	-13	-10	-7	-4	-1	-4	4	-1	-1	-1	-6	-11	-16	-21	-26	-31
	C	-45	-37	-29	-26	-18	-15	-12	-9	-6	-3	-1	7	2	-3	-3	-3	-8	-13	-18	-23
	G	-50	-42	-34	-31	-23	-20	-17	-14	-11	-8	-5	2	5	0	-5	0	-5	-5	-10	-15
G	-55	-47	-39	-31	-28	-20	-22	-19	-16	-8	-10	-3	0	3	3	-2	-2	-7	-7	-7	

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global: Alinhamento Downmost

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
G	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
G	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
G	-15	-7	-4	4	-1	-1	-6	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
T	-20	-12	-9	-1	2	-3	2	-3	-8	-13	-18	-23	-28	-33	-38	-43	-48	-53	-58	-63
T	-25	-17	-14	-6	-3	0	0	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
G	-30	-22	-19	-11	-8	0	-2	0	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
C	-35	-27	-19	-16	-8	-5	-2	-4	-2	1	1	1	-4	-9	-14	-19	-24	-29	-34	-39
A	-40	-32	-24	-21	-13	-10	-7	-4	-1	-4	4	-1	-1	-1	-6	-11	-16	-21	-26	-31
C	-45	-37	-29	-26	-18	-15	-12	-9	-6	-3	-1	7	2	-3	-3	-3	-8	-13	-18	-23
C	-50	-42	-34	-31	-23	-20	-17	-14	-11	-8	-5	2	5	0	-5	0	-5	-5	-10	-15
G	-55	-47	-39	-31	-28	-20	-22	-19	-16	-8	-10	-3	0	3	3	-2	-2	-7	-7	-7
$\alpha$	=	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	=	G	-	G	-	G	T	T	-	G	-	C	-	A	-	C	-	C	-	G
-7	=	3	-5	3	-5	3	3	3	-5	3	-5	3	-5	3	-5	3	-5	3	-5	3

# Alinhamento Global: Alinhamento Upmost

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
G	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
G	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
G	-15	-7	-4	4	-1	-1	-6	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
T	-20	-12	-9	-1	2	-3	2	-3	-8	-13	-18	-23	-28	-33	-38	-43	-48	-53	-58	-63
T	-25	-17	-14	-6	-3	0	0	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55
G	-30	-22	-19	-11	-8	0	-2	0	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47
C	-35	-27	-19	-16	-8	-5	-2	-4	-2	1	1	1	-4	-9	-14	-19	-24	-29	-34	-39
A	-40	-32	-24	-21	-13	-10	-7	-4	-1	-4	4	-1	-1	-1	-6	-11	-16	-21	-26	-31
C	-45	-37	-29	-26	-18	-15	-12	-9	-6	-3	-1	7	2	-3	-3	-3	-8	-13	-18	-23
C	-50	-42	-34	-31	-23	-20	-17	-14	-11	-8	-5	2	5	0	-5	0	-5	-5	-10	-15
G	-55	-47	-39	-31	-28	-20	-22	-19	-16	-8	-10	-3	0	3	3	-2	-2	-7	-7	-7
$\alpha$	=	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	=	G	-	G	-	G	-	T	-	-	-	-	T	-	G	C	A	C	C	G
-7	=	3	-5	3	-5	3	-5	3	-5	-5	-5	-5	3	-5	3	3	3	3	3	3

## Exercício

Escreva um algoritmo, em pseudocódigo, que dadas duas sequências  $\alpha$  e  $\beta$ , com, respectivamente,  $m$  e  $n$  caracteres, e uma matriz de pontuação de alinhamentos  $M$ , entre todos os pares de prefixos de  $\alpha$  e de  $\beta$ , retorne um alinhamento ótimo entre  $\alpha$  e  $\beta$ .

## Exercício

Calcule a complexidade assintótica (usando a notação  $\Theta$ ) dos algoritmos Memoization e AlignMemoization.

## Exercício

Prove ou dê um contraexemplo: se  $\alpha$  for muito maior do que  $\beta$  ( $|\alpha| \gg |\beta|$ ), então existe um alinhamento global ótimo entre  $\alpha$  e  $\beta$  que não adiciona buracos em  $\alpha$ . Considere um esquema de pontuação de alinhamento tal que  $\text{gap} \leq \text{mismatch} < 0 < \text{match}$ .

## Exercício

*Dadas duas sequências  $\alpha$  e  $\beta$ , com respectivamente  $m$  e  $n$  caracteres, mostre como determinar em tempo polinomial se o alinhamento ótimo entre  $\alpha$  e  $\beta$  é único.*

## Exercício

*Mostre que, para todo  $k \in \mathbb{N}$ , existem duas sequências  $\alpha$  e  $\beta$ , com respectivamente  $m$  e  $n$  caracteres ( $k \leq n \leq m$ ), tal que o número de alinhamentos ótimos entre  $\alpha$  e  $\beta$  é  $\Omega(2^k)$ .*

## Exercício

*Mostre que, para todo  $k \in \mathbb{N}$ , existem duas sequências  $\alpha$  e  $\beta$ , com respectivamente  $m$  e  $n$  caracteres ( $m, n \in \mathcal{O}(k)$ ), tal que o número de alinhamentos ótimos entre  $\alpha$  e  $\beta$  é  $\Omega(2^k)$ .*

## Problema

*A distância Levenshtein ou distância de edição entre duas seqüências de caracteres é dada pelo número mínimo de operações necessárias para transformar uma seqüência na outra. Uma operação é definida como uma inserção, uma remoção ou uma substituição de um caracter. Foi definida pela primeira vez em 1965 por Vladimir Levenshtein.*

- A distância de edição tem aplicações em várias áreas:
  - Verificação de erros escrita
  - Reconhecimento de fala
  - Análise de DNA
  - Detecção de plágio

## Exemplo

genoma

gnoma (*remoção do "e"*)

gnomo (*substituição do "a" por um "o"*)

gnomos (*inserção do "s"*)

## Exercício

*Dadas duas sequências  $\alpha$  e  $\beta$  com, respectivamente,  $m$  e  $n$  caracteres, mostre como calcular eficientemente a distância de edição entre elas.*

- Custos das operações:
  - Inserção = 1
  - Remoção = 1
  - Substituição = 1
- Importante: a distância de edição é um problema de minimização.
- Como usar o algoritmo de Alinhamento Global como uma caixa preta para calcular a distância de edição entre duas sequências?
  - Match = 0
  - Mismatch = -1
  - Gap = -1



# Distância de Edição

$\times$	$\alpha$	<b>C</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11
<b>A</b>	-1											
<b>T</b>	-2											
<b>T</b>	-3											
<b>G</b>	-4											
<b>G</b>	-5											
<b>C</b>	-6											
<b>C</b>	-7											
<b>A</b>	-8											
<b>C</b>	-9											

Gap = -1

Match = 0

Mismatch = -1

# Distância de Edição

$\times$ $\beta$	$\alpha$	<b>C</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11
<b>A</b>	-1	-1	-2	-3	-4	-5	-6	-7	-8	-8	-9	-10
<b>T</b>	-2	-2	-2	-2	-3	-4	-5	-6	-7	-8	-9	-10
<b>T</b>	-3	-3	-3	-2	-3	-3	-4	-5	-6	-7	-8	-9
<b>G</b>	-4	-4	-4	-3	-2	-3	-3	-4	-5	-6	-7	-8
<b>G</b>	-5	-5	-5	-4	-3	-3	-3	-3	-4	-5	-6	-7
<b>C</b>	-6	-5	-5	-5	-4	-4	-4	-4	-3	-4	-5	-6
<b>C</b>	-7	-6	-5	-6	-5	-5	-5	-5	-4	-4	-5	-5
<b>A</b>	-8	-7	-6	-6	-6	-6	-6	-6	-5	-4	-4	-5
<b>C</b>	-9	-8	-7	-7	-7	-7	-7	-7	-6	-5	-5	-4

Gap = -1

Match = 0

Mismatch = -1

# Distância de Edição

$\times$	$\alpha$	<b>C</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11
<b>A</b>	-1	-1	-2	-3	-4	-5	-6	-7	-8	-8	-9	-10
<b>T</b>	-2	-2	-2	-2	-3	-4	-5	-6	-7	-8	-9	-10
<b>T</b>	-3	-3	-3	-2	-3	-3	-4	-5	-6	-7	-8	-9
<b>G</b>	-4	-4	-4	-3	-2	-3	-3	-4	-5	-6	-7	-8
<b>G</b>	-5	-5	-5	-4	-3	-3	-3	-3	-4	-5	-6	-7
<b>C</b>	-6	-5	-5	-5	-4	-4	-4	-4	-3	-4	-5	-6
<b>C</b>	-7	-6	-5	-6	-5	-5	-5	-5	-4	-4	-5	-5
<b>A</b>	-8	-7	-6	-6	-6	-6	-6	-6	-5	-4	-4	-5
<b>C</b>	-9	-8	-7	-7	-7	-7	-7	-7	-6	-5	-5	-4
$\alpha$	=	<b>C</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>
$\beta$	=	<b>-</b>	<b>A</b>	<b>T</b>	<b>-</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>C</b>	<b>C</b>	<b>A</b>	<b>C</b>

# **Alinhamento Semi-Global**

---

- Como alterar o algoritmo de Needleman e Wunsch para produzir alinhamentos semi-globais?
  - Não penalizar buracos no começo das sequências.
    - Alterar a inicialização da matriz, atribuindo valor zero para o alinhamento de qualquer prefixo com a subsequência vazia.
  - Não penalizar buracos no final das sequências.
    - Buscar o valor do alinhamento ótimo em todas as posições da última linha ou da última coluna da matriz.

---

## Algoritmo 6: Semi-Global

---

**Input:**  $\alpha, m, \beta, n$

**for** *all*  $i \in [0..n]$  **do**

  |  $M[i, 0] \leftarrow 0$

**end**

**for** *all*  $j \in [1..m]$  **do**

  |  $M[0, j] \leftarrow 0$

**end**

**for** *all*  $i \in [1..n]$  **do**

  | **for** *all*  $j \in [1..m]$  **do**

$$M[i, j] \leftarrow \max \left\{ \begin{array}{l} M[i, j-1] + \text{gap}, \\ M[i-1, j-1] + \sigma(\alpha[j], \beta[i]), \\ M[i-1, j] + \text{gap} \end{array} \right\}$$

  | **end**

**end**

$\text{max} \leftarrow -\infty$

**for** *all*  $i \in [0..n]$  **do**

  | **if**  $M[i, m] > \text{max}$  **then**  $\text{max} \leftarrow M[i, m]$ ;

**end**

**for** *all*  $j \in [0..m-1]$  **do**

  | **if**  $M[n, j] > \text{max}$  **then**  $\text{max} \leftarrow M[n, j]$ ;

**end**

**return**  $\text{max}$

---

# Alinhamento Semi-Global

$\times$	$\alpha$	A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A
$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0																				
T	0																				
T	0																				
G	0																				
G	0																				
C	0																				
C	0																				
A	0																				
A	0																				
T	0																				
C	0																				
C	0																				
C	0																				
G	0																				
C	0																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Semi-Global

$\times$	$\alpha$	A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A
$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>C</b>	0	-2	-2	3	-2	-2	3	-2	-2	-2	-2	-2	3	-2	3	-2	3	-2	3	-2	-2
<b>T</b>	0	-2	1	-2	6	1	-2	1	1	1	-4	1	-2	1	-2	1	-2	1	-2	6	1
<b>T</b>	0	-2	1	-1	1	9	4	-1	4	4	-1	-1	-1	-4	-1	-4	-1	-4	-1	1	4
<b>G</b>	0	-2	-4	-1	-3	4	7	7	2	2	2	-3	-3	-3	-6	2	-3	-3	-6	-3	-1
<b>G</b>	0	-2	-4	-6	-3	-1	2	10	5	0	0	0	-5	-5	-5	-3	0	-5	-5	-8	-5
<b>C</b>	0	-2	-4	-1	-6	-5	2	5	8	3	-2	-2	3	-2	-2	-7	0	-2	-2	-7	-10
<b>C</b>	0	-2	-4	-1	-3	-8	-2	0	3	6	1	-4	1	1	1	-4	-4	-2	1	-4	-9
<b>A</b>	0	3	-2	-6	-3	-5	-7	-4	-2	1	9	4	-1	4	-1	-1	-6	-1	-4	-1	-1
<b>A</b>	0	3	1	-4	-8	-5	-7	-9	-6	-4	4	7	2	2	2	-3	-3	-3	-3	-6	2
<b>T</b>	0	-2	6	1	-1	-5	-7	-9	-6	-3	-1	7	5	0	0	0	-5	-5	-5	0	-3
<b>C</b>	0	-2	1	9	4	-1	-2	-7	-11	-8	-5	2	10	5	3	-2	3	-2	-2	-5	-2
<b>C</b>	0	-2	-4	4	7	2	2	-3	-8	-13	-10	-3	5	8	8	3	1	1	1	-4	-7
<b>C</b>	0	-2	-4	-1	2	5	5	0	-5	-10	-15	-8	0	3	11	6	6	1	4	-1	-6
<b>G</b>	0	-2	-4	-6	-3	0	3	8	3	-2	-7	-12	-5	-2	6	14	9	4	-1	2	-3
<b>C</b>	0	-2	-4	-1	-6	-5	3	3	6	1	-4	-9	-9	-7	1	9	17	12	7	2	0

Gap = -5

Match = +3

Mismatch = -2



# Alinhamento Semi-Global

$\times$	$\alpha$	A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A
$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	-2	-2	3	-2	-2	3	-2	-2	-2	-2	-2	3	-2	3	-2	3	-2	3	-2	-2
T	0	-2	1	-2	6	1	-2	1	1	1	-4	1	-2	1	-2	1	-2	1	-2	6	1
T	0	-2	1	-1	1	9	4	-1	4	4	-1	-1	-1	-4	-1	-4	-1	-4	-1	1	4
G	0	-2	-4	-1	-3	4	7	7	2	2	2	-3	-3	-3	-6	2	-3	-3	-6	-3	-1
G	0	-2	-4	-6	-3	-1	2	10	5	0	0	0	-5	-5	-5	-3	0	-5	-5	-8	-5
C	0	-2	-4	-1	-6	-5	2	5	8	3	-2	-2	3	-2	-2	-7	0	-2	-2	-7	-10
C	0	-2	-4	-1	-3	-8	-2	0	3	6	1	-4	1	1	1	-4	-4	-2	1	-4	-9
A	0	3	-2	-6	-3	-5	-7	-4	-2	1	9	4	-1	4	-1	-1	-6	-1	-4	-1	-1
A	0	3	1	-4	-8	-5	-7	-9	-6	-4	4	7	2	2	2	-3	-3	-3	-3	-6	2
T	0	-2	6	1	-1	-5	-7	-9	-6	-3	-1	7	5	0	0	0	-5	-5	-5	0	-3
C	0	-2	1	9	4	-1	-2	-7	-11	-8	-5	2	10	5	3	-2	3	-2	-2	-5	-2
C	0	-2	-4	4	7	2	2	-3	-8	-13	-10	-3	5	8	8	3	1	1	1	-4	-7
C	0	-2	-4	-1	2	5	5	0	-5	-10	-15	-8	0	3	11	6	6	1	4	-1	-6
G	0	-2	-4	-6	-3	0	3	8	3	-2	-7	-12	-5	-2	6	14	9	4	-1	2	-3
C	0	-2	-4	-1	-6	-5	3	3	6	1	-4	-9	-9	-7	1	9	17	12	7	2	0

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Semi-Global

		$\alpha$																				
		A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A	
$\times$	$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	C	0	-2	-2	3	-2	-2	3	-2	-2	-2	-2	3	-2	3	-2	3	-2	3	-2	3	-2
T	0	-2	1	-2	6	1	-2	1	1	1	-4	1	-2	1	-2	1	-2	1	-2	6	1	
T	0	-2	1	-1	1	9	4	-1	4	4	-1	-1	-1	-4	-1	-4	-1	-4	-1	1	4	
G	0	-2	-4	-1	-3	4	7	7	2	2	2	-3	-3	-3	-6	2	-3	-3	-6	-3	-1	
G	0	-2	-4	-6	-3	-1	2	10	5	0	0	0	-5	-5	-5	-3	0	-5	-5	-8	-5	
C	0	-2	-4	-1	-6	-5	2	5	8	3	-2	-2	3	-2	-2	-7	0	-2	-2	-7	-10	
C	0	-2	-4	-1	-3	-8	-2	0	3	6	1	-4	1	1	1	-4	-4	-2	1	-4	-9	
A	0	3	-2	-6	-3	-5	-7	-4	-2	1	9	4	-1	4	-1	-1	-6	-1	-4	-1	-1	
A	0	3	1	-4	-8	-5	-7	-9	-6	-4	4	7	2	2	2	-3	-3	-3	-3	-6	2	
T	0	-2	6	1	-1	-5	-7	-9	-6	-3	-1	7	5	0	0	0	-5	-5	-5	0	-3	
C	0	-2	1	9	4	-1	-2	-7	-11	-8	-5	2	10	5	3	-2	3	-2	-2	-5	-2	
C	0	-2	-4	4	7	2	2	-3	-8	-13	-10	-3	5	8	8	3	1	1	1	-4	-7	
C	0	-2	-4	-1	2	5	5	0	-5	-10	-15	-8	0	3	11	6	6	1	4	-1	-6	
G	0	-2	-4	-6	-3	0	3	8	3	-2	-7	-12	-5	-2	6	14	9	4	-1	2	-3	
C	0	-2	-4	-1	-6	-5	3	3	6	1	-4	-9	-9	-7	1	9	17	12	7	2	0	
																		C	A	C	T	A
																		C	-	-	-	-

# Alinhamento Semi-Global

× β	α	A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	-2	-2	3	-2	-2	3	-2	-2	-2	-2	-2	3	-2	3	-2	3	-2	3	-2	3	-2
T	0	-2	1	-2	6	1	-2	1	1	1	-4	1	-2	1	-2	1	-2	1	-2	6	1	
T	0	-2	1	-1	1	9	4	-1	4	4	-1	-1	-1	-4	-1	-4	-1	-4	-1	1	4	
G	0	-2	-4	-1	-3	4	7	7	2	2	2	-3	-3	-3	-6	2	-3	-3	-6	-3	-1	
G	0	-2	-4	-6	-3	-1	2	10	5	0	0	0	-5	-5	-5	-3	0	-5	-5	-8	-5	
C	0	-2	-4	-1	-6	-5	2	5	8	3	-2	-2	3	-2	-2	-7	0	-2	-2	-7	-10	
C	0	-2	-4	-1	-3	-8	-2	0	3	6	1	-4	1	1	1	-4	-4	-2	1	-4	-9	
A	0	3	-2	-6	-3	-5	-7	-4	-2	1	9	4	-1	4	-1	-1	-6	-1	-4	-1	-1	
A	0	3	1	-4	-8	-5	-7	-9	-6	-4	4	7	2	2	2	-3	-3	-3	-3	-6	2	
T	0	-2	6	1	-1	-5	-7	-9	-6	-3	-1	7	5	0	0	0	-5	-5	-5	0	-3	
C	0	-2	1	9	4	-1	-2	-7	-11	-8	-5	2	10	5	3	-2	3	-2	-2	-5	-2	
C	0	-2	-4	4	7	2	2	-3	-8	-13	-10	-3	5	8	8	3	1	1	1	-4	-7	
C	0	-2	-4	-1	2	5	5	0	-5	-10	-15	-8	0	3	11	6	6	1	4	-1	-6	
G	0	-2	-4	-6	-3	0	3	8	3	-2	-7	-12	-5	-2	6	14	9	4	-1	2	-3	
C	0	-2	-4	-1	-6	-5	3	3	6	1	-4	-9	-9	-7	1	9	17	12	7	2	0	
α													-	A	C	G	C	A	C	T	A	
β													C	C	C	G	C	-	-	-	-	

# Alinhamento Semi-Global

$\times$	$\alpha$	A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A	
$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	-2	-2	3	-2	-2	3	-2	-2	-2	-2	-2	3	-2	3	-2	3	-2	3	-2	3	-2
T	0	-2	1	-2	6	1	-2	1	1	1	-4	1	-2	1	-2	1	-2	1	-2	6	1	1
T	0	-2	1	-1	1	9	4	-1	4	4	-1	-1	-1	-4	-1	-4	-1	-4	-1	1	4	4
G	0	-2	-4	-1	-3	4	7	7	2	2	2	-3	-3	-3	-6	2	-3	-3	-6	-3	-1	-1
G	0	-2	-4	-6	-3	-1	2	10	5	0	0	0	-5	-5	-5	-3	0	-5	-5	-8	-5	-5
C	0	-2	-4	-1	-6	-5	2	5	8	3	-2	-2	3	-2	-2	-7	0	-2	-2	-7	-10	-10
C	0	-2	-4	-1	-3	-8	-2	0	3	6	1	-4	1	1	1	-4	-4	-2	1	-4	-9	-9
A	0	3	-2	-6	-3	-5	-7	-4	-2	1	9	4	-1	4	-1	-1	-6	-1	-4	-1	-1	-1
A	0	3	1	-4	-8	-5	-7	-9	-6	-4	4	7	2	2	2	-3	-3	-3	-3	-6	2	2
T	0	-2	6	1	-1	-5	-7	-9	-6	-3	-1	7	5	0	0	0	-5	-5	-5	0	-3	-3
C	0	-2	1	9	4	-1	-2	-7	-11	-8	-5	2	10	5	3	-2	3	-2	-2	-5	-2	-2
C	0	-2	-4	4	7	2	2	-3	-8	-13	-10	-3	5	8	8	3	1	1	1	-4	-7	-7
C	0	-2	-4	-1	2	5	5	0	-5	-10	-15	-8	0	3	11	6	6	1	4	-1	-6	-6
G	0	-2	-4	-6	-3	0	3	8	3	-2	-7	-12	-5	-2	6	14	9	4	-1	2	-3	-3
C	0	-2	-4	-1	-6	-5	3	3	6	1	-4	-9	-9	-7	1	9	17	12	7	2	0	0
$\alpha$			C	T	T	C	G	T	T	A	-	T	C	A	C	G	C	A	C	T	A	A
$\beta$			C	T	T	G	G	C	C	A	A	T	C	C	C	G	C	-	-	-	-	-

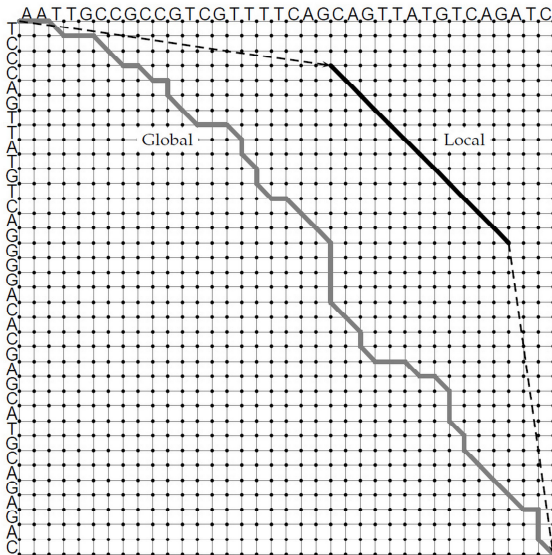
# Alinhamento Semi-Global

$\times$	$\alpha$	A	T	C	T	T	C	G	T	T	A	T	C	A	C	G	C	A	C	T	A	
$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	-2	-2	3	-2	-2	3	-2	-2	-2	-2	-2	3	-2	3	-2	3	-2	3	-2	-2	-2
T	0	-2	1	-2	6	1	-2	1	1	1	-4	1	-2	1	-2	1	-2	1	-2	6	1	1
T	0	-2	1	-1	1	9	4	-1	4	4	-1	-1	-1	-4	-1	-4	-1	-4	-1	1	4	4
G	0	-2	-4	-1	-3	4	7	7	2	2	2	-3	-3	-3	-6	2	-3	-3	-6	-3	-1	-1
G	0	-2	-4	-6	-3	-1	2	10	5	0	0	0	-5	-5	-5	-3	0	-5	-5	-8	-5	-5
C	0	-2	-4	-1	-6	-5	2	5	8	3	-2	-2	3	-2	-2	-7	0	-2	-2	-7	-10	-10
C	0	-2	-4	-1	-3	-8	-2	0	3	6	1	-4	1	1	1	-4	-4	-2	1	-4	-9	-9
A	0	3	-2	-6	-3	-5	-7	-4	-2	1	9	4	-1	4	-1	-1	-6	-1	-4	-1	-1	-1
A	0	3	1	-4	-8	-5	-7	-9	-6	-4	4	7	2	2	2	-3	-3	-3	-3	-6	2	2
T	0	-2	6	1	-1	-5	-7	-9	-6	-3	-1	7	5	0	0	0	-5	-5	-5	0	-3	-3
C	0	-2	1	9	4	-1	-2	-7	-11	-8	-5	2	10	5	3	-2	3	-2	-2	-5	-2	-2
C	0	-2	-4	4	7	2	2	-3	-8	-13	-10	-3	5	8	8	3	1	1	1	-4	-7	-7
C	0	-2	-4	-1	2	5	5	0	-5	-10	-15	-8	0	3	11	6	6	1	4	-1	-6	-6
G	0	-2	-4	-6	-3	0	3	8	3	-2	-7	-12	-5	-2	6	14	9	4	-1	2	-3	-3
C	0	-2	-4	-1	-6	-5	3	3	6	1	-4	-9	-9	-7	1	9	17	12	7	2	0	0
$\alpha$	A	T	C	T	T	C	G	T	T	A	-	T	C	A	C	G	C	A	C	T	A	A
$\beta$	-	-	C	T	T	G	G	C	C	A	A	T	C	C	C	G	C	-	-	-	-	-

## **Alinhamento Local**

---

# Alinhamento Global x Alinhamento Local



## Alinhamento Local - Abordagem Ingênuo

- O alinhamento local ótimo entre duas sequências  $\alpha$  e  $\beta$  é o melhor alinhamento (global) entre uma subsequência de  $\alpha$  e uma subsequência de  $\beta$ .
- É possível usar o algoritmo de alinhamento global de Needleman e Wunsch para produzir alinhamentos locais?
- Quantas subsequências (não nulas) as sequências  $\alpha$  e  $\beta$ , formadas respectivamente por  $m$  e  $n$  caracteres, possuem?
  - Respectivamente,  $m(m+1)/2 = \Theta(m^2)$  e  $n(n+1)/2 = \Theta(n^2)$  subsequências.
- Algoritmo ingênuo: testar todos os pares de subsequências e escolher aquele para com o maior valor de alinhamento global.
- Qual a complexidade do algoritmo ingênuo para alinhamento local?



# Análise de Complexidade

- Complexidade:

$$T(m, n) = \sum_{\alpha'=1}^m \sum_{\alpha''=\alpha'}^m \sum_{\beta'=1}^n \sum_{\beta''=\beta'}^n \Theta((\alpha'' - \alpha' + 1)(\beta'' - \beta' + 1))$$

- Limite superior:

$$T(m, n) \leq \frac{m(m+1)}{2} \frac{n(n+1)}{2} O(mn) = O(m^3 n^3)$$

- Limite inferior:

$$T(m, n) \geq \sum_{\alpha'=1}^{\lfloor \frac{m}{3} \rfloor} \sum_{\alpha''=m-\lfloor \frac{m}{3} \rfloor+1}^m \sum_{\beta'=1}^{\lfloor \frac{n}{3} \rfloor} \sum_{\beta''=n-\lfloor \frac{n}{3} \rfloor+1}^n \Omega\left(\left\lceil \frac{m}{3} \right\rceil \left\lceil \frac{n}{3} \right\rceil\right)$$

$$T(m, n) \geq \left\lfloor \frac{m}{3} \right\rfloor \left\lfloor \frac{m}{3} \right\rfloor \left\lfloor \frac{n}{3} \right\rfloor \left\lfloor \frac{n}{3} \right\rfloor \Omega\left(\left\lceil \frac{m}{3} \right\rceil \left\lceil \frac{n}{3} \right\rceil\right) = \Omega(m^3 n^3)$$

- Complexidade:
  - $T(m, n) = \Theta(m^3 n^3)$
- Exemplo: computador de 3GHz (1 instrução por ciclo)
  - $m = n = 10$ :
    - Tempo:  $10^6 / (3 \times 2^{30}) = 310$  bilionésimos de segundo
  - $m = n = 100$ :
    - Tempo:  $100^6 / (3 \times 2^{30}) = 310$  segundos
  - $m = n = 1000$ :
    - Tempo:  $1000^6 / (3 \times 2^{30}) = 10$  anos
  - $m = n = 10000$ :
    - Tempo:  $10000^6 / (3 \times 2^{30}) = 10$  bilhões de anos

# Alinhamento Local

- Como alterar o algoritmo de Needleman e Wunsch para produzir alinhamentos locais?
  - Redefinir o significado das células da matriz  $M$ .
    - A célula  $M[i,j]$  deve representar o valor do alinhamento ótimo entre todos os sufixos das sequências  $\alpha[1..j]$  e  $\beta[1..i]$ .
  - Qual o valor mínimo que uma célula pode armazenar?
    - O valor correspondente ao alinhamento de um sufixo vazio com qualquer outro, ou seja, zero.
  - Qual o valor do alinhamento local ótimo entre as duas sequências?
    - O maior valor armazenado em qualquer uma das posições da matriz.
  - Como recuperar o alinhamento local ótimo?
    - O alinhamento deve começar numa célula com valor zero e terminar na célula de maior valor da matriz (obter alinhamento na ordem inversa).
- Em 1981, Temple Smith e Michael Waterman propuseram o algoritmo de programação dinâmica para alinhamento local de sequências.

## Algoritmo 7: Local

**Input:**  $\alpha, m, \beta, n$

**for** *all*  $i \in [0..n]$  **do**

$M[i, 0] \leftarrow 0$

**end**

**for** *all*  $j \in [1..m]$  **do**

$M[0, j] \leftarrow 0$

**end**

**for** *all*  $i \in [1..n]$  **do**

**for** *all*  $j \in [1..m]$  **do**

$$M[i, j] \leftarrow \max \left\{ \begin{array}{l} 0, \\ M[i, j-1] + \text{gap}, \\ M[i-1, j-1] + \sigma(\alpha[j], \beta[i]), \\ M[i-1, j] + \text{gap} \end{array} \right\}$$

**end**

**end**

$\max \leftarrow 0$

**for** *all*  $i \in [1..n]$  **do**

**for** *all*  $j \in [1..m]$  **do**

**if**  $M[i, j] > \max$  **then**  $\max \leftarrow M[i, j]$ ;

**end**

**end**

**return**  $\max$

# Alinhamento Local

$\times$	$\alpha$	<b>G</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>
$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>C</b>	0																				
<b>T</b>	0																				
<b>T</b>	0																				
<b>G</b>	0																				
<b>G</b>	0																				
<b>C</b>	0																				
<b>C</b>	0																				
<b>A</b>	0																				
<b>C</b>	0																				
<b>T</b>	0																				
<b>C</b>	0																				
<b>C</b>	0																				
<b>C</b>	0																				
<b>G</b>	0																				
<b>C</b>	0																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Local

x β	α	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Local

x β	α	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Local

x	$\alpha$	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>C</b>	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
<b>T</b>	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
<b>T</b>	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
<b>G</b>	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
<b>G</b>	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
<b>C</b>	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
<b>C</b>	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
<b>A</b>	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
<b>C</b>	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
<b>T</b>	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
<b>C</b>	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
<b>C</b>	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
<b>C</b>	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
<b>G</b>	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
<b>C</b>	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1

C  
C

$\alpha$   
 $\beta$



# Alinhamento Local

x	α	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		β	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1

C   C  
G   C

# Alinhamento Local

x	$\alpha$	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		$\beta$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1

$\alpha$

$\beta$

T C C  
T G C

# Alinhamento Local

x	α	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		β	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1
α																					
β																					

C T C C  
C T G C

# Alinhamento Local

x	α	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A
		β	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	3	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1

α  
 β

A C T C C  
A C T G C

# Alinhamento Local

x	α	G	A	C	A	A	C	G	T	T	A	C	T	G	C	T	T	A	C	T	A	
		β	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	3	0	0	3	0	0	0	3	0	0	3	0	0	0	3	0	0	0
T	0	0	0	0	1	0	0	1	3	3	0	0	6	1	0	6	3	0	0	6	1	0
T	0	0	0	0	0	0	0	0	4	6	1	0	3	4	0	3	9	4	0	3	4	0
G	0	3	0	0	0	0	0	3	0	2	4	0	0	6	2	0	4	7	2	0	1	0
G	0	3	1	0	0	0	0	3	1	0	0	2	0	3	4	0	0	2	5	0	0	0
C	0	0	1	4	0	0	3	0	1	0	0	3	0	0	6	2	0	0	5	3	0	0
C	0	0	0	4	2	0	3	1	0	0	0	3	1	0	3	4	0	0	3	3	1	0
A	0	0	3	0	7	5	0	1	0	0	3	0	1	0	0	1	2	3	0	1	6	0
C	0	0	0	6	2	5	8	3	0	0	0	6	1	0	3	0	0	0	6	1	1	0
T	0	0	0	1	4	0	3	6	6	3	0	1	9	4	0	6	3	0	1	9	4	0
C	0	0	0	3	0	2	3	1	4	4	1	3	4	7	7	2	4	1	3	4	7	0
C	0	0	0	3	1	0	5	1	0	2	2	4	1	2	10	5	0	2	4	1	2	0
C	0	0	0	3	1	0	3	3	0	0	0	5	2	0	5	8	3	0	5	2	0	0
G	0	3	0	0	1	0	0	6	1	0	0	0	3	5	0	3	6	1	0	3	0	0
C	0	0	1	3	0	0	3	1	4	0	0	3	0	1	8	3	1	4	4	0	1	0

α  
 β

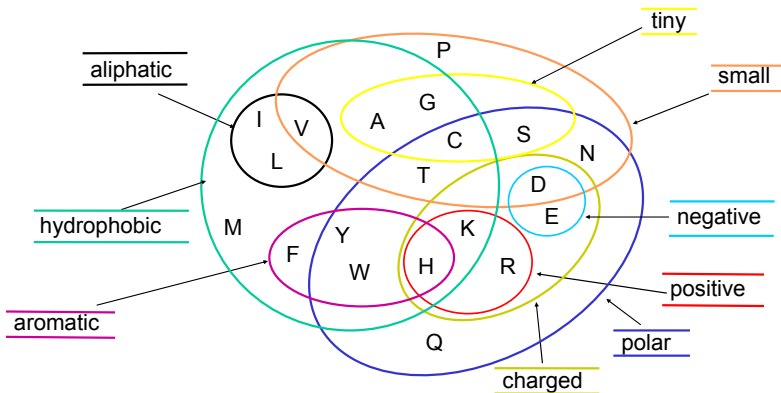
A C T C C  
A C T G C

# Pontuação para Alinhamento de Caracteres

---

- Simples (para nucleotídeos):
  - Match: alinhamento de dois caracteres iguais
  - Mismatch: alinhamento de dois caracteres diferentes
- Matrizes de Substituições (para aminoácidos):
  - Família PAM
  - Família BLOSUM

# Propriedades dos Aminoácidos





- Desenvolvida por Margaret Dayhoff (1978).
- PAM: *Percent (or Point) of Accepted Mutations*.
- Distância 1-PAM: uma mutação (substituição) a cada 100 aminoácidos em média.
- Duas proteínas com distância  $k$ -PAM não necessariamente possuem  $k\%$  de diferença entre seus aminoácidos.
- A similaridade esperada entre duas sequências protéicas com distância 200-PAM é de cerca de 25%.

- Seja  $F^k(a_i, a_j)$  a probabilidade de um aminoácido  $a_i$  mutar para o aminoácido  $a_j$  em sequências com distância  $k$ -PAM e seja  $freq(a_j)$  a frequência do aminoácido  $a_j$ .
- Fórmula geral para matrizes PAM:

$$PAM_k(a_i, a_j) = \left[ 10 \times \log \frac{F^k(a_i, a_j)}{freq(a_j)} \right]$$

- Valores mais utilizados:  $k = 40$ ,  $k = 120$  e  $k = 250$ .

- Desenvolvida por Steven Henikoff e Jorja Henikoff (1992).
- BLOSUM: *BLOcks of Amino Acid SUBstitution Matrix*
- BLOCKS: banco de dados de alinhamentos múltiplos de blocos conservados (sem gaps) de 504 grupos de proteínas.
- $BLOSUM_k$  considera apenas sequências com uma similaridade maior do que  $k\%$  entre suas bases.

- Seja  $freq_k(a_i, a_j)$  a frequência de alinhamentos de aminoácidos  $a_i$  com aminoácidos  $a_j$  e seja  $freq_k(a_i)$  a frequência do aminoácido  $a_i$ , em ambos os casos, em sequências com similaridade maior do que  $k\%$  entre suas bases (considerando o banco de dados BLOCKS).
- Fórmula geral para matrizes BLOSUM:

$$BLOSUM_k(a_i, a_j) = \left[ \log_2 \frac{freq_k(a_i, a_j)}{freq_k(a_i) \times freq_k(a_j)} \right]$$

- $BLOSUM_{62}$ : matriz padrão para alinhamento de proteínas (BLAST).

# BLOCKS

A	A	B	C	D	A	-	-	-	B	B	C	D	A
-	A	B	C	D	A	-	A	-	B	B	C	B	B
B	B	B	C	D	A	B	A	-	B	C	C	A	A
A	A	A	C	D	A	C	-	D	C	B	C	D	-
C	C	B	A	D	A	B	-	D	B	B	D	C	C
A	A	A	C	A	A	-	-	-	B	B	C	C	C

	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W	
C	9																				C
S	-1	4																			S
T	-1	1	5																		T
P	-3	-1	-1	7																	P
A	0	1	0	-1	4																A
G	-3	0	-2	-2	0	6															G
N	-3	1	0	-2	-2	0	6														N
D	-3	0	-1	-1	-2	-1	1	6													D
E	-4	0	-1	-1	-1	-2	0	2	5												E
Q	-3	0	-1	-1	-1	-2	0	0	2	5											Q
H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8										H
R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5									R
K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5								K
M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5							M
I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	-3	1	4						I
L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	-2	2	2	4					L
V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	-3	-2	1	3	1	4				V
F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6			F
Y	-2	-2	-2	-3	-2	-3	-2	-3	-2	-1	2	-2	-2	-1	-1	-1	-1	3	7		Y
W	-2	-3	-2	-4	-3	-2	-4	-4	-3	-2	-2	-3	-3	-1	-3	-2	-3	1	2	11	W

- A família PAM assume um modelo evolucionário baseado em árvores filogenéticas. A família BLOSUM não assume nenhum modelo evolucionário, mas considera blocos conservados de proteínas.
- A Entropia Relativa ( $H$ ) de uma matriz de substituição é dada por:

$$H = \sum_{i=1}^{20} \sum_{j=1}^i \text{freq}(a_i) \text{freq}(a_j) \sigma(a_i, a_j)$$

- Equivalência entre matrizes (com base em sua Entropia Relativa):

$$PAM_{100} \iff BLOSUM_{90} \mid PAM_{200} \iff BLOSUM_{52}$$

$$PAM_{120} \iff BLOSUM_{80} \mid PAM_{250} \iff BLOSUM_{45}$$

$$PAM_{160} \iff BLOSUM_{60} \mid PAM_{400} \iff BLOSUM_{30}$$

## Erro no Cálculo das Matrizes da Família BLOSUM

- Em 2008, Mark Styczynski e Kyle Jensen mostraram que as matrizes BLOSUM, comumente utilizadas na prática, foram calculadas de forma errada, devido a um erro de software.
- Este erro, que afeta cerca de 15% das posições das matrizes BLOSUM, passou despercebido por mais de 15 anos.
- Teste computacionais, realizados pela mesma dupla de pesquisadores, mostraram que as matrizes “erradas” apresentam melhores resultados na prática do que as matrizes “corretas”.



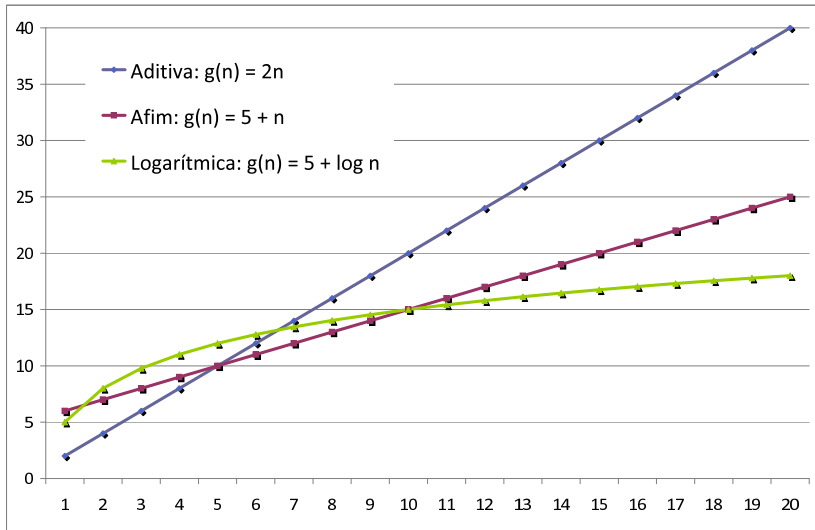
# **Penalidades para Blocos de Buracos**

---

## Penalidades para Blocos de Buracos

- Ao invés de penalizar buracos individualmente, podemos penalizar blocos de buracos consecutivos num alinhamento.
- Existem várias opções de funções de penalização de blocos de buracos.
- Funções Aditivas:  $g(k_1 + k_2) = g(k_1) + g(k_2)$ 
  - Linear:  $g(k) = a \times k$
- Funções Sub-Aditivas:  $g(k_1 + k_2) < g(k_1) + g(k_2)$ 
  - Afim:  $g(k) = a + b \times k$
  - Funções Convexas:  $g(k) - g(k - 1) > g(k + 1) - g(k)$ 
    - Logarítmica:  $g(k) = a + b \times \log k$

# Exemplos de Funções para Penalidades de Buracos



# Penalidades para Blocos de Buracos

- Exemplos de diferenças entre alinhamentos com mesma matriz de pontuação de alinhamento de bases, mas com penalidades diferentes para blocos de buracos.
- Alinhamento global ótimo com pontuação aditiva:

```
G C G C G T T A G A C T A G C A C C G  
G - G - G T T - G - C - A - C - C - G
```

# Penalidades para Blocos de Buracos

- Exemplos de diferenças entre alinhamentos com mesma matriz de pontuação de alinhamento de bases, mas com penalidades diferentes para blocos de buracos.
- Alinhamento global ótimo com pontuação sub-aditiva:

```
G C G C G T T A G A C T A G C A C C G
- - G G T T - - - - - G C A C C G
```

# Função Genérica de Penalidade para Buracos

- *Importante:* inserir um novo buraco em um alinhamento tem custo diferente caso o alinhamento corrente termine num buraco ou não.
- Para preencher a célula  $M[i, j]$  devemos considerar 3 opções:
  - Valor do alinhamento ótimo representado pela célula  $M[i - 1, j - 1]$  acrescido do alinhamento dos caracteres  $\alpha[j]$  e  $\beta[i]$ .
  - Valor do alinhamento ótimo representado por uma célula  $M[i, j - k]$ , acrescido de um alinhamento entre  $\alpha[j - k + 1..j]$  e um bloco formado por  $k$  buracos (para  $1 \leq k \leq j$ ).
  - Valor do alinhamento ótimo representado por uma célula  $M[i - k, j]$ , acrescido de um alinhamento entre  $\beta[i - k + 1..i]$  e um bloco formado por  $k$  buracos (para  $1 \leq k \leq i$ ).
- Algoritmo proposto por Michael Waterman, Temple Smith e William Beyer em 1976.

## Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6						*					
C	7											
A	8											
C	9											

$$M[6,6] + \sigma(G,C)$$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7						*					
A	8											
C	9											

$k = 1$

$M[7,6] + g(1)$



# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7					*						
A	8											
C	9											

$k = 2$

$M[7,5] + g(2)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7				*							
A	8											
C	9											

$k = 3$

$M[7,4] + g(3)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7			*								
A	8											
C	9											

$k = 4$

$M[7,3] + g(4)$

## Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7		*									
A	8											
C	9											

$k = 5$

$M[7,2] + g(5)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7	*										
A	8											
C	9											

$k = 6$

$M[7,1] + g(6)$

## Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	*											
A	8											
C	9											

$k = 7$

$M[7,0] + g(7)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7											
A	8											
C	9											

$k = 1$

$M[6,7] + g(1)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5							*				
C	6											
C	7											
A	8											
C	9											

$k = 2$

$M[5,7] + g(2)$



## Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3											
G	4							*				
G	5											
C	6											
C	7											
A	8											
C	9											

$k = 3$

$M[4,7] + g(3)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2											
T	3							*				
G	4											
G	5											
C	6											
C	7											
A	8											
C	9											

$k = 4$

$M[3,7] + g(4)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1											
T	2							*				
T	3											
G	4											
G	5											
C	6											
C	7											
A	8											
C	9											

$k = 5$

$M[2,7] + g(5)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	7	8	9	10	11
A	1							*				
T	2											
T	3											
G	4											
G	5											
C	6											
C	7											
A	8											
C	9											

$k = 6$

$M[1,7] + g(6)$

## Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	C	C	T	G	T	G	G	C	A	A	C
$\beta$	0	1	2	3	4	5	6	*	8	9	10	11
A	1											
T	2											
T	3											
G	4											
G	5											
C	6											
C	7											
A	8											
C	9											

$k = 7$

$M[0,7] + g(7)$

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24
G	-6	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
G	-7	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
G	-8	-4	-5	4	-2	-3	-9													
T	-9																			
T	-10																			
G	-11																			
C	-12																			
A	-13																			
C	-14																			
C	-15																			
G	-16																			

$\alpha$  = G C G C G T  
 $\beta$  = G G G - - -  
 -9 = 3 -2 3 -6 -1 -6

Open Gap = -5

Match = +3

Extend Gap = -1

Mismatch = -2

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24
G	-6	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
G	-7	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
G	-8	-4	-5	4	-2	-3	-9													
T	-9																			
T	-10																			
G	-11																			
C	-12																			
A	-13																			
C	-14																			
C	-15																			
G	-16																			

$\alpha$  = G C G C | G T

Open Gap = -5

Match = +3

$\beta$  = G G G - - -

Extend Gap = -1

Mismatch = -2

-9 = 3 -2 3 -6 | -6 -1

# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24
G	-6	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
G	-7	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
G	-8	-4	-5	4	-2	-3	-4													
T	-9																			
T	-10																			
G	-11																			
C	-12																			
A	-13																			
C	-14																			
C	-15																			
G	-16																			

$\alpha$  = G C G | C G T

Open Gap = -5

Match = +3

$\beta$  = G G G | - - -

Extend Gap = -1

Mismatch = -2

-4 = 3 -2 3 | -6 -1 -1



# Função Genérica de Penalidade para Buracos

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24
<b>G</b>	-6	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
<b>G</b>	-7	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
<b>G</b>	-8	-4	-5	4	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17
<b>T</b>	-9	-5	-6	-2	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
<b>T</b>	-10	-6	-7	-3	-4	0	-1	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14
<b>G</b>	-11	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
<b>C</b>	-12	-8	-4	-5	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-8	-9	-10	-11
<b>A</b>	-13	-9	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-7	-8	1	-5	-6	-7
<b>C</b>	-14	-10	-6	-7	-3	-9	-5	-6	-7	-3	-4	5	-1	-2	-3	-4	-5	4	-2	-3
<b>C</b>	-15	-11	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	1
<b>G</b>	-16	-12	-13	-4	-10	-1	-7	-8	-9	-5	-6	-2	-3	1	0	-6	-2	-3	1	10

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Função Genérica de Penalidade para Buracos

×	α	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
		β	0	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22
G	-6	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
G	-7	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
G	-8	-4	-5	4	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17
T	-9	-5	-6	-2	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
T	-10	-6	-7	-3	-4	0	-1	3	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14
G	-11	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
C	-12	-8	-4	-5	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-8	-9	-10	-11
A	-13	-9	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-7	-8	1	-5	-6	-7
C	-14	-10	-6	-7	-3	-9	-5	-6	-7	-3	-4	5	-1	-2	-3	-4	-5	4	-2	-3
C	-15	-11	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	1
G	-16	-12	-13	-4	-10	-1	-7	-8	-9	-5	-6	-2	-3	1	0	-6	-2	-3	1	10
α	=	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
β	=	-	-	G	G	G	T	T	-	-	-	-	-	-	G	C	A	C	C	G
10	=	-6	-1	3	-2	3	3	3	-6	-1	-1	-1	-1	-1	3	3	3	3	3	3

# Função Genérica de Penalidade para Buracos

---

## Algoritmo 8: Global

---

**Input:**  $\alpha, m, \beta, n$

$M[0, 0] \leftarrow 0$

**for** all  $i \in [1..n]$  **do**  $M[i, 0] \leftarrow g(i)$  ;

**for** all  $j \in [1..m]$  **do**  $M[0, j] \leftarrow g(j)$  ;

**for** all  $i \in [1..n]$  **do**

**for** all  $j \in [1..m]$  **do**

$\max \leftarrow M[i-1, j-1] + \sigma(\alpha[j], \beta[i])$

**for** all  $k \in [1..j]$  **do**

**if**  $\max < M[i, j-k] + g(k)$  **then**  $\max \leftarrow M[i, j-k] + g(k)$  ;

**end**

**for** all  $k \in [1..i]$  **do**

**if**  $\max < M[i-k, j] + g(k)$  **then**  $\max \leftarrow M[i-k, j] + g(k)$  ;

**end**

$M[i, j] \leftarrow \max$

**end**

**end**

**return**  $M[n, m]$

---

# Análise de Complexidade

- Complexidade:

$$T(m, n) = \sum_{i=1}^m \sum_{j=1}^n \Theta(i + j + 1)$$

- Limite superior:

$$T(m, n) \leq mnO(m + n + 1) = O(m^2n + mn^2)$$

- Limite inferior:

$$T(m, n) \geq \sum_{i=\lfloor \frac{m}{2} \rfloor + 1}^m \sum_{j=\lfloor \frac{n}{2} \rfloor + 1}^n \Omega(i + j + 1)$$

$$T(m, n) \geq \lfloor \frac{m}{2} \rfloor \lfloor \frac{n}{2} \rfloor \Omega(\lfloor \frac{m}{2} \rfloor + \lfloor \frac{n}{2} \rfloor) = \Omega(m^2n + mn^2)$$

# Complexidade e Tempo de Execução

- Complexidade:
  - $T(m, n) = \Theta(mn(m + n)) = \Theta(m^2n + mn^2)$
- Exemplo: computador de 3GHz (1 instrução por ciclo)
  - $m = n = 10$ :
    - Tempo:  $(10 \times 10 \times 20)/(3 \times 2^{30}) = 31$  bilionésimos de segundo
  - $m = n = 100$ :
    - Tempo:  $(100 \times 100 \times 200)/(3 \times 2^{30}) = 621$  milionésimos de segundo
  - $m = n = 1000$ :
    - Tempo:  $(1000 \times 1000) \times 2000/(3 \times 2^{30}) = 621$  milésimos de segundo
  - $m = n = 10000$ :
    - Tempo:  $(10000 \times 10000) \times 20000/(3 \times 2^{30}) = 10$  minutos
  - $m = n = 100000$ :
    - Tempo:  $(100000 \times 100000 \times 200000)/(3 \times 2^{30}) = 7$  dias
  - $m = n = 1000000$ :
    - Tempo:  $(1000000 \times 1000000 \times 2000000)/(3 \times 2^{30}) = 19$  anos

## Função Afim de Penalidade para Buracos

- Para obter um algoritmo de melhor complexidade, e ainda assim penalizar um bloco de buracos de forma mais adequada, podemos considerar funções sub-aditivas afins, ao invés de funções genéricas.
- Construção de três matrizes:
  - $M_{\searrow}[i, j]$ : alinhamento ótimo que termina com o alinhamento dos caracteres  $\alpha[j]$  e  $\beta[i]$
  - $M_{\leftarrow}[i, j]$ : alinhamento ótimo que termina com o alinhamento do caracter  $\alpha[j]$  com um buraco
  - $M_{\uparrow}[i, j]$ : alinhamento ótimo que termina com o alinhamento de um buraco com o caracter  $\beta[i]$
- Valor do alinhamento global máximo:
  - $\max\{M_{\searrow}[n, m], M_{\leftarrow}[n, m], M_{\uparrow}[n, m]\}$
- Algoritmo proposto por Osamu Gotoh em 1982.

---

## Algoritmo 9: Initialize

---

**Input:**  $m, n$

$M_{\swarrow}[0, 0] \leftarrow 0$

$M_{\leftarrow}[0, 0] \leftarrow -\infty$

$M_{\uparrow}[0, 0] \leftarrow -\infty$

**for** all  $i \in [1..n]$  **do**

$M_{\swarrow}[i, 0] \leftarrow -\infty$

$M_{\leftarrow}[i, 0] \leftarrow -\infty$

$M_{\uparrow}[i, 0] \leftarrow -(a + b \times i)$

**end**

**for** all  $j \in [1..m]$  **do**

$M_{\swarrow}[0, j] \leftarrow -\infty$

$M_{\leftarrow}[0, j] \leftarrow -(a + b \times j)$

$M_{\uparrow}[0, j] \leftarrow -\infty$

**end**

---

---

## Algoritmo 10: GlobalAfim

---

**Input:**  $\alpha, m, \beta, n$

Initialize(m,n)

**for** all  $i \in [1..n]$  **do**

**for** all  $j \in [1..m]$  **do**

$$M_{\swarrow}[i, j] \leftarrow \sigma(\alpha[j], \beta[i]) + \max \left\{ \begin{array}{l} M_{\swarrow}[i-1, j-1], \\ M_{\leftarrow}[i-1, j-1], \\ M_{\uparrow}[i-1, j-1] \end{array} \right\}$$

$$M_{\leftarrow}[i, j] \leftarrow \max \left\{ \begin{array}{l} M_{\swarrow}[i, j-1] - (a+b), \\ M_{\leftarrow}[i, j-1] - b, \\ M_{\uparrow}[i, j-1] - (a+b) \end{array} \right\}$$

$$M_{\uparrow}[i, j] \leftarrow \max \left\{ \begin{array}{l} M_{\swarrow}[i-1, j] - (a+b), \\ M_{\leftarrow}[i-1, j] - (a+b), \\ M_{\uparrow}[i-1, j] - b \end{array} \right\}$$

**end**

**end**

**return**  $\max\{M_{\swarrow}[n, m], M_{\leftarrow}[n, m], M_{\uparrow}[n, m]\}$

---



# Função Afim para Penalidade de Buracos: Matriz $M_{\alpha, \beta}$

$\times$		$\alpha$																			
		<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	<b>0</b>	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	
	<b>G</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>G</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>G</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>T</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>T</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>G</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>C</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>A</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>C</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>C</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	<b>G</b>	$-\infty$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Função Afim para Penalidade de Buracos: Matriz $M_{\leftarrow}$

$\times$		$\alpha$																			
		G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G	
$\beta$	$-\infty$	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	
	G	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	G	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	G	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	T	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	T	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	G	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	C	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	A	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	C	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
	C	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	
G	$-\infty$	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←	←		

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Função Afim para Penalidade de Buracos: Matriz $M_{\uparrow}$

$\times$		$\alpha$																		
		G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	G	-6	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	G	-7	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	G	-8	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	T	-9	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	T	-10	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	G	-11	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	C	-12	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	A	-13	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	C	-14	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	C	-15	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	G	-16	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Função Afim para Penalidade de Buracos: Matriz $M_{\alpha, \beta}$

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
<b>G</b>	$-\infty$	3	-8	-4	-10	-6	-12	-13	-14	-10	-16	-17	-18	-19	-15	-21	-22	-23	-24	-20
<b>G</b>	$-\infty$	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
<b>G</b>	$-\infty$	-4	-5	4	-2	-3	-4	-10	-11	-7	-8	-14	-15	-16	-12	-13	-19	-20	-21	-17
<b>T</b>	$-\infty$	-10	-6	-7	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
<b>T</b>	$-\infty$	-11	-7	-8	-4	0	-1	3	-3	-9	-10	-11	-7	-8	-14	-15	-16	-17	-18	-19
<b>G</b>	$-\infty$	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
<b>C</b>	$-\infty$	-13	-4	-10	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-13	-9	-10	-16
<b>A</b>	$-\infty$	-14	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-12	-13	1	-10	-11	-12
<b>C</b>	$-\infty$	-15	-6	-12	-3	-9	-5	-11	-7	-3	-9	5	-6	-7	-8	-4	-10	4	-2	-8
<b>C</b>	$-\infty$	-16	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	-4
<b>G</b>	$-\infty$	-12	-13	-4	-10	-1	-7	-13	-9	-5	-11	-7	-3	1	0	-6	-2	-8	-4	10

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Função Afim para Penalidade de Buracos: Matriz $M_{\leftarrow}$

$\times$		$\alpha$																			
		<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	$-\infty$	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	
	<b>G</b>	$-\infty$	-12	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
	<b>G</b>	$-\infty$	-13	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21
	<b>G</b>	$-\infty$	-14	-10	-11	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17
	<b>T</b>	$-\infty$	-15	-11	-12	-8	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18
	<b>T</b>	$-\infty$	-16	-12	-13	-9	-10	-6	-7	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14
	<b>G</b>	$-\infty$	-17	-13	-14	-10	-11	-7	-8	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15
	<b>C</b>	$-\infty$	-18	-14	-10	-11	-7	-8	-9	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-10	-11
	<b>A</b>	$-\infty$	-19	-15	-16	-12	-13	-9	-10	-11	-7	-8	-4	-5	-6	-7	-8	-9	-5	-6	-7
	<b>C</b>	$-\infty$	-20	-16	-12	-13	-9	-10	-11	-12	-13	-9	-10	-1	-2	-3	-4	-5	-6	-2	-3
	<b>C</b>	$-\infty$	-21	-17	-13	-14	-10	-11	-12	-13	-14	-15	-11	-7	-3	-4	-5	-6	-7	-8	1
<b>G</b>	$-\infty$	-22	-18	-19	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-5	-6	-7	-8	-9	-5	

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Função Afim para Penalidade de Buracos: Matriz $M_{\uparrow}$

$\times$		$\alpha$																			
		<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	
	<b>G</b>	-6	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	-25	-26	-27	-28	-29	-30
	<b>G</b>	-7	-3	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	-25	-26
	<b>G</b>	-8	-4	-5	-6	-12	-8	-14	-15	-16	-12	-18	-19	-20	-21	-17	-23	-24	-25	-26	-22
	<b>T</b>	-9	-5	-6	-2	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23
	<b>T</b>	-10	-6	-7	-3	-4	-10	-6	-7	-13	-14	-15	-16	-12	-18	-19	-20	-21	-22	-23	-24
	<b>G</b>	-11	-7	-8	-4	-5	-6	-7	-3	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
	<b>C</b>	-12	-8	-9	-5	-6	-7	-8	-4	-5	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
	<b>A</b>	-13	-9	-10	-6	-7	-8	-9	-5	-6	-7	-8	-9	-15	-16	-12	-8	-14	-15	-16	-17
	<b>C</b>	-14	-10	-11	-7	-8	-9	-10	-6	-7	-8	-4	-10	-11	-12	-13	-9	-5	-11	-12	-13
	<b>C</b>	-15	-11	-12	-8	-9	-10	-11	-7	-8	-9	-5	-1	-7	-8	-9	-10	-6	-2	-8	-9
<b>G</b>	-16	-12	-13	-9	-10	-11	-12	-8	-9	-10	-6	-2	-3	-9	-10	-6	-7	-3	1	-5	

Open Gap = -5

Extend Gap = -1

Match = +3

Mismatch = -2

# Alinhamento Upmost (Matriz $M_{\searrow}$ )

×		α																			
		G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G	
β	0	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	
	G	-∞	3	-8	-4	-10	-6	-12	-13	-14	-10	-16	-17	-18	-19	-15	-21	-22	-23	-24	-20
	G	-∞	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
	G	-∞	-4	-5	4	-2	-3	-4	-10	-11	-7	-8	-14	-15	-16	-12	-13	-19	-20	-21	-17
	T	-∞	-10	-6	-7	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
	T	-∞	-11	-7	-8	-4	0	-1	3	-3	-9	-10	-11	-7	-8	-14	-15	-16	-17	-18	-19
	G	-∞	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
	C	-∞	-13	-4	-10	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-13	-9	-10	-16
	A	-∞	-14	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-12	-13	1	-10	-11	-12
	C	-∞	-15	-6	-12	-3	-9	-5	-11	-7	-3	-9	5	-6	-7	-8	-4	-10	4	-2	-8
	C	-∞	-16	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	-4
G	-∞	-12	-13	-4	-10	-1	-7	-13	-9	-5	-11	-7	-3	1	0	-6	-2	-8	-4	10	
α	=												T	A	G	C	A	C	C	G	
β	=												T	T	G	C	A	C	C	G	
	=												3	-2	3	3	3	3	3	3	

# Alinhamento Upmost (Matriz $M_{\leftarrow}$ )

$\times$		$\alpha$																			
		G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G	
$\beta$	$-\infty$	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	
	G	$-\infty$	-12	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
	G	$-\infty$	-13	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21
	G	$-\infty$	-14	-10	-11	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17
	T	$-\infty$	-15	-11	-12	-8	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18
	T	$-\infty$	-16	-12	-13	-9	-10	-6	-7	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14
	G	$-\infty$	-17	-13	-14	-10	-11	-7	-8	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15
	C	$-\infty$	-18	-14	-10	-11	-7	-8	-9	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-10	-11
	A	$-\infty$	-19	-15	-16	-12	-13	-9	-10	-11	-7	-8	-4	-5	-6	-7	-8	-9	-5	-6	-7
	C	$-\infty$	-20	-16	-12	-13	-9	-10	-11	-12	-13	-9	-10	-1	-2	-3	-4	-5	-6	-2	-3
	C	$-\infty$	-21	-17	-13	-14	-10	-11	-12	-13	-14	-15	-11	-7	-3	-4	-5	-6	-7	-8	1
	G	$-\infty$	-22	-18	-19	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-5	-6	-7	-8	-9	-5
$\alpha$	=				C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G	
$\beta$	=				-	-	-	-	-	-	-	-	T	T	G	C	A	C	C	G	
=					-6	-1	-1	-1	-1	-1	-1	-1	3	-2	3	3	3	3	3	3	



# Alinhamento Upmost (Matriz $M_{\searrow}$ )

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
G	$-\infty$	3	-8	-4	-10	-6	-12	-13	-14	-10	-16	-17	-18	-19	-15	-21	-22	-23	-24	-20
G	$-\infty$	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
G	$-\infty$	-4	-5	4	-2	-3	-4	-10	-11	-7	-8	-14	-15	-16	-12	-13	-19	-20	-21	-17
T	$-\infty$	-10	-6	-7	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
T	$-\infty$	-11	-7	-8	-4	0	-1	3	-3	-9	-10	-11	-7	-8	-14	-15	-16	-17	-18	-19
G	$-\infty$	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
C	$-\infty$	-13	-4	-10	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-13	-9	-10	-16
A	$-\infty$	-14	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-12	-13	1	-10	-11	-12
C	$-\infty$	-15	-6	-12	-3	-9	-5	-11	-7	-3	-9	5	-6	-7	-8	-4	-10	4	-2	-8
C	$-\infty$	-16	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	-4
G	$-\infty$	-12	-13	-4	-10	-1	-7	-13	-9	-5	-11	-7	-3	1	0	-6	-2	-8	-4	10
$\alpha$	=	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	=	G	G	G	-	-	-	-	-	-	-	-	T	T	G	C	A	C	C	G
10	=	3	-2	3	-6	-1	-1	-1	-1	-1	-1	-1	3	-2	3	3	3	3	3	3

# Alinhamento Downmost (Matriz $M_{\leftarrow}$ )

×		α																			
		G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G	
β	0	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	-∞	
	G	-∞	3	-8	-4	-10	-6	-12	-13	-14	-10	-16	-17	-18	-19	-15	-21	-22	-23	-24	-20
	G	-∞	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
	G	-∞	-4	-5	4	-2	-3	-4	-10	-11	-7	-8	-14	-15	-16	-12	-13	-19	-20	-21	-17
	T	-∞	-10	-6	-7	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
	T	-∞	-11	-7	-8	-4	0	-1	3	-3	-9	-10	-11	-7	-8	-14	-15	-16	-17	-18	-19
	G	-∞	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
	C	-∞	-13	-4	-10	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-13	-9	-10	-16
	A	-∞	-14	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-12	-13	1	-10	-11	-12
	C	-∞	-15	-6	-12	-3	-9	-5	-11	-7	-3	-9	5	-6	-7	-8	-4	-10	4	-2	-8
	C	-∞	-16	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	-4
G	-∞	-12	-13	-4	-10	-1	-7	-13	-9	-5	-11	-7	-3	1	0	-6	-2	-8	-4	10	
α	=														G	C	A	C	C	G	
β	=														G	C	A	C	C	G	
	=														3	3	3	3	3	3	

# Alinhamento Downmost (Matriz $M_{\leftarrow}$ )

$\times$		$\alpha$																			
		<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	$-\infty$	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	
	<b>G</b>	$-\infty$	-12	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
	<b>G</b>	$-\infty$	-13	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21
	<b>G</b>	$-\infty$	-14	-10	-11	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17
	<b>T</b>	$-\infty$	-15	-11	-12	-8	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18
	<b>T</b>	$-\infty$	-16	-12	-13	-9	-10	-6	-7	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14
	<b>G</b>	$-\infty$	-17	-13	-14	-10	-11	-7	-8	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15
	<b>C</b>	$-\infty$	-18	-14	-10	-11	-7	-8	-9	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-10	-11
	<b>A</b>	$-\infty$	-19	-15	-16	-12	-13	-9	-10	-11	-7	-8	-4	-5	-6	-7	-8	-9	-5	-6	-7
	<b>C</b>	$-\infty$	-20	-16	-12	-13	-9	-10	-11	-12	-13	-9	-10	-1	-2	-3	-4	-5	-6	-2	-3
	<b>C</b>	$-\infty$	-21	-17	-13	-14	-10	-11	-12	-13	-14	-15	-11	-7	-3	-4	-5	-6	-7	-8	1
	<b>G</b>	$-\infty$	-22	-18	-19	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-5	-6	-7	-8	-9	-5
$\alpha$	=								<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	=								-	-	-	-	-	-	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
	=								-6	-1	-1	-1	-1	-1	-1	3	3	3	3	3	3

# Alinhamento Downmost (Matriz $M_{\leftarrow}$ )

$\times$		$\alpha$																			
		<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	
	<b>G</b>	$-\infty$	3	-8	-4	-10	-6	-12	-13	-14	-10	-16	-17	-18	-19	-15	-21	-22	-23	-24	-20
	<b>G</b>	$-\infty$	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
	<b>G</b>	$-\infty$	-4	-5	4	-2	-3	-4	-10	-11	-7	-8	-14	-15	-16	-12	-13	-19	-20	-21	-17
	<b>T</b>	$-\infty$	-10	-6	-7	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
	<b>T</b>	$-\infty$	-11	-7	-8	-4	0	-1	3	-3	-9	-10	-11	-7	-8	-14	-15	-16	-17	-18	-19
	<b>G</b>	$-\infty$	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
	<b>C</b>	$-\infty$	-13	-4	-10	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-13	-9	-10	-16
	<b>A</b>	$-\infty$	-14	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-12	-13	1	-10	-11	-12
	<b>C</b>	$-\infty$	-15	-6	-12	-3	-9	-5	-11	-7	-3	-9	5	-6	-7	-8	-4	-10	4	-2	-8
	<b>C</b>	$-\infty$	-16	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	-4
<b>G</b>	$-\infty$	-12	-13	-4	-10	-1	-7	-13	-9	-5	-11	-7	-3	1	0	-6	-2	-8	-4	10	
$\alpha$	=						<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	=						<b>T</b>	<b>T</b>	-	-	-	-	-	-	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
=							3	3	-6	-1	-1	-1	-1	-1	3	3	3	3	3	3	

# Alinhamento Downmost (Matriz $M_{\leftarrow}$ )

$\times$		$\alpha$																			
		<b>G</b>	<b>C</b>	<b>G</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
$\beta$	$-\infty$	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21	-22	-23	-24	
	<b>G</b>	$-\infty$	-12	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20
	<b>G</b>	$-\infty$	-13	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19	-20	-21
	<b>G</b>	$-\infty$	-14	-10	-11	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17
	<b>T</b>	$-\infty$	-15	-11	-12	-8	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15	-16	-17	-18
	<b>T</b>	$-\infty$	-16	-12	-13	-9	-10	-6	-7	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14
	<b>G</b>	$-\infty$	-17	-13	-14	-10	-11	-7	-8	-9	-5	-6	-7	-8	-9	-10	-11	-12	-13	-14	-15
	<b>C</b>	$-\infty$	-18	-14	-10	-11	-7	-8	-9	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-10	-11
	<b>A</b>	$-\infty$	-19	-15	-16	-12	-13	-9	-10	-11	-7	-8	-4	-5	-6	-7	-8	-9	-5	-6	-7
	<b>C</b>	$-\infty$	-20	-16	-12	-13	-9	-10	-11	-12	-13	-9	-10	-1	-2	-3	-4	-5	-6	-2	-3
	<b>C</b>	$-\infty$	-21	-17	-13	-14	-10	-11	-12	-13	-14	-15	-11	-7	-3	-4	-5	-6	-7	-8	1
	<b>G</b>	$-\infty$	-22	-18	-19	-10	-11	-7	-8	-9	-10	-11	-12	-8	-9	-5	-6	-7	-8	-9	-5
	$\alpha$	=				<b>C</b>	<b>G</b>	<b>T</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>
$\beta$	=				-	-	<b>T</b>	<b>T</b>	-	-	-	-	-	-	<b>G</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	
	=				-6	-1	3	3	-6	-1	-1	-1	-1	-1	3	3	3	3	3	3	

# Alinhamento Downmost (Matriz $M_{\leftarrow}$ )

$\times$	$\alpha$	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
G	$-\infty$	3	-8	-4	-10	-6	-12	-13	-14	-10	-16	-17	-18	-19	-15	-21	-22	-23	-24	-20
G	$-\infty$	-3	1	0	-6	-2	-8	-9	-10	-6	-12	-13	-14	-15	-11	-17	-18	-19	-20	-16
G	$-\infty$	-4	-5	4	-2	-3	-4	-10	-11	-7	-8	-14	-15	-16	-12	-13	-19	-20	-21	-17
T	$-\infty$	-10	-6	-7	2	-4	0	-1	-7	-8	-9	-10	-6	-12	-13	-14	-15	-16	-17	-18
T	$-\infty$	-11	-7	-8	-4	0	-1	3	-3	-9	-10	-11	-7	-8	-14	-15	-16	-17	-18	-19
G	$-\infty$	-7	-8	-4	-5	-1	-2	-3	1	0	-6	-7	-8	-9	-5	-11	-12	-13	-14	-10
C	$-\infty$	-13	-4	-10	-1	-7	-3	-4	-5	-1	-2	-3	-9	-10	-11	-2	-13	-9	-10	-16
A	$-\infty$	-14	-10	-6	-7	-3	-9	-5	-1	-7	2	-4	-5	-6	-12	-13	1	-10	-11	-12
C	$-\infty$	-15	-6	-12	-3	-9	-5	-11	-7	-3	-9	5	-6	-7	-8	-4	-10	4	-2	-8
C	$-\infty$	-16	-7	-8	-4	-5	-11	-7	-8	-9	-5	-1	3	-3	-4	0	-6	-2	7	-4
G	$-\infty$	-12	-13	-4	-10	-1	-7	-13	-9	-5	-11	-7	-3	1	0	-6	-2	-8	-4	10
$\alpha$	=	G	C	G	C	G	T	T	A	G	A	C	T	A	G	C	A	C	C	G
$\beta$	=	G	G	G	-	-	T	T	-	-	-	-	-	-	G	C	A	C	C	G
10	=	3	-2	3	-6	-1	3	3	-6	-1	-1	-1	-1	-1	3	3	3	3	3	3

## Exercício

*Calcule a complexidade assintótica (usando a notação  $\Theta$ ) do algoritmo de alinhamento global com função afim para penalidade de blocos de buracos proposto por Osamu Gotoh.*

## Exercício

*Escreva um algoritmo, em pseudocódigo, que dadas duas sequências  $\alpha$  e  $\beta$ , com respectivamente  $m$  e  $n$  caracteres, e três matrizes de pontuação de alinhamentos entre  $\alpha$  e  $\beta$  usando função afim para penalidade de blocos de buracos ( $M_{\leftarrow}$ ,  $M_{\searrow}$  e  $M_{\uparrow}$ ), retorne um alinhamento ótimo entre  $\alpha$  e  $\beta$ .*

# **Alinhamento Global usando Espaço Linear**

---



## Alinhamento Global usando Espaço Linear

- Todos os algoritmos de alinhamento que vimos até agora usam espaço quadrático em relação ao tamanho das sequências ( $\Theta(mn)$ ).
- Este requisito de memória pode tornar estes algoritmos proibitivos.
- Por exemplo, ao tentar alinhar duas sequências com um milhão de caracteres seria necessário aproximadamente 4TB de memória, considerando que um inteiro ocupe 4 bytes (32 bits) de memória.
- Note que é possível obter o valor do alinhamento ótimo usando apenas duas linhas da matriz de programação dinâmica.
- Como obter o valor ótimo e o alinhamento usando apenas espaço linear em relação ao tamanho das sequências de entrada?

# Alinhamento Global usando Espaço Linear

- Usar o paradigma de Divisão e Conquista em cima da matriz de Programação Dinâmica:
  - Calcular o vetor  $V_1$  com os valores dos alinhamentos ótimos entre todos os prefixos da sequência  $\alpha$  com a sequência  $\beta[1..\lfloor \frac{n}{2} \rfloor]$ .
  - Calcular o vetor  $V_2$  com os valores dos alinhamentos ótimos entre todos os prefixos da sequência  $\bar{\alpha}$  com a sequência  $\beta[\lfloor \frac{n}{2} \rfloor + 1..n]$ .
    - *Observação:*  $\bar{\alpha}$  significa a sequência reversa de  $\alpha$ .
  - Encontrar  $k$  tal que  $V_1[k] + V_2[m - k]$  seja o máximo possível.
  - O valor  $k$  indica que o alinhamento ótimo entre as sequências  $\alpha$  e  $\beta$  passa pela posição  $M[\lfloor \frac{n}{2} \rfloor, k]$ .
  - Resolver recursivamente o alinhamento entre  $\alpha[1..k]$  e  $\beta[1..\lfloor \frac{n}{2} \rfloor]$ , assim como o alinhamento entre  $\bar{\alpha}[k + 1..m]$  e  $\beta[\lfloor \frac{n}{2} \rfloor + 1..n]$ .
- Nos slides seguintes consideraremos que além dos vetores  $V_1$  e  $V_2$ , que representam as últimas linhas das submatrizes, temos acesso aos vetores que representam as penúltimas das submatrizes.
- Algoritmo proposto por Daniel Hirschberg em 1975.

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C																				
A																				
C																				
A																				
G																				
G																				
C																				
A																				
T																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27																		
A																				
C																				
A																				
G																				
G																				
C																				
A																				
T																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19																	
A																				
C																				
A																				
G																				
G																				
C																				
A																				
T																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1										
A																				
C																				
A																				
G																				
G																				
C																				
A																				
T																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44
A																				
C																				
A																				
G																				
G																				
C																				
A																				
T																				

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44
A																				
C																				
A																				
G																				
G																				
C																				
A																				
T																				

Gap = -5

Match = +3

Mismatch = -2



# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1	-6	-11	-16	-21	-26	-31	-36	-41
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7	2	-3	-8	-13	-18	-23	-28	-33
A	-50	-42	-34	-31	-23	-15	-12	-9	-1	7	7	5	5	0	-5	-10	-15	-20	-25	-30
G	-55	-47	-39	-36	-28	-20	-12	-14	-6	2	5	5	3	8	3	-2	-7	-12	-17	-22
G	-60	-52	-44	-41	-33	-25	-17	-14	-11	-3	0	3	3	6	6	1	-4	-4	-9	-14
C	-65	-57	-49	-46	-38	-30	-22	-19	-11	-8	-5	3	6	1	4	9	4	-1	-1	-6
A	-70	-62	-54	-51	-43	-35	-27	-24	-16	-8	-5	-2	1	4	-1	4	7	2	-3	-3
T	-75	-67	-59	-51	-48	-40	-32	-24	-21	-13	-10	-7	-4	-1	7	2	7	5	0	-5

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	α	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1	-6	-11	-16	-21	-26	-31	-36	-41
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7	2	-3	-8	-13	-18	-23	-28	-33
A	-50	-42	-34	-31	-23	-15	-12	-9	-1	7	7	5	5	0	-5	-10	-15	-20	-25	-30
G	-55	-47	-39	-36	-28	-20	-12	-14	-6	2	5	5	3	8	3	-2	-7	-12	-17	-22
G	-60	-52	-44	-41	-33	-25	-17	-14	-11	-3	0	3	3	6	6	1	-4	-4	-9	-14
C	-65	-57	-49	-46	-38	-30	-22	-19	-11	-8	-5	3	6	1	4	9	4	-1	-1	-6
A	-70	-62	-54	-51	-43	-35	-27	-24	-16	-8	-5	-2	1	4	-1	4	7	2	-3	-3
T	-75	-67	-59	-51	-48	-40	-32	-24	-21	-13	-10	-7	-4	-1	7	2	7	5	0	-5

Gap = -5

Match = +3

Mismatch = -2

# Alinhamento Global usando Espaço Linear

×	$\alpha$	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1	-6	-11	-16	-21	-26	-31	-36	-41
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7	2	-3	-8	-13	-18	-23	-28	-33
A	-50	-42	-34	-31	-23	-15	-12	-9	-1	7	7	5	5	0	-5	-10	-15	-20	-25	-30
G	-55	-47	-39	-36	-28	-20	-12	-14	-6	2	5	5	3	8	3	-2	-7	-12	-17	-22
G	-60	-52	-44	-41	-33	-25	-17	-14	-11	-3	0	3	3	6	6	1	-4	-4	-9	-14
C	-65	-57	-49	-46	-38	-30	-22	-19	-11	-8	-5	3	6	1	4	9	4	-1	-1	-6
A	-70	-62	-54	-51	-43	-35	-27	-24	-16	-8	-5	-2	1	4	-1	4	7	2	-3	-3
T	-75	-67	-59	-51	-48	-40	-32	-24	-21	-13	-10	-7	-4	-1	7	2	7	5	0	-5

-5	=	3	-2	3	3	3	-5	-5	3	-2	3	3	-2	3	-2	3	-5	-2	-5	-2
$\alpha$	=	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
$\beta$	=	A	T	T	G	G	-	-	C	C	A	C	A	G	G	C	-	A	-	T

# Alinhamento Global usando Espaço Linear

×	α	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>																																													
β	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95																																													
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87																																													
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79																																													
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71																																													
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63																																													
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55																																													
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47																																													
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44																																													
		-36	-41	-41	-36	-31	-26	-21	-16	-11	-6	-14	-14	-9	-7	-10	-18	-21	-24	-32	-40	<b>A</b>																																											
		-44	-39	-39	-39	-34	-29	-24	-19	-19	-14	-9	-12	-12	-7	-5	-13	-16	-19	-27	-35	<b>C</b>																																											
		-47	-47	-42	-37	-37	-37	-32	-27	-22	-17	-17	-12	-15	-10	-5	-8	-11	-19	-22	-30	<b>A</b>																																											
		-55	-50	-45	-40	-35	-35	-40	-35	-30	-25	-20	-15	-10	-13	-8	-3	-6	-14	-17	-25	<b>G</b>																																											
		-63	-58	-53	-48	-43	-38	-38	-38	-33	-28	-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	<b>G</b>																																											
		-71	-66	-61	-56	-51	-46	-41	-36	-36	-31	-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>																																											
		-79	-74	-69	-64	-59	-54	-49	-44	-39	-34	-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>																																											
		-87	-82	-77	-72	-67	-62	-57	-52	-47	-42	-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>																																											
		-95	-90	-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	-35	-30	-25	-20	-15	-10	-5	0	β																																											
		<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>	α	×																																											
Gap = -5																						Match = +3																						Mismatch = -2																					

# Alinhamento Global usando Espaço Linear

×	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>																																													
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95																																													
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87																																													
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79																																													
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71																																													
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63																																													
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55																																													
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47																																													
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44																																													
		-36	-41	-41	-36	-31	-26	-21	-16	-11	-6	-14	-14	-9	-7	-10	-18	-21	-24	-32	-40	<b>A</b>																																											
		-44	-39	-39	-39	-34	-29	-24	-19	-19	-14	-9	-12	-12	-7	-5	-13	-16	-19	-27	-35	<b>C</b>																																											
		-47	-47	-42	-37	-37	-37	-32	-27	-22	-17	-17	-12	-15	-10	-5	-8	-11	-19	-22	-30	<b>A</b>																																											
		-55	-50	-45	-40	-35	-35	-40	-35	-30	-25	-20	-15	-10	-13	-8	-3	-6	-14	-17	-25	<b>G</b>																																											
		-63	-58	-53	-48	-43	-38	-38	-38	-33	-28	-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	<b>G</b>																																											
		-71	-66	-61	-56	-51	-46	-41	-36	-36	-31	-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>																																											
		-79	-74	-69	-64	-59	-54	-49	-44	-39	-34	-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>																																											
		-87	-82	-77	-72	-67	-62	-57	-52	-47	-42	-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>																																											
		-95	-90	-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	-35	-30	-25	-20	-15	-10	-5	0	$\beta$																																											
		<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>	$\alpha$	×																																											
Gap = -5																						Match = +3																						Mismatch = -2																					

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>																																													
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95																																													
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87																																													
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79																																													
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71																																													
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63																																													
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55																																													
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47																																													
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44																																													
		-36	-41	-41	-36	-31	-26	-21	-16	-11	-6	-14	-14	-9	-7	-10	-18	-21	-24	-32	-40	<b>A</b>																																											
		-44	-39	-39	-39	-34	-29	-24	-19	-19	-14	-9	-12	-12	-7	-5	-13	-16	-19	-27	-35	<b>C</b>																																											
		-47	-47	-42	-37	-37	-37	-32	-27	-22	-17	-17	-12	-15	-10	-5	-8	-11	-19	-22	-30	<b>A</b>																																											
		-55	-50	-45	-40	-35	-35	-40	-35	-30	-25	-20	-15	-10	-13	-8	-3	-6	-14	-17	-25	<b>G</b>																																											
		-63	-58	-53	-48	-43	-38	-38	-38	-33	-28	-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	<b>G</b>																																											
		-71	-66	-61	-56	-51	-46	-41	-36	-36	-31	-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>																																											
		-79	-74	-69	-64	-59	-54	-49	-44	-39	-34	-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>																																											
		-87	-82	-77	-72	-67	-62	-57	-52	-47	-42	-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>																																											
		-95	-90	-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	-35	-30	-25	-20	-15	-10	-5	0	$\beta$																																											
		<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>	$\alpha$	$\times$																																											
<b>Gap = -5</b>																						<b>Match = +3</b>																						<b>Mismatch = -2</b>																					

# Alinhamento Global usando Espaço Linear

×	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95		
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87		
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79		
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71		
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63		
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55		
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47		
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44		
		-36	-41	-41	-36	-31	-26	-21	-16	-11	-6	-14	-14	-9	-7	-10	-18	-21	-24	-32	-40	<b>A</b>
		-44	-39	-39	-39	-34	-29	-24	-19	-19	-14	-9	-12	-12	-7	-5	-13	-16	-19	-27	-35	<b>C</b>
		-47	-47	-42	-37	-37	-37	-32	-27	-22	-17	-17	-12	-15	-10	-5	-8	-11	-19	-22	-30	<b>A</b>
		-55	-50	-45	-40	-35	-35	-40	-35	-30	-25	-20	-15	-10	-13	-8	-3	-6	-14	-17	-25	<b>G</b>
		-63	-58	-53	-48	-43	-38	-38	-38	-33	-28	-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	<b>G</b>
		-71	-66	-61	-56	-51	-46	-41	-36	-36	-31	-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>
		-79	-74	-69	-64	-59	-54	-49	-44	-39	-34	-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>
		-87	-82	-77	-72	-67	-62	-57	-52	-47	-42	-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>
		-95	-90	-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	-35	-30	-25	-20	-15	-10	-5	0	$\beta$
		<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>	$\alpha$	×
											<b>A</b>	<b>A</b>										
											<b>C</b>	<b>A</b>										

# Alinhamento Global usando Espaço Linear

×	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95		
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87		
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79		
<b>T</b>	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71		
<b>G</b>	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63		
<b>G</b>	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55		
<b>C</b>	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47		
<b>C</b>	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44		
		-36	-41	-41	-36	-31	-26	-21	-16	-11	-6	-14	-14	-9	-7	-10	-18	-21	-24	-32	-40	<b>A</b>
		-44	-39	-39	-39	-34	-29	-24	-19	-19	-14	-9	-12	-12	-7	-5	-13	-16	-19	-27	-35	<b>C</b>
		-47	-47	-42	-37	-37	-37	-32	-27	-22	-17	-17	-12	-15	-10	-5	-8	-11	-19	-22	-30	<b>A</b>
		-55	-50	-45	-40	-35	-35	-40	-35	-30	-25	-20	-15	-10	-13	-8	-3	-6	-14	-17	-25	<b>G</b>
		-63	-58	-53	-48	-43	-38	-38	-38	-33	-28	-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	<b>G</b>
		-71	-66	-61	-56	-51	-46	-41	-36	-36	-31	-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>
		-79	-74	-69	-64	-59	-54	-49	-44	-39	-34	-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>
		-87	-82	-77	-72	-67	-62	-57	-52	-47	-42	-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>
		-95	-90	-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	-35	-30	-25	-20	-15	-10	-5	0	$\beta$
		<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>	$\alpha$	×

**A** **A**  
**C** **A**



# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	A	C	T	G	G	G	T	C		
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40		
A	-5	3	-2	-7	-12	-17	-22	-27	-32		
T	-10	-2	1	1	-4	-9	-14	-19	-24		
T	-15	-7	-4	4	-1	-6	-11	-11	-16		
		-16	-11	-6	-1	4	4	-4	-7	-15	G
		-24	-19	-14	-9	-4	1	1	-2	-10	G
		-32	-27	-22	-17	-12	-7	-2	3	-5	C
		-40	-35	-30	-25	-20	-15	-10	-5	0	
		A	C	T	G	G	G	T	C		

A A  
C A

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	A	C	T	G	G	G	T	C		
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40		
A	-5	3	-2	-7	-12	-17	-22	-27	-32		
T	-10	-2	1	1	-4	-9	-14	-19	-24		
T	-15	-7	-4	4	-1	-6	-11	-11	-16		
		-16	-11	-6	-1	4	4	-4	-7	-15	G
		-24	-19	-14	-9	-4	1	1	-2	-10	G
		-32	-27	-22	-17	-12	-7	-2	3	-5	C
		-40	-35	-30	-25	-20	-15	-10	-5	0	
		A	C	T	G	G	G	T	C		

A A  
C A

# Alinhamento Global usando Espaço Linear

×	$\alpha$	<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		
$\beta$	<b>0</b>	-5	-10	-15	-20	-25	-30	-35	-40		
<b>A</b>	-5	3	-2	-7	-12	-17	-22	-27	-32		
<b>T</b>	-10	-2	1	1	-4	-9	-14	-19	-24		
<b>T</b>	-15	-7	-4	<b>4</b>	-1	-6	-11	-11	-16		
		-16	-11	-6	<b>-1</b>	4	4	-4	-7	-15	<b>G</b>
		-24	-19	-14	-9	-4	1	1	-2	-10	<b>G</b>
		-32	-27	-22	-17	-12	-7	-2	3	-5	<b>C</b>
		-40	-35	-30	-25	-20	-15	-10	-5	0	
		<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		

**A** **A**  
**C** **A**

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	A	C	T	G	G	G	T	C		
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40		
A	-5	3	-2	-7	-12	-17	-22	-27	-32		
T	-10	-2	1	1	-4	-9	-14	-19	-24		
T	-15	-7	-4	4	-1	-6	-11	-11	-16		
		-16	-11	-6	-1	4	4	-4	-7	-15	G
		-24	-19	-14	-9	-4	1	1	-2	-10	G
		-32	-27	-22	-17	-12	-7	-2	3	-5	C
		-40	-35	-30	-25	-20	-15	-10	-5	0	
		A	C	T	G	G	G	T	C		

T G  
T G

A A  
C A

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	A	C	T	G	G	G	T	C		
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40		
A	-5	3	-2	-7	-12	-17	-22	-27	-32		
T	-10	-2	1	1	-4	-9	-14	-19	-24		
T	-15	-7	-4	4	-1	-6	-11	-11	-16		
		-16	-11	-6	-1	4	4	-4	-7	-15	G
		-24	-19	-14	-9	-4	1	1	-2	-10	G
		-32	-27	-22	-17	-12	-7	-2	3	-5	C
		-40	-35	-30	-25	-20	-15	-10	-5	0	
		A	C	T	G	G	G	T	C		

T	G			A	A
T	G			C	A

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	<b>A</b>	<b>C</b>			
$\beta$	0	-5	-10			
<b>A</b>	-5	3	-2			
		-7	-2	-5	<b>T</b>	
		-10	-5	0		
		<b>A</b>	<b>C</b>			

**T G**  
**T G**

**A A**  
**C A**

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	<b>A</b>	<b>C</b>			
$\beta$	0	-5	-10			
<b>A</b>	-5	3	-2			
		-7	-2	-5	<b>T</b>	
		-10	-5	0		
		<b>A</b>	<b>C</b>			

T G  
T G

A A  
C A

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	<b>A</b>	<b>C</b>		
$\beta$	0	-5	-10		
<b>A</b>	-5	<b>3</b>	-2		
		-7	<b>-2</b>	-5	<b>T</b>
		-10	-5	0	
		<b>A</b>	<b>C</b>		

T G  
T G

A A  
C A



# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	<b>A</b>	<b>C</b>		
$\beta$	0	-5	-10		
<b>A</b>	-5	<b>3</b>	-2		
		-7	<b>-2</b>	-5	<b>T</b>
		-10	-5	0	
		<b>A</b>	<b>C</b>		

A C T G  
A T T G

A A  
C A

# Alinhamento Global usando Espaço Linear

		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		
	0	-5	-10	-15	-20		
<b>G</b>	-5	3	-2	-7	-12		
		-12	-7	-2	3	-5	<b>C</b>
		-20	-15	-10	-5	0	
		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		

A C T G  
A T T G

A A  
C A

# Alinhamento Global usando Espaço Linear

		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		
	0	-5	-10	-15	-20		
<b>G</b>	-5	3	-2	-7	-12		
		-12	-7	-2	3	-5	<b>C</b>
		-20	-15	-10	-5	0	
		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		

A C T G  
A T T G

A A  
C A

# Alinhamento Global usando Espaço Linear

		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		
	0	-5	-10	-15	-20		
<b>G</b>	-5	<b>3</b>	-2	-7	-12		
		-12	<b>-7</b>	-2	3	-5	<b>C</b>
		-20	-15	-10	-5	0	
		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		

A C T G  
A T T G

A A  
C A

# Alinhamento Global usando Espaço Linear

		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		
	0	-5	-10	-15	-20		
<b>G</b>	-5	<b>3</b>	-2	-7	-12		
		-12	<b>-7</b>	-2	3	-5	<b>C</b>
		-20	-15	-10	-5	0	
		<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>		

A C T G G G            A A  
A T T G G -            C A

# Alinhamento Global usando Espaço Linear

		<b>T</b>	<b>C</b>
	0	-5	-10
<b>C</b>	-5	-2	-2

A C T G G G            A A  
A T T G G -            C A

# Alinhamento Global usando Espaço Linear

		<b>T</b>	<b>C</b>
	0	-5	-10
<b>C</b>	-5	-2	-2

A C T G G G T C A A  
A T T G G - - C C A

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
	0	-5	-10	-15	-20	-25	-30	-35	-40	-45		
<b>C</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37		
<b>A</b>	-10	-2	1	-4	-9	-14	-19	-24	-29	-34		
<b>G</b>	-15	-7	-4	4	-1	-6	-11	-16	-21	-26		
		-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	<b>G</b>
		-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>
		-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>
		-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>
		-45	-40	-35	-30	-25	-20	-15	-10	-5	0	
		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		

**A C T G G G T C A A**  
**A T T G G - - C C A**



# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			
	<b>0</b>	<b>-5</b>	<b>-10</b>	<b>-15</b>	<b>-20</b>	<b>-25</b>	<b>-30</b>	<b>-35</b>	<b>-40</b>	<b>-45</b>			
<b>C</b>	<b>-5</b>	<b>3</b>	<b>-2</b>	<b>-7</b>	<b>-12</b>	<b>-17</b>	<b>-22</b>	<b>-27</b>	<b>-32</b>	<b>-37</b>			
<b>A</b>	<b>-10</b>	<b>-2</b>	<b>1</b>	<b>-4</b>	<b>-9</b>	<b>-14</b>	<b>-19</b>	<b>-24</b>	<b>-29</b>	<b>-34</b>			
<b>G</b>	<b>-15</b>	<b>-7</b>	<b>-4</b>	<b>4</b>	<b>-1</b>	<b>-6</b>	<b>-11</b>	<b>-16</b>	<b>-21</b>	<b>-26</b>			
		<b>-23</b>	<b>-18</b>	<b>-13</b>	<b>-13</b>	<b>-11</b>	<b>-6</b>	<b>-1</b>	<b>-9</b>	<b>-12</b>	<b>-20</b>	<b>G</b>	
		<b>-26</b>	<b>-21</b>	<b>-21</b>	<b>-16</b>	<b>-11</b>	<b>-11</b>	<b>-6</b>	<b>-4</b>	<b>-12</b>	<b>-15</b>	<b>C</b>	
		<b>-34</b>	<b>-29</b>	<b>-24</b>	<b>-19</b>	<b>-14</b>	<b>-14</b>	<b>-9</b>	<b>-4</b>	<b>-7</b>	<b>-10</b>	<b>A</b>	
		<b>-37</b>	<b>-32</b>	<b>-27</b>	<b>-22</b>	<b>-17</b>	<b>-12</b>	<b>-12</b>	<b>-7</b>	<b>-2</b>	<b>-5</b>	<b>T</b>	
		<b>-45</b>	<b>-40</b>	<b>-35</b>	<b>-30</b>	<b>-25</b>	<b>-20</b>	<b>-15</b>	<b>-10</b>	<b>-5</b>	<b>0</b>		
	<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>				

**A C T G G G T C A A**  
**A T T G G - - C C A**

# Alinhamento Global usando Espaço Linear

		C	C	G	T	C	T	G	C	G		
	0	-5	-10	-15	-20	-25	-30	-35	-40	-45		
C	-5	3	-2	-7	-12	-17	-22	-27	-32	-37		
A	-10	-2	1	-4	-9	-14	-19	-24	-29	-34		
G	-15	-7	-4	4	-1	-6	-11	-16	-21	-26		
		-23	-18	-13	-13	-11	-6	-1	-9	-12	-20	G
		-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	C
		-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	A
		-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	T
		-45	-40	-35	-30	-25	-20	-15	-10	-5	0	
C	C	G	T	C	T	G	C	G				

A C T G G G T C A A  
A T T G G - - C C A

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
	<b>0</b>	-5	-10	-15	-20	-25	-30	-35	-40	-45		
<b>C</b>	-5	3	-2	-7	-12	-17	-22	-27	-32	-37		
<b>A</b>	-10	-2	1	-4	-9	-14	-19	-24	-29	-34		
<b>G</b>	-15	-7	-4	<b>4</b>	-1	-6	-11	-16	-21	-26		
		-23	-18	-13	<b>-13</b>	-11	-6	-1	-9	-12	-20	<b>G</b>
		-26	-21	-21	-16	-11	-11	-6	-4	-12	-15	<b>C</b>
		-34	-29	-24	-19	-14	-14	-9	-4	-7	-10	<b>A</b>
		-37	-32	-27	-22	-17	-12	-12	-7	-2	-5	<b>T</b>
		-45	-40	-35	-30	-25	-20	-15	-10	-5	0	
		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>			
<b>A</b>	<b>T</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>-</b>	<b>-</b>	<b>C</b>	<b>C</b>	<b>A</b>			
					<b>G</b>	<b>T</b>						
					<b>G</b>	<b>G</b>						

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
	<b>0</b>	<b>-5</b>	<b>-10</b>	<b>-15</b>	<b>-20</b>	<b>-25</b>	<b>-30</b>	<b>-35</b>	<b>-40</b>	<b>-45</b>		
<b>C</b>	<b>-5</b>	<b>3</b>	<b>-2</b>	<b>-7</b>	<b>-12</b>	<b>-17</b>	<b>-22</b>	<b>-27</b>	<b>-32</b>	<b>-37</b>		
<b>A</b>	<b>-10</b>	<b>-2</b>	<b>1</b>	<b>-4</b>	<b>-9</b>	<b>-14</b>	<b>-19</b>	<b>-24</b>	<b>-29</b>	<b>-34</b>		
<b>G</b>	<b>-15</b>	<b>-7</b>	<b>-4</b>	<b>4</b>	<b>-1</b>	<b>-6</b>	<b>-11</b>	<b>-16</b>	<b>-21</b>	<b>-26</b>		
		<b>-23</b>	<b>-18</b>	<b>-13</b>	<b>-13</b>	<b>-11</b>	<b>-6</b>	<b>-1</b>	<b>-9</b>	<b>-12</b>	<b>-20</b>	<b>G</b>
		<b>-26</b>	<b>-21</b>	<b>-21</b>	<b>-16</b>	<b>-11</b>	<b>-11</b>	<b>-6</b>	<b>-4</b>	<b>-12</b>	<b>-15</b>	<b>C</b>
		<b>-34</b>	<b>-29</b>	<b>-24</b>	<b>-19</b>	<b>-14</b>	<b>-14</b>	<b>-9</b>	<b>-4</b>	<b>-7</b>	<b>-10</b>	<b>A</b>
		<b>-37</b>	<b>-32</b>	<b>-27</b>	<b>-22</b>	<b>-17</b>	<b>-12</b>	<b>-12</b>	<b>-7</b>	<b>-2</b>	<b>-5</b>	<b>T</b>
		<b>-45</b>	<b>-40</b>	<b>-35</b>	<b>-30</b>	<b>-25</b>	<b>-20</b>	<b>-15</b>	<b>-10</b>	<b>-5</b>	<b>0</b>	
		<b>C</b>	<b>C</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>		
<b>A</b>	<b>C</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>G</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>A</b>			
<b>A</b>	<b>T</b>	<b>T</b>	<b>G</b>	<b>G</b>	<b>-</b>	<b>-</b>	<b>C</b>	<b>C</b>	<b>A</b>			
						<b>G</b>	<b>T</b>					
						<b>G</b>	<b>G</b>					

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>																	
		0	-5	-10																
<b>C</b>		-5	3	-2																
			-7	-2	-5	<b>A</b>														
			-10	-5	0															
			<b>C</b>	<b>C</b>																

A C T G G G T C A A                    G T  
A T T G G - - C C A                    G G

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>										
		0	-5	-10									
<b>C</b>	-5	<b>3</b>	-2										
		-7	<b>-2</b>	-5	<b>A</b>								
		-10	-5	0									
		<b>C</b>	<b>C</b>										

A C T G G G T C A A            G T  
A T T G G - - C C A            G G

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>											
		0	-5	-10										
<b>C</b>	-5	<b>3</b>	-2											
		-7	<b>-2</b>	-5	<b>A</b>									
		-10	-5	0										
		<b>C</b>	<b>C</b>											

A C T G G G T C A A                    G T  
A T T G G - - C C A                    G G

# Alinhamento Global usando Espaço Linear

		<b>C</b>	<b>C</b>											
		0	-5	-10										
<b>C</b>	-5	<b>3</b>	-2											
		-7	<b>-2</b>	-5	<b>A</b>									
		-10	-5	0										
		<b>C</b>	<b>C</b>											

A C T G G G T C A A C C G T  
A T T G G - - C C A C A G G



# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T  
A T T G G - - C C A C A G G

		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			
	<b>0</b>	-5	-10	-15	-20	-25			
<b>C</b>	-5	3	-2	-7	-12	-17			
		-14	-14	-9	-4	-7	-10	<b>A</b>	
		-17	-12	-12	-7	-2	-5	<b>T</b>	
		-25	-20	-15	-10	-5	0		
		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			

# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T  
A T T G G - - C C A C A G G

		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			
	0	-5	-10	-15	-20	-25			
<b>C</b>	-5	3	-2	-7	-12	-17			
		-14	-14	-9	-4	-7	-10	<b>A</b>	
		-17	-12	-12	-7	-2	-5	<b>T</b>	
		-25	-20	-15	-10	-5	0		
		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			

# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T  
A T T G G - - C C A C A G G

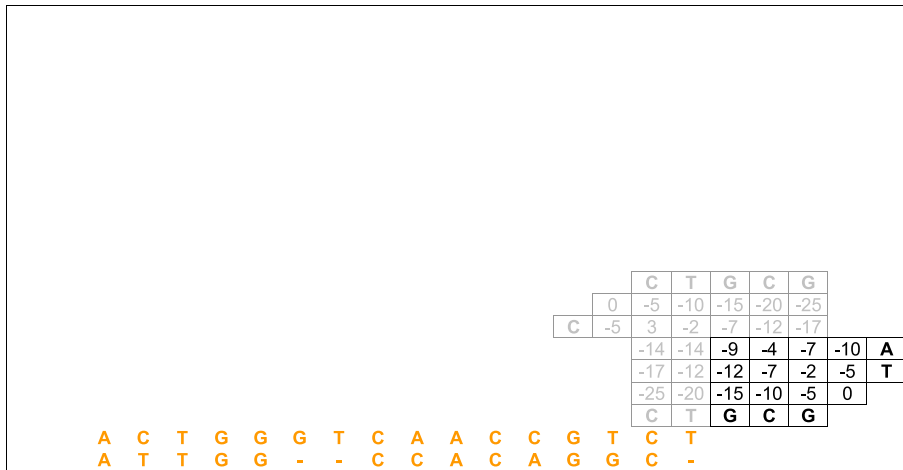
		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			
	<b>0</b>	-5	-10	-15	-20	-25			
<b>C</b>	-5	<b>3</b>	-2	-7	-12	-17			
		-14	<b>-14</b>	-9	-4	-7	-10	<b>A</b>	
		-17	-12	-12	-7	-2	-5	<b>T</b>	
		-25	-20	-15	-10	-5	0		
		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			

# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T C T  
A T T G G - - C C A C A G G C -

		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			
	<b>0</b>	-5	-10	-15	-20	-25			
<b>C</b>	-5	<b>3</b>	-2	-7	-12	-17			
		-14	<b>-14</b>	-9	-4	-7	-10	<b>A</b>	
		-17	-12	-12	-7	-2	-5	<b>T</b>	
		-25	-20	-15	-10	-5	0		
		<b>C</b>	<b>T</b>	<b>G</b>	<b>C</b>	<b>G</b>			

# Alinhamento Global usando Espaço Linear



# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T C T  
A T T G G - - C C A C A G G C -

		<b>G</b>	<b>C</b>	<b>G</b>					
	0	-5	-10	-15					
<b>A</b>	-5	-2	-7	-12					
		-12	-7	-2	-5	<b>T</b>			
		-15	-10	-5	0				
		<b>G</b>	<b>C</b>	<b>G</b>					

# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T C T  
A T T G G - - C C A C A G G C -

		<b>G</b>	<b>C</b>	<b>G</b>					
	0	-5	-10	-15					
<b>A</b>	-5	-2	<b>-7</b>	-12					
		-12	-7	<b>-2</b>	-5	<b>T</b>			
		-15	-10	-5	0				
		<b>G</b>	<b>C</b>	<b>G</b>					

# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T C T  
A T T G G - - C C A C A G G C -

		<b>G</b>	<b>C</b>	<b>G</b>															
		0	-5	-10	-15														
<b>A</b>	-5	-2	<b>-7</b>	-12															
		-12	-7	<b>-2</b>	-5	<b>T</b>													
		-15	-10	-5	0														
		<b>G</b>	<b>C</b>	<b>G</b>															



# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T C T  
A T T G G - - C C A C A G G C -

			<b>G</b>	<b>C</b>	<b>G</b>														
		0	-5	-10	-15														
<b>A</b>	-5	-2	-7	-12															
		-12	-7	-2	-5	<b>T</b>													
		-15	-10	-5	0														
		<b>G</b>	<b>C</b>	<b>G</b>															

# Alinhamento Global usando Espaço Linear

A C T G G G T C A A C C G T C T  
A T T G G - - C C A C A G G C -

		<b>G</b>	<b>C</b>	<b>G</b>															
	0	-5	-10	-15															
<b>A</b>	-5	-2	-7	-12															
		-12	-7	-2	-5	<b>T</b>													
		-15	-10	-5	0														
		<b>G</b>	<b>C</b>	<b>G</b>															

# Alinhamento Global usando Espaço Linear

		<b>G</b>
	0	-5
<b>A</b>	-5	-2

A C T G G G T C A A C C G T C T G C G  
A T T G G - - C C A C A G G C - A - T

# Alinhamento Global usando Espaço Linear

$\times$	$\alpha$	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
A	-5	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
T	-10	-2	1	1	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
T	-15	-7	-4	4	-1	-6	-11	-11	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-1	7	2	-3	-8	-13	-18	-23	-28	-33	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	2	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-45	-50	-55
C	-30	-22	-14	-11	-3	5	8	3	3	-2	-7	-12	-17	-22	-27	-32	-37	-42	-42	-47
C	-35	-27	-19	-16	-8	0	3	6	6	1	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44
A	-40	-32	-24	-21	-13	-5	-2	1	4	9	4	-1	-6	-11	-16	-21	-26	-31	-36	-41
C	-45	-37	-29	-26	-18	-10	-7	-4	4	4	7	7	2	-3	-8	-13	-18	-23	-28	-33
A	-50	-42	-34	-31	-23	-15	-12	-9	-1	7	7	5	5	0	-5	-10	-15	-20	-25	-30
G	-55	-47	-39	-36	-28	-20	-12	-14	-6	2	5	5	3	8	3	-2	-7	-12	-17	-22
G	-60	-52	-44	-41	-33	-25	-17	-14	-11	-3	0	3	3	6	6	1	-4	-4	-9	-14
C	-65	-57	-49	-46	-38	-30	-22	-19	-11	-8	-5	3	6	1	4	9	4	-1	-1	-6
A	-70	-62	-54	-51	-43	-35	-27	-24	-16	-8	-5	-2	1	4	-1	4	7	2	-3	-3
T	-75	-67	-59	-51	-48	-40	-32	-24	-21	-13	-10	-7	-4	-1	7	2	7	5	0	-5

-5	=	3	-2	3	3	3	-5	-5	3	-2	3	3	-2	3	-2	3	-5	-2	-5	-2
$\alpha$	=	A	C	T	G	G	G	T	C	A	A	C	C	G	T	C	T	G	C	G
$\beta$	=	A	T	T	G	G	-	-	C	C	A	C	A	G	G	C	-	A	-	T

# Complexidade de Tempo e de Espaço

- Complexidade de Espaço:
  - $E(m, n) = \Theta(m + n)$
- Complexidade de Tempo:
  1. O número de chamadas recursivas é no máximo proporcional aos tamanhos das seqüência  $\alpha$  e  $\beta$ , já que conseguimos alinhar pelo menos um caracter a cada chamada recursiva:
    - $T(m, n) = O(m + n) \times O(mn) = O(m^2n + mn^2)$
  2. A cada passo do algoritmo, metade da matriz  $M$  é descartada, logo o número de chamadas recursivas deve ser no máximo proporcional ao logaritmo do tamanho da matriz:
    - $T(m, n) = O(\lceil \log_2 mn \rceil) \times O(mn) = O(mn \log mn)$
  3. Podemos obter um resultado melhor fazendo uma análise assintótica agregada (soma de todos os passos do algoritmo):
    - $T(m, n) = \sum_{k=0}^{\lceil \log_2 mn \rceil} \Theta\left(\frac{mn}{2^k}\right)$
    - $T(m, n) = \Theta\left(\sum_{k=0}^{\lceil \log_2 mn \rceil} \frac{mn}{2^k}\right)$
    - $T(m, n) = \Theta\left(mn \sum_{k=0}^{\lceil \log_2 mn \rceil} \frac{1}{2^k}\right) = \Theta\left(mn \sum_{k=0}^{\infty} \frac{1}{2^k}\right)$
    - $T(m, n) = \Theta(mn)$

- Algoritmo de Daniel Hirschberg (1975):
  - Pontuação Aditiva
  - Divisão e Conquista + Programação Dinâmica (1 matriz)
  - Complexidade de Espaço:  $\Theta(m + n)$
  - Complexidade de Tempo:  $\Theta(mn)$
- Algoritmo de Osamu Gotoh (1982):
  - Pontuação Afim
  - Programação Dinâmica (3 matrizes)
  - Complexidade de Espaço:  $\Theta(mn)$
  - Complexidade de Tempo:  $\Theta(mn)$
- Algoritmo de Eugene Myers e Webb Miller (1988):
  - Pontuação Afim
  - Divisão e Conquista + Programação Dinâmica (3 matrizes)
  - Complexidade de Espaço:  $\Theta(m + n)$
  - Complexidade de Tempo:  $\Theta(mn)$

- Algoritmo proposto por Webb Miller e Eugene Myers (1988).
- Complexidade:
  - Tempo (pior caso):  $\Theta(mn \log mn)$ .
  - Espaço (esperado):  $\Theta(m + n)$ .
- Pouco utilizado na prática.

## Exercício

*Escreva algoritmos, em pseudocódigo, que dadas duas sequências  $\alpha$  e  $\beta$ , com respectivamente  $m$  e  $n$  caracteres, calcule:*

- *O valor de um alinhamento global ótimo entre  $\alpha$  e  $\beta$ .*
- *O valor de um alinhamento semi-global ótimo entre  $\alpha$  e  $\beta$ .*
- *O valor de um alinhamento local ótimo entre  $\alpha$  e  $\beta$ .*

*Seus algoritmos devem considerar função aditiva para penalidade de blocos de buracos e usar espaço linear em relação ao tamanho da entrada.*



# **Alinhamento Global de Sequências Similares**

---

## Alinhamento Global de Sequências Similares

- Suponha que queremos alinhar duas sequências  $\alpha$  e  $\beta$  de mesmo tamanho  $n$  (é fácil adaptar o algoritmo para funcionar com duas sequências de tamanho diferentes).
- Se as duas sequências são similares, é razoável supor que existam poucos buracos no alinhamento ótimo entre as duas sequências.
- Seja  $2k$  o número de buracos no alinhamento ótimo.
  - Note que no caso das duas sequências terem o mesmo tamanho, o número de buracos será sempre par.
- O que podemos afirmar em relação as células da matriz que representam alinhamentos ótimos entre  $\alpha$  e  $\beta$ ?
  - Elas estão, no máximo, a  $k$  células de distância da diagonal principal.
- Ideia: preencher apenas as células que estão a  $k$  células de distância da diagonal principal da matriz de Programação Dinâmica.

# Alinhamento Global de Sequências Similares

- Como determinar o valor de  $k$ , sem conhecer as sequências?
- Seja  $M$  a pontuação para *match*,  $m$  para *mismatch* e  $g$  para *gap*.
- Qual a maior pontuação possível para um alinhamento que não esteja completamente contido na faixa de  $k$  células de distância da diagonal principal?
  - $M \times (n - k - 1) + 2 \times (k + 1) \times g$
- Algoritmo:
  1. Inicialização:  $k = 1$
  2. Preencha a matriz de Programação Dinâmica considerando apenas as células com distância no máximo  $k$  para a diagonal principal.
  3. Se  $M[n, n] > M \times (n - k - 1) + 2 \times (k + 1) \times g$ , pare e retorne o alinhamento máximo.
  4. Caso contrário, incremente o valor de  $k$  e volte ao passo 2.
- Algoritmo proposto por James Fickett (1983).

# Alinhamento Global de Sequências Similares

$\times$	$\alpha$	G	C	G	A	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
C	-5	-2	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
A	-10	-7	-4	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
G	-15	-7	-9	-1	-6	-6	-11	-16	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-6	-3	-8	-8	-13	-13	-18	-18	-23	-28	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	-8	-5	-10	-10	-10	-15	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60
C	-30	-22	-14	-11	-8	-10	-7	-12	-12	-12	-17	-17	-22	-27	-32	-32	-37	-42	-47	-52
G	-35	-27	-19	-11	-13	-10	-12	-9	-9	-14	-9	-14	-19	-24	-29	-34	-34	-39	-44	-49
A	-40	-32	-24	-16	-8	-10	-7	-9	-11	-11	-14	-6	-11	-16	-21	-26	-31	-36	-41	-46
A	-45	-37	-29	-21	-13	-5	-7	-4	-9	-13	-13	-11	-8	-13	-18	-23	-28	-33	-38	-43
A	-50	-42	-34	-26	-18	-10	-2	-4	-6	-11	-15	-10	-13	-10	-15	-20	-25	-30	-35	-40
A	-55	-47	-39	-31	-23	-15	-7	1	-4	-8	-13	-12	-12	-15	-12	-17	-22	-27	-32	-37
G	-60	-52	-44	-36	-28	-20	-12	-4	4	-1	-5	-10	-14	-14	-17	-14	-19	-24	-29	-34
T	-65	-57	-49	-41	-33	-25	-17	-9	-1	7	2	-3	-7	-11	-11	-16	-16	-16	-21	-26
G	-70	-62	-54	-46	-38	-30	-22	-14	-6	2	10	5	0	-5	-10	-13	-18	-18	-18	-23
A	-75	-67	-59	-51	-43	-35	-27	-19	-11	-3	5	13	8	3	-2	-7	-12	-17	-20	-20
T	-80	-72	-64	-56	-48	-40	-32	-24	-16	-8	0	8	16	11	6	1	-4	-9	-14	-19
G	-85	-77	-69	-61	-53	-45	-37	-29	-21	-13	-5	3	11	14	9	4	-1	-6	-11	-16
T	-90	-82	-74	-66	-58	-50	-42	-34	-26	-18	-10	-2	6	14	17	12	7	2	-3	-8
A	-95	-87	-79	-71	-63	-55	-47	-39	-31	-23	-15	-7	1	9	12	15	10	5	0	-5

n	=	19	M	=	3	m	=	-2	g	=	-5
---	---	----	---	---	---	---	---	----	---	---	----

# Alinhamento Global de Sequências Similares

$\times$	$\alpha$	G	C	G	A	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C
$\beta$	0	-5																		
C	-5	-2	-2																	
A		-7	-4	-4																
G			-9	-1	-6															
G				-6	-3	-8														
G					-8	-5	-10													
C						-10	-7	-12												
G							-12	-9	-9											
A								-9	-11	-11										
A									-11	-13	-13									
A										-13	-15	-10								
A											-15	-12	-12							
G												-17	-14	-14						
T													-14	-11	-11					
G														-16	-13	-13				
A															-18	-15	-15			
T																-20	-17	-12		
G																	-22	-17	-14	
T																		-19	-19	-16
A																			-21	-21

n = 19   M = 3   m = -2   g = -5   k = 1

máximo = 31

# Alinhamento Global de Sequências Similares

×	α	G	C	G	A	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C
β	0	-5	-10																	
C	-5	-2	-2	-7																
A	-10	-7	-4	-4	-4															
G		-7	-9	-1	-6	-6														
G			-9	-6	-3	-8	-8													
G				-6	-8	-5	-10	-10												
C					-8	-10	-7	-12	-12											
G						-10	-12	-9	-9	-14										
A							-7	-9	-11	-11	-16									
A								-4	-9	-13	-13	-13								
A									-6	-11	-15	-10	-15							
A										-8	-13	-12	-12	-17						
G											-5	-10	-14	-14	-19					
T												-7	-7	-11	-11	-16				
G													-9	-9	-13	-13	-18			
A														-11	-11	-15	-15	-20		
T															-8	-13	-17	-12	-17	
G																-10	-15	-17	-14	-19
T																	-12	-12	-17	-16
A																		-14	-14	-19

n = 19   M = 3   m = -2   g = -5   k = 2

máximo = 18

# Alinhamento Global de Sequências Similares

$\times$	$\alpha$	G	C	G	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C	
$\beta$	0	-5	-10	-15																
C	-5	-2	-2	-7	-12															
A	-10	-7	-4	-4	-4	-9														
G	-15	-7	-9	-1	-6	-6	-11													
G		-12	-9	-6	-3	-8	-8	-13												
G			-14	-6	-8	-5	-10	-10	-10											
C				-11	-8	-10	-7	-12	-12	-12										
G					-13	-10	-12	-9	-9	-14	-9									
A						-10	-7	-9	-11	-11	-14	-6								
A							-7	-4	-9	-13	-13	-11	-8							
A								-4	-6	-11	-15	-10	-13	-10						
A									-6	-8	-13	-12	-12	-15	-12					
G										-8	-5	-10	-14	-14	-17	-14				
T											-10	-7	-7	-11	-11	-16	-16			
G												-12	-9	-9	-13	-13	-18	-18		
A													-14	-11	-11	-15	-15	-20	-20	
T														-11	-8	-13	-17	-12	-17	-22
G															-13	-10	-15	-17	-14	-19
T																-15	-12	-12	-17	-16
A																	-17	-14	-14	-19

n = 19   M = 3   m = -2   g = -5   k = 3

máximo = 5

# Alinhamento Global de Sequências Similares

$\times$	$\alpha$	G	C	G	A	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C
$\beta$	0	-5	-10	-15	-20															
C	-5	-2	-2	-7	-12	-17														
A	-10	-7	-4	-4	-4	-9	-14													
G	-15	-7	-9	-1	-6	-6	-11	-16												
G	-20	-12	-9	-6	-3	-8	-8	-13	-13											
G		-17	-14	-6	-8	-5	-10	-10	-10	-15										
C			-14	-11	-8	-10	-7	-12	-12	-12	-17									
G				-11	-13	-10	-12	-9	-9	-14	-9	-14								
A					-8	-10	-7	-9	-11	-11	-14	-6	-11							
A						-5	-7	-4	-9	-13	-13	-11	-8	-13						
A							-2	-4	-6	-11	-15	-10	-13	-10	-15					
A								1	-4	-8	-13	-12	-12	-15	-12	-17				
G									4	-1	-5	-10	-14	-14	-17	-14	-19			
T										7	2	-3	-7	-11	-11	-16	-16	-16		
G											10	5	0	-5	-10	-13	-18	-18	-18	
A												13	8	3	-2	-7	-12	-17	-20	-20
T													16	11	6	1	-4	-9	-14	-19
G														14	9	4	-1	-6	-11	-16
T															17	12	7	2	-3	-8
A																15	10	5	0	-5

n = 19   M = 3   m = -2   g = -5   k = 4

máximo = -8



# Alinhamento Global de Sequências Similares

$\times$	$\alpha$	G	C	G	A	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C
$\beta$	0	-5	-10	-15	-20	-25														
C	-5	-2	-2	-7	-12	-17	-22													
A	-10	-7	-4	-4	-4	-9	-14	-19												
G	-15	-7	-9	-1	-6	-6	-11	-16	-16											
G	-20	-12	-9	-6	-3	-8	-8	-13	-13	-18										
G	-25	-17	-14	-6	-8	-5	-10	-10	-10	-15	-15									
C		-22	-14	-11	-8	-10	-7	-12	-12	-12	-17	-17								
G			-19	-11	-13	-10	-12	-9	-9	-14	-9	-14	-19							
A				-16	-8	-10	-7	-9	-11	-11	-14	-6	-11	-16						
A					-13	-5	-7	-4	-9	-13	-13	-11	-8	-13	-18					
A						-10	-2	-4	-6	-11	-15	-10	-13	-10	-15	-20				
A							-7	1	-4	-8	-13	-12	-12	-15	-12	-17	-22			
G								-4	4	-1	-5	-10	-14	-14	-17	-14	-19	-24		
T									-1	7	2	-3	-7	-11	-11	-16	-16	-16	-21	
G										2	10	5	0	-5	-10	-13	-18	-18	-18	-23
A											5	13	8	3	-2	-7	-12	-17	-20	-20
T												8	16	11	6	1	-4	-9	-14	-19
G													11	14	9	4	-1	-6	-11	-16
T														14	17	12	7	2	-3	-8
A															12	15	10	5	0	-5

n = 19   M = 3   m = -2   g = -5   k = 5

máximo = -21

# Alinhamento Global de Sequências Similares

$\times$	$\alpha$	G	C	G	A	A	A	A	G	T	G	A	T	T	T	C	C	T	C	C
$\beta$	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60	-65	-70	-75	-80	-85	-90	-95
C	-5	-2	-2	-7	-12	-17	-22	-27	-32	-37	-42	-47	-52	-57	-62	-67	-72	-77	-82	-87
A	-10	-7	-4	-4	-4	-9	-14	-19	-24	-29	-34	-39	-44	-49	-54	-59	-64	-69	-74	-79
G	-15	-7	-9	-1	-6	-6	-11	-16	-16	-21	-26	-31	-36	-41	-46	-51	-56	-61	-66	-71
G	-20	-12	-9	-6	-3	-8	-8	-13	-13	-18	-18	-23	-28	-33	-38	-43	-48	-53	-58	-63
G	-25	-17	-14	-6	-8	-5	-10	-10	-10	-15	-15	-20	-25	-30	-35	-40	-45	-50	-55	-60
C	-30	-22	-14	-11	-8	-10	-7	-12	-12	-12	-17	-17	-22	-27	-32	-32	-37	-42	-47	-52
G	-35	-27	-19	-11	-13	-10	-12	-9	-9	-14	-9	-14	-19	-24	-29	-34	-34	-39	-44	-49
A	-40	-32	-24	-16	-8	-10	-7	-9	-11	-11	-14	-6	-11	-16	-21	-26	-31	-36	-41	-46
A	-45	-37	-29	-21	-13	-5	-7	-4	-9	-13	-13	-11	-8	-13	-18	-23	-28	-33	-38	-43
A	-50	-42	-34	-26	-18	-10	-2	-4	-6	-11	-15	-10	-13	-10	-15	-20	-25	-30	-35	-40
A	-55	-47	-39	-31	-23	-15	-7	1	-4	-8	-13	-12	-12	-15	-12	-17	-22	-27	-32	-37
G	-60	-52	-44	-36	-28	-20	-12	-4	4	-1	-5	-10	-14	-14	-17	-14	-19	-24	-29	-34
T	-65	-57	-49	-41	-33	-25	-17	-9	-1	7	2	-3	-7	-11	-11	-16	-16	-16	-21	-26
G	-70	-62	-54	-46	-38	-30	-22	-14	-6	2	10	5	0	-5	-10	-13	-18	-18	-18	-23
A	-75	-67	-59	-51	-43	-35	-27	-19	-11	-3	5	13	8	3	-2	-7	-12	-17	-20	-20
T	-80	-72	-64	-56	-48	-40	-32	-24	-16	-8	0	8	16	11	6	1	-4	-9	-14	-19
G	-85	-77	-69	-61	-53	-45	-37	-29	-21	-13	-5	3	11	14	9	4	-1	-6	-11	-16
T	-90	-82	-74	-66	-58	-50	-42	-34	-26	-18	-10	-2	6	14	17	12	7	2	-3	-8
A	-95	-87	-79	-71	-63	-55	-47	-39	-31	-23	-15	-7	1	9	12	15	10	5	0	-5

n	=	19	M	=	3	m	=	-2	g	=	-5
---	---	----	---	---	---	---	---	----	---	---	----

# Complexidade de Tempo e de Espaço

- Complexidade de Espaço:
  - $E(n) = \Theta(kn)$
  - Pior caso  $k = \Theta(n)$ , logo:  $E(n) = \Theta(n^2)$
  - Usando as ideias de Hirschberg, é possível implementar o algoritmo usando apenas espaço linear no tamanho das sequências ( $E(n) = \Theta(n)$ ).
- Complexidade de Tempo:
  - Primeira abordagem:
    - Incremento (passo 4):  $k = k + 1$
    - $T(n) = \sum_{i=1}^k \Theta(in) = \Theta(n) \sum_{i=1}^k i$
    - $T(n) = \Theta(n) \frac{(k+1)k}{2} = \Theta(nk^2)$
    - Pior caso  $k = \Theta(n)$ , logo:  $T(n) = \Theta(n^3)$
  - Segunda abordagem:
    - Incremento (passo 4):  $k = 2 \times k$
    - $T(n) = \sum_{i=0}^{\log_2 k} \Theta(2^i n) = \Theta(n) \sum_{i=0}^{\log_2 k} 2^i$
    - $T(n) = \Theta(n)(2^{1+\log_2 k} - 1) = \Theta(n)(2k - 1) = \Theta(nk)$
    - Pior caso  $k = \Theta(n)$ , logo:  $T(n) = \Theta(n^2)$

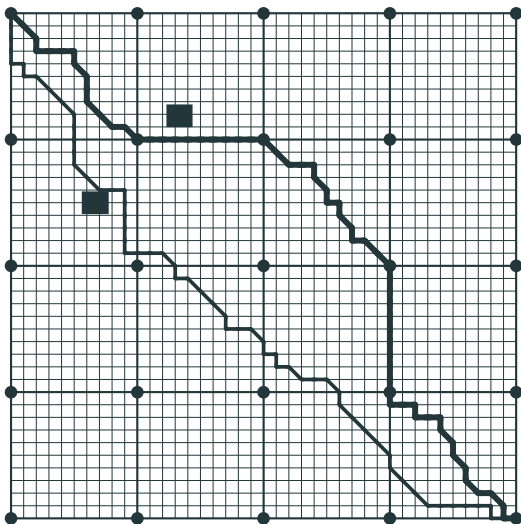
# **Algoritmo Sub-Quadrático para Alinhamento Global**

---

## Algoritmo Sub-Quadrático para Alinhamento Global

- É possível obter um algoritmo de tempo sub-quadrático para alinhamento global considerando pontuação aditiva para buracos?
- Por simplificação, vamos considerar a seguinte variação do problema de alinhamento global chamada de alinhamento de blocos:
  - Dadas duas seqüências  $\alpha$  e  $\beta$  de tamanhos, respectivamente,  $m = m' \times t$  e  $n = n' \times t$ , ambas divididas em blocos de  $t$  caracteres consecutivos, encontrar o melhor alinhamento entre os blocos da seqüências  $\alpha$  com os blocos da seqüência  $\beta$ .
  - Note que cada um dos blocos das seqüências deve ser alinhado com um bloco da outra seqüência ou com um bloco de buracos de tamanho  $t$ .
- Na figura do próximo slide é mostrado (a) um alinhamento de blocos e (b) um alinhamento global entre um par de seqüências.

# Alinhamento de Blocos $\times$ Alinhamento Global



## Algoritmo Sub-Quadrático para Alinhamento de Blocos

- Four-Russian Algorithm (Arlazarov, Dinic, Kronrod e Faradzev), 1970.
- O método foi desenvolvido apenas por Kronrod, único russo do grupo.
- Considere que ambas as sequências são formadas por caracteres de um alfabeto de quatro letras, por exemplo,  $\{A, C, G, T\}$ .
- Passos do algoritmo:
  - Dividir a matriz de programação dinâmica em quadrados de dimensão  $t \times t$ .
  - Pré-computar os alinhamentos de todos os pares de sequências de tamanho  $t$ , em tempo  $\Theta(4^t 4^t t^2) = \Theta(16^t t^2)$ .
  - Preencher a matriz de Programação Dinâmica, usando os valores pré-computados, em tempo  $\Theta(\frac{m}{t} \frac{n}{t} t) = \Theta(\frac{mn}{t})$ .
- Complexidade:  $T(n) = \Theta(16^t t^2 + \frac{mn}{t})$

# Algoritmo Sub-Quadrático para Alinhamento de Blocos

- Complexidade:  $T(n) = \Theta(16^t t^2 + \frac{mn}{t})$
- Valor de  $t$  que minimiza a complexidade do algoritmo:
  - $t = \log_{16}(m + n)$ .
- Complexidade:
  - $T(n) = \Theta((m + n) \log^2(m + n) + \frac{mn}{\log(m+n)}) = \Theta(\frac{mn}{\log(m+n)})$
- William Masek e Michael Paterson (1980) mostraram como calcular o alinhamento global em tempo  $\Theta(\frac{mn}{\log(m+n)})$ , usando uma adaptação do Four-Russian Algorithm.
- Algoritmo não utilizado na prática.