

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mohamed Khider-BISKRA.

Faculté des sciences exactes et des
sciences de la nature et de la vie



Département d'informatique

Calcul à l'ADN et Bio-informatique

Réalisé par :

Dr. TBERMACINE Ahmed

ALIGNEMENT DE SÉQUENCES

❖ En quoi consiste l'alignement

- **Alignement de séquences**= *alignement séquentiel*
- Est une manière de disposer les composantes des ADN, des ARN, ou des séquences primaires de protéines pour identifier les zones de concordance qui traduisent des similarités ou dissemblances de nature historique.

ALIGNEMENT DE SÉQUENCES

- Pourquoi

- Recherche homologie
- identifier des sites fonctionnels
- prédire la/les fonction(s) d'une protéine
- prédire la structure secondaire (voire tertiaire) d'une protéine
- établir une phylogénie

ALIGNEMENT DE SÉQUENCES

❖ Types d'alignement

- *l'alignement par paires*: aligner 2 séquences.
- Il est possible de réaliser un alignement :
 - **global**, c'est à dire entre les 2 séquences sur toutes leurs longueurs
 - **local** entre une séquence et une partie de l'autre séquence
- *l'alignement multiple*, qui est un alignement global, consiste à aligner plus de 2 séquences.

ALIGNEMENT DE SÉQUENCES

◦ Cadre formel

Définition 1: Alphabet

Un alphabet Σ est un ensemble fini de symboles distincts. Dans le cas de séquences d'ADN ou d'acide aminés on définit le symbole vide ou gap par -.

- L'alphabet de l'ADN est composé par les symboles :
 - -, A, C, G, T
- L'alphabet des acides aminés est composé des symboles:
 - -, A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y

ALIGNEMENT DE SÉQUENCES

◦ Cadre formel

Définition 2: Séquence, segment

On appelle séquence S une suite ordonnée de caractères pris dans un alphabet Σ :

$$S = \{x_1, x_2, \dots, x_n\}$$

On note $|S| = n$ la longueur de la séquence.

- on appelle **segment** tout ou partie d'une séquence.
- les **séquences** étant des mots sur un alphabet on utilise les termes de **préfixe** et **suffixe** pour désigner respectivement un segment de début ou de fin de la séquence.

ALIGNEMENT DE SÉQUENCES

◦ Cadre formel

Définition 3: Alignement

Soit $S = \{S_1, S_2, \dots, S_k\}$ un ensemble de k séquences définies sur un alphabet Σ :

$\forall u, 1 \leq u \leq k, S_u = \{x_1^u, x_2^u, \dots, x_{|S_u|}^u\}$. Un alignement de S , noté $A(\{S_1, S_2, \dots, S_k\})$ est une matrice :

$$\begin{bmatrix} a_1^1 & \cdots & a_q^1 \\ \vdots & \ddots & \vdots \\ a_1^k & \cdots & a_q^k \end{bmatrix}$$

$$\forall u \in \{1, \dots, k\}, \forall v \in \{1, \dots, q\}, a_v^u \in \Sigma$$

avec

$$\forall u \in \{1, \dots, k\}, \max_{1 \leq v \leq q} (|S_v|) \leq q \leq \sum_{u=1}^{u=k} |S_u|$$

ALIGNEMENT DE SÉQUENCES

- **Matrice de score**

- l'alignement consiste à mettre en correspondance les caractères de deux séquences, S1 et S2, en autorisant le décalage avec l'insertion d'espaces (gap) pour compenser les différences.

ALIGNEMENT DE SÉQUENCES

○ Matrice de score

- un alignement peut être construit avec quatre opérations de base :
 - l'appariement (a,a) pour des caractères identiques,
 - la substitution (a,b) pour des caractères différents,
 - l'insertion d'un gap dans S1 (-,b), et l'insertion d'un gap dans S2 (a,-).
- **Remarque:**
- l'insertion dans une séquence peut être vue comme une suppression dans la séquence correspondante: **InDel** (Insertion-Deletion).

ALIGNEMENT DE SÉQUENCES

- **Exemple**

- Soient les séquences $S_1 = \{\text{CATTGC}\}$ et $S_2 = \{\text{ACAGTC}\}$.
Un exemple d'alignement est :

-CATTGC
ACAGT-C

ALIGNEMENT DE SÉQUENCES

○ Exemple

- La construction de cet alignement correspond à la suite d'opérations suivantes :

(-,A)	insertion d'un gap dans S_1
(C,C)	appariement
(A,A)	appariement
(T,G)	substitution
(T,T)	appariement
(G,-)	insertion d'un gap dans S_2
(C,C)	appariement

- d'autres alignements semblent tout aussi valables :

-CAT-TGC

ACA-GT-C

-CA-TTGC

ACAGT--C

ALIGNEMENT DE SÉQUENCES

```
AAB24882 TYHMCQFHCYVNNHSGEKLIECNERSKAFSCPSHLQCHKRRQIGERTHEHNQCGKAFPT 60
AAB24881 -----YECNQCGKAFQAQSSSLKCHYRTHIGEKPYECNQCGKAFSK 40
          ****:***: * :** * :****:* *****..

AAB24882 PSHLQYHRTHTGEKPYECHQCGQAFKKCSLLQRHKRTHIGEKPYE-CNQCGKAFQAQ- 116
AAB24881 HSHLQCHRTHTGEKPYECNQCGKAFSQHGLLQRHKRTHIGEKPYMNVINMVKPLHNS 98
          **** *:*****:***:*.: .*****
                                     *.::
```

http://en.wikipedia.org/wiki/Sequence_alignment

identité

Insertion / Délétion

substitution

3 situations sont possibles :

- Les caractères sont les mêmes (appariement) : **identité**
- Les caractères ne sont pas les mêmes (mésappariement) : **substitution**
- L'une des positions est un espace (-) : **insertion/délétion** (InDel ou gap)

ALIGNEMENT DE SÉQUENCES

○ Matrices de Scores pour l'ADN

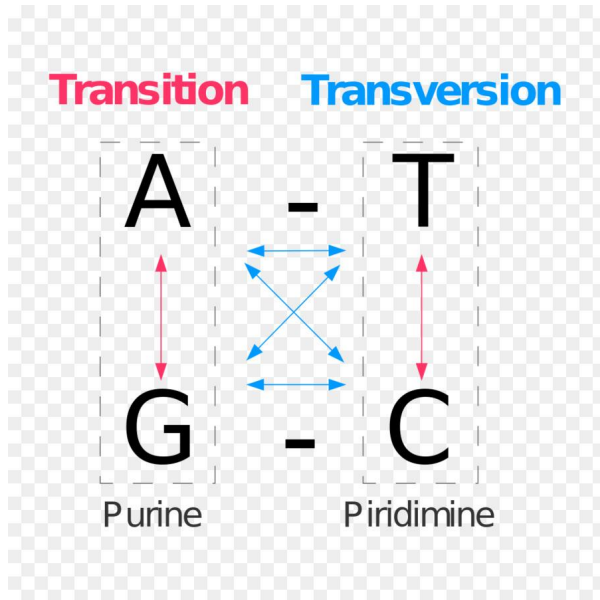
a) La matrice Identité:

w(x,y)	-	A	C	G	T
-	?	0	0	0	0
A	0	1	0	0	0
C	0	0	1	0	0
G	0	0	0	1	0
T	0	0	0	0	1

ALIGNEMENT DE SÉQUENCES

○ Matrices de Scores pour l'ADN

b) La matrice de Transition/ Transversion



$w(x,y)$	-	A	C	G	T
-	?	0	0	0	0
A	0	3	0	1	0
C	0	0	3	0	1
G	0	1	0	3	0
T	0	0	1	0	3

ALIGNEMENT DE SÉQUENCES

○ Matrices de Scores pour l'ADN

b) BLAST

w(x,y)	-	A	C	G	T
-	?	0	0	0	0
A	0	5	-4	-4	-4
C	0	-4	5	-4	-4
G	0	-4	-4	5	-4
T	0	-4	-4	-4	5

ALIGNEMENT DE SÉQUENCES

- **Matrices de Scores pour les protéines**

- **a) PAM** (Percent Accepted Mutation).

La matrice la plus couramment utilisée est PAM 250.

- **b) BLOSUM** (Blocks Substitution Matrices),

La matrice la plus couramment utilisée est BLOSUM 62 ce qui correspond à des protéines n'ayant pas plus de 62% d'identité.

- **c) Gonnet**

ALIGNEMENT PAR PAIRES (PAIRWISE SEQUENCE ALIGNMENT)

❖ Problème d'alignement par paires

- Soient deux séquences S_1 et S_2 , une matrice de score w . Le problème d'alignement par paires consiste à déterminer un alignement de coût optimal selon w .
- obtenir un alignement de coût optimal $\rightarrow$ une énumération de tous les alignements possibles et à retenir ceux de coût optimal.

ALIGNEMENT PAR PAIRES

- **Définition : Problème d'alignement par paires**
- **Problème** le nombre d'alignements différents de deux séquences de même longueur n est donné par la formule:

$$\sum_{k=0}^{k=n} C_{n+k}^k \times C_n^k = 1 + C_{2n}^n = 1 + \frac{(2n)!}{(n!)^2} \sim \frac{2^{2n}}{\sqrt{\pi n}}$$

- Pour deux séquences de longueur n , voici le nombre d'alignements :

n	alignements
10	1.84e5
20	1.37e11
30	1.18e17
40	1.07e23
50	1.00e29
100	9.05e58

ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE

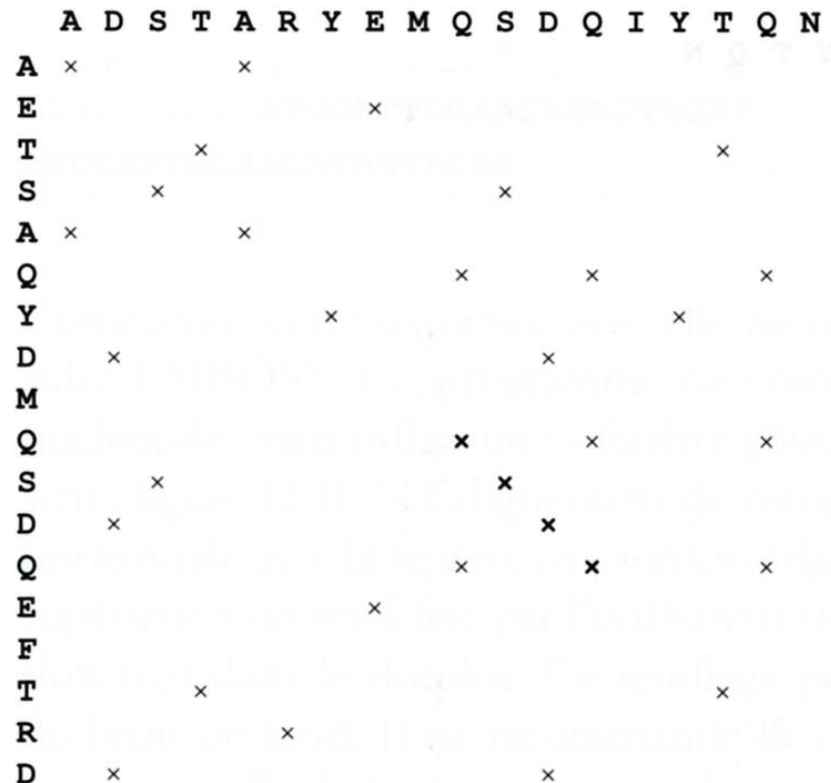
◦ DotPlot

ADSTARYEMQSDQIYTQN

| | | | | |

AETSAQYDMQSDQEFTRD

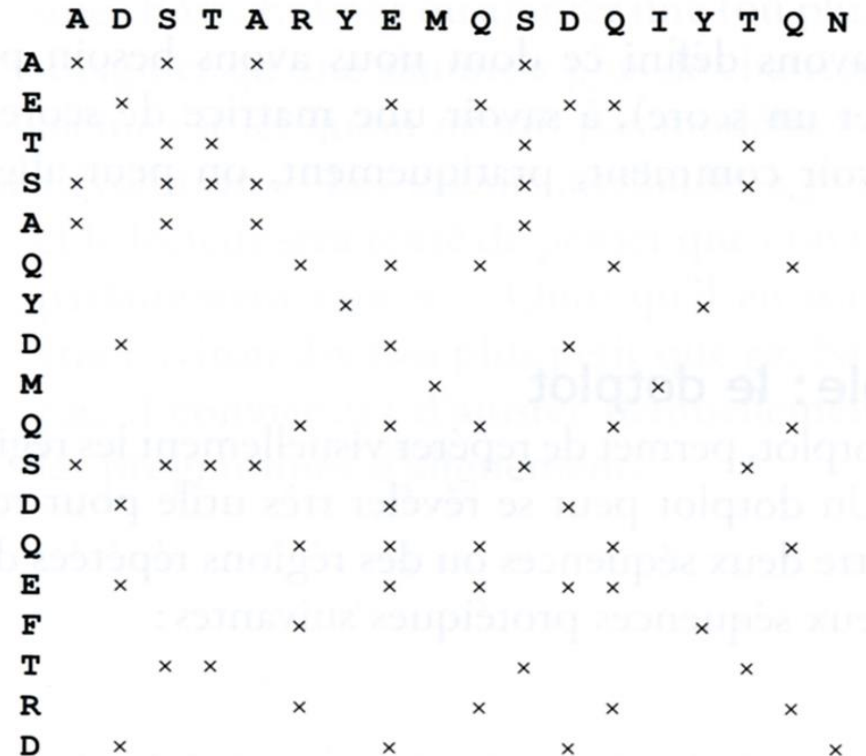
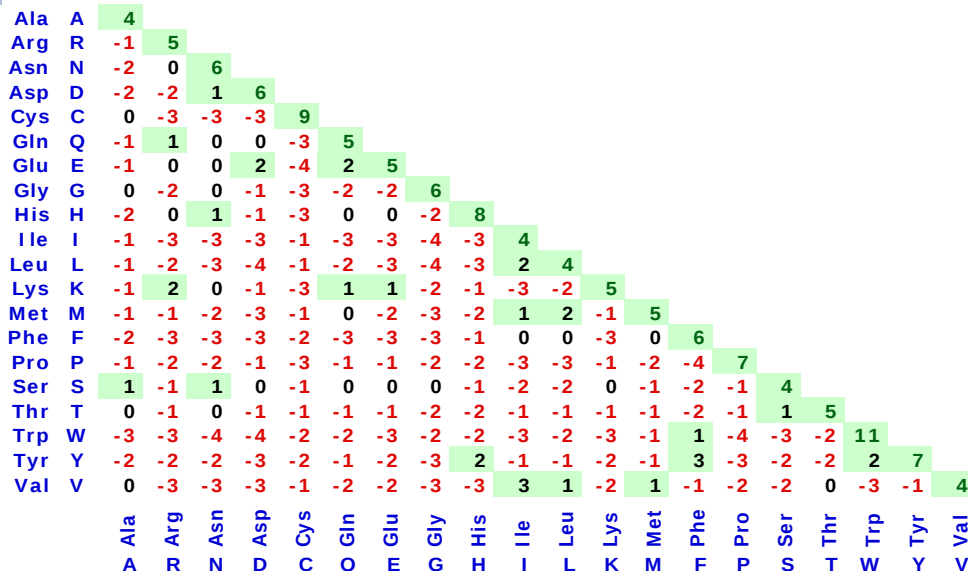
- Le dot plot est une représentation graphique simple des résidus identiques entre les deux séquences.
 - ▣ Les deux séquences sont représentées sur les deux axes
 - ▣ Un point (dot) est tracé pour chaque correspondance entre deux résidus de séquences.
 - ▣ Les lignes diagonales révèlent les régions alignables entre les deux séquences.



TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

- Outre les identités, on peut marquer les similarités entre acides aminés (substitutions conservatives, d'après une matrice de substitution donnée)
 - Exemple: marquage des paires de résidus ayant un score BLOSUM62 > 1 .

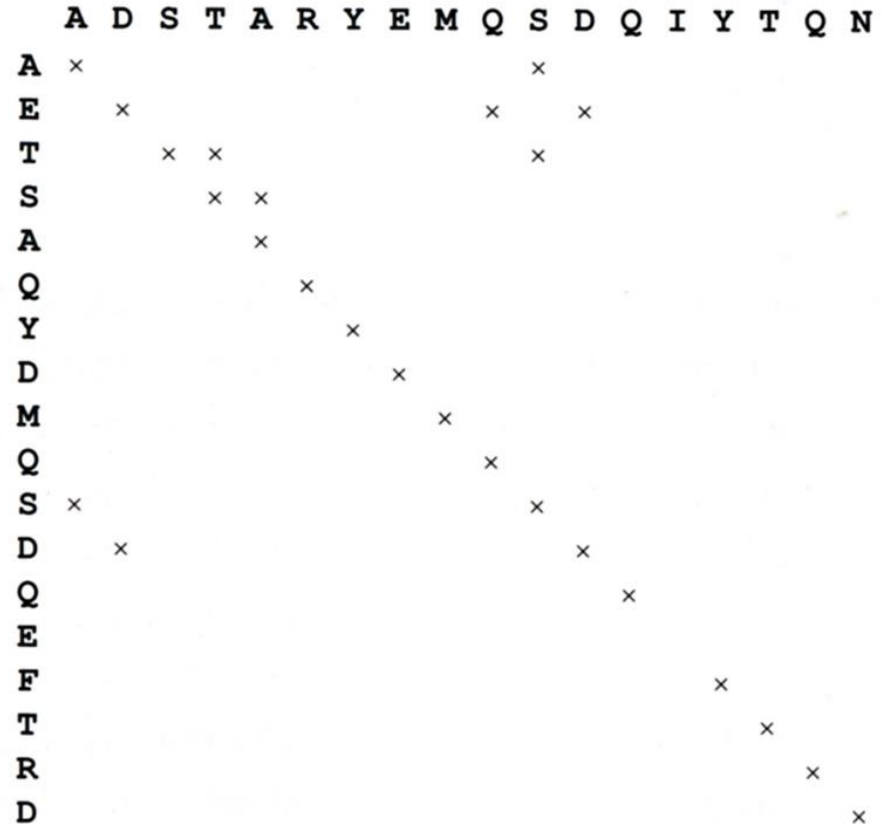
ADSTARYEMQSDQIYTQN
|:::|:|:||||| :|::
AETSAQYDMQSDQEFTRD



ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

- On peut appliquer à la matrice de points un filtrage par fenêtre glissante.
 - ▣ On n'affiche que les diagonale comportant au moins 2 points marqués successifs.
- À chaque position de la matrice on extrait la paire de mots de taille w qui commence aux positions correspondantes des deux séquences.
- Le score de la paire de mots est calculé en additionnant les scores des paires des résidus (d'après la matrice de substitution).
- Si le score dépasse un seuil défini par l'utilisateur, une diagonale noire s'affiche à la position correspondante.

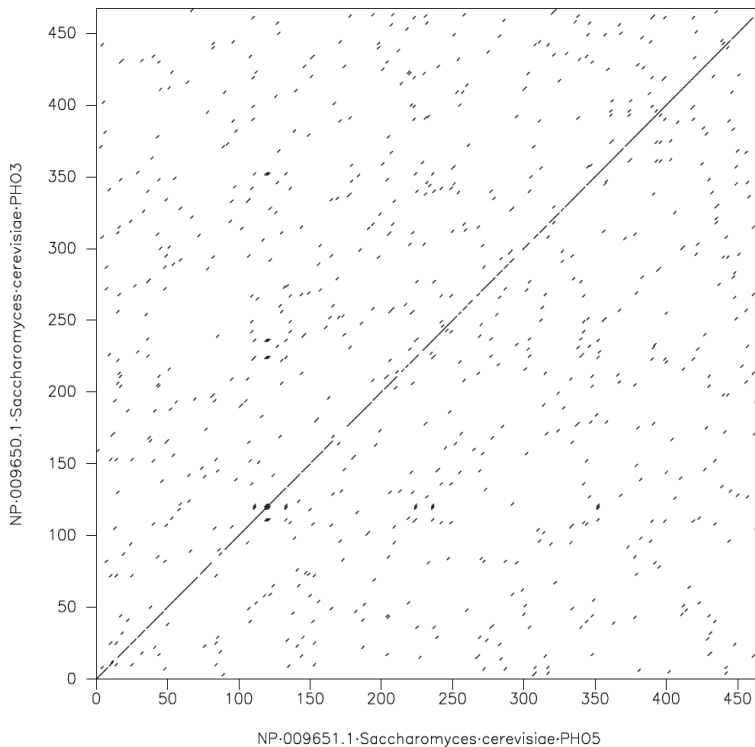


ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE

(DOTPLOT)

dottup (19-May-2006)



Example: protein sequences of the Pho5p and Pho3p phosphatases in the yeast *Saccharomyces cerevisiae*

- Exemple de dot plot: comparaison de deux séquences peptidiques de phosphatases de la levure *Saccharomyces cerevisiae*.

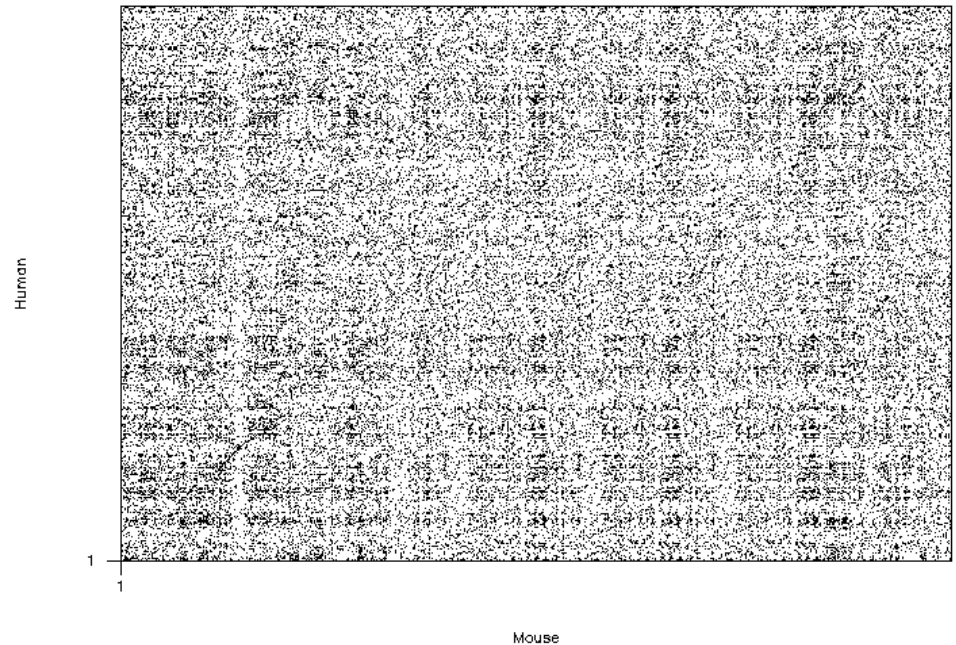
- L'axe X représente la première séquence (PHO5),
- L'axe Y représente la seconde séquence (PHO3)
- Un point est affiché pour chaque correspondance entre deux résidus des séquences.
- Les lignes obliques représentent des régions d'identité entre deux séquences.
- La diagonale est fortement marquée, car les deux séquences sont homologues et conservées sur toute leur longueur.
- Le reste du dot plot est clairsemé de points, qui représentent les identités sporadiques entre acides aminés.

ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

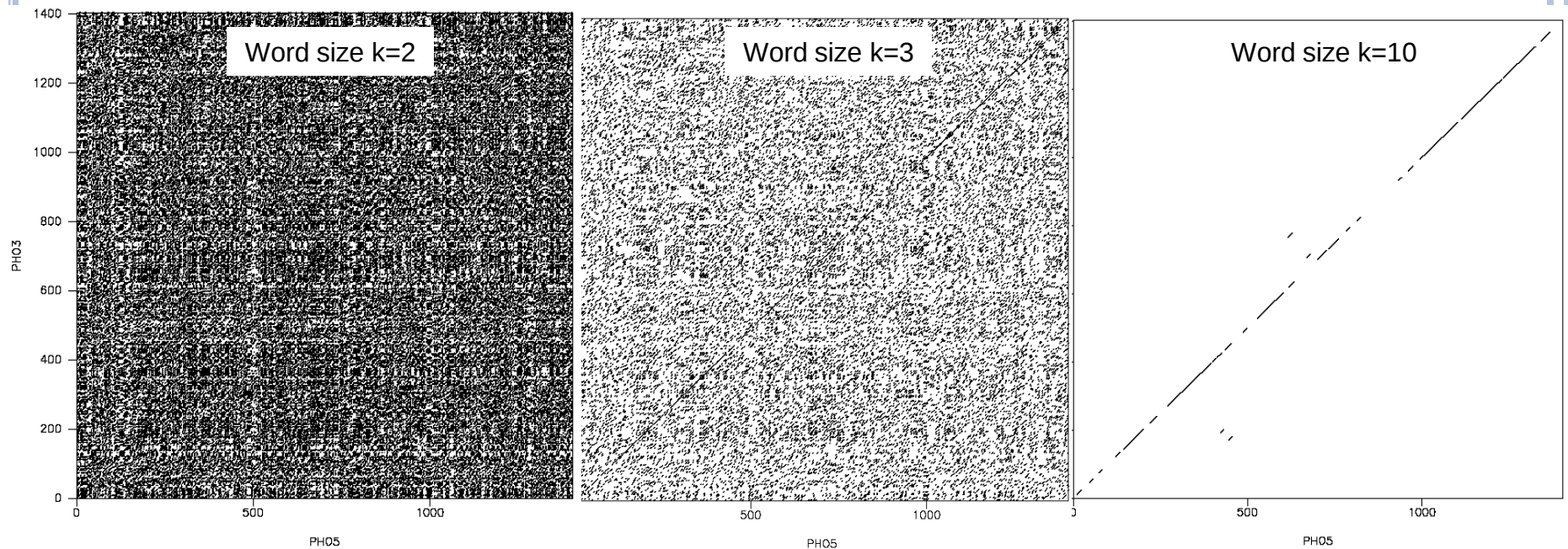
- Dot plot des séquences nucléiques de deux opsines de souris.
- Le plot « brut » est difficilement lisible, car on trouve des identités entre nucléotides un peu partout (1 chance sur 4).
- On distingue néanmoins un renforcement de la diagonale, même sur ce graphique extrêmement bruité.

Human versus mouse
genomic region surrounding OPN1MW



ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)



ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

○ Exercice 1

Afficher la matrice de points

+Filtrage: fenêtre =2

+Filtrage: fenêtre =3

+Filtrage: fenêtre =4

○ Solution

ACTCGAGCTATCG

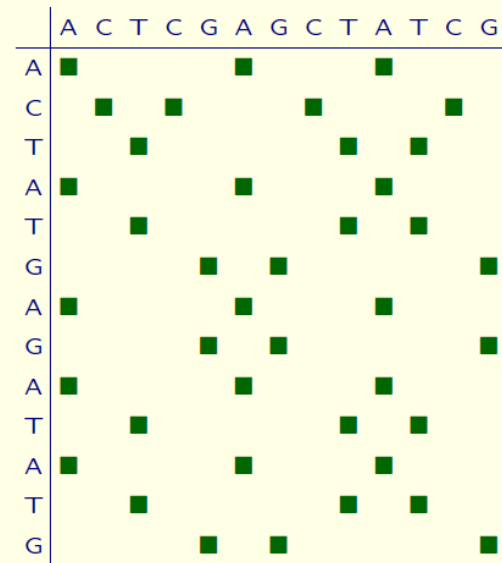
ACTATGAGATATG

match (identité) → ■

mismatch → □

diagonale = région similaire

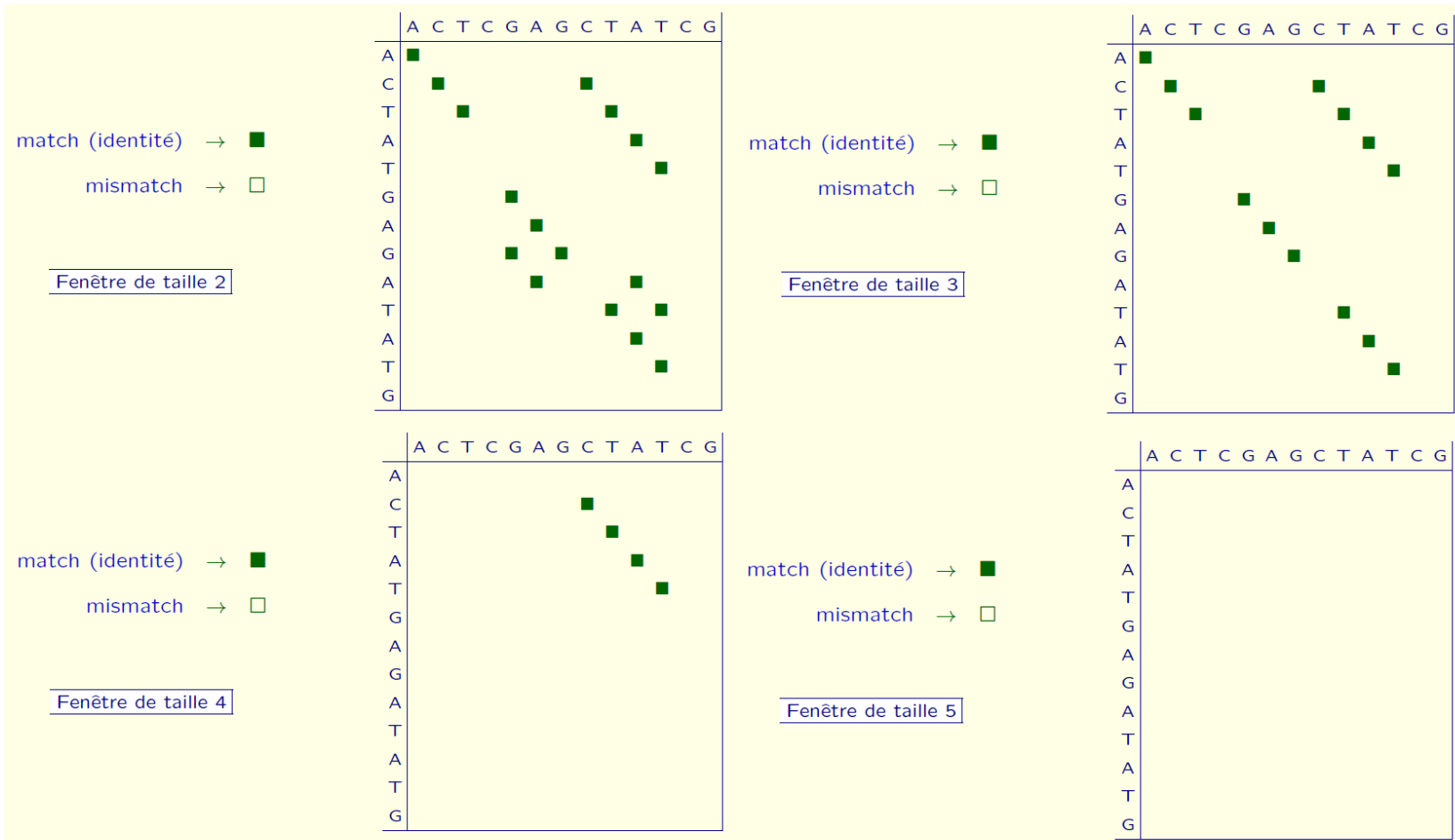
diagonale inversée = région miroir



ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

o Solution



ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

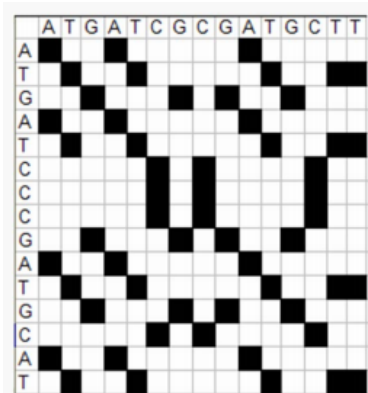
○ Exercice

Afficher la matrice de points

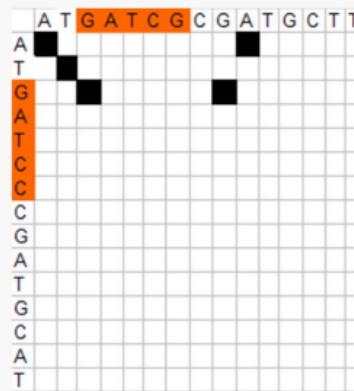
Filtrage: fenêtre (4,5)

ATGATCGCGATGCTT
ATGATCCCGATGCAT

○ Solution



Dotplot non filtré



exemple :

fenêtres de 5 nucléotides

Critère retenu : au moins 4 positions identiques

ALIGNEMENT PAR PAIRES

TYPE DE L'ALIGNEMENT PAR PAIRE (DOTPLOT)

○ Avantages

- ❖ simple, visuel
- ❖ très informatif
- ❖ pas ou peu de perte d'information

○ Inconvénients

- ❖ *Identification* : pas de méthode de détection automatique
- ❖ *Interprétation*: pas de mesure objective et d'évaluation de la pertinence de l'information

➔ **Besoin d'une mesure quantitative de similarité**

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

- La PD est la méthode la plus intéressante pour l'alignement de deux séquences car sa complexité temporelle et spatiale est en $O(n^2)$.
- La PD est appliquée à des problèmes d'optimisation pour lesquels un choix doit être fait entre plusieurs solutions possibles afin d'aboutir à une solution optimale.
- Le terme *Programmation* fait ici référence à une méthode basée sur le calcul de tableaux de valeurs.

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

- L'algorithme d'alignement global basé sur la programmation dynamique a été mis au point par Needleman et Wunsch en 1970.
- Une version pour l'alignement local fut ensuite introduite par Smith et Waterman en 1981.

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

- **Algorithme**

- Soient deux séquences à aligner:

$$S = \{ x_1, x_2, \dots, x_N \} \text{ et } T = \{ y_1, y_2, \dots, y_M \}$$

- L'algorithme d'alignement utilise deux matrices de dimensions $(N+1) \times (M+1)$:
- $M[0..N, 0..P]$: *la matrice des scores optimaux d'alignement* qui stocke le coût des meilleurs alignements intermédiaires.
- Au final $M[N, P]$ contient le coût du meilleur alignement global suivant w .

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

- **Algorithme**

- $\text{Dir}[0..N, 0..P]$: *la matrice des directions* qui indique la suite d'opérations d'édition à réaliser afin d'obtenir un alignement optimal.
- L'algorithme se déroule en 2 étapes :
 1. calcul des matrices M et Dir
 2. création d'un alignement optimal par lecture de la matrice des directions

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

◦ Algorithme

1ere étape: calcul des matrices M et Dir

- Soient deux séquences S et T de longueurs respectives N et P. La recherche d'un alignement global optimal entre S et T suivant une fonction de score w est obtenu par construction d'une *matrice des scores optimaux d'alignement* $M[0..N,0..P]$ telle que :

- initialisation :

- $M[0,0] = 0$
- $M[i,0] = M[i-1,0] + w(x_i,-)$ pour tout i de 1 à N
- $M[0,j] = M[0,j-1] + w(-,y_j)$ pour tout j de 1 à P

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

◦ Algorithme

1ere étape: calcul des matrices M et Dir

◦ calcul du score optimal:

$$\circ M[i,j] = \text{optimise} \begin{cases} M[i-1,j-1] + w(x_i, y_j) \\ M[i-1,j] + w(x_i, -) \\ M[i,j-1] + w(-, y_j) \end{cases}$$

◦ où $M[i,j]$ représente le score de l'alignement de $S[1..i]$ avec $T[1..j]$ et **optimise** est la fonction :

❖ *maximum* si w est une matrice de score maximisante

❖ *minimum* si w est une matrice de score minimisante

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

◦ Algorithme

1ere étape: calcul des matrices M et Dir

- La matrice des directions est obtenue par les formules suivantes:
- initialisation:
 - $\text{Dir}[0,0] = x$
 - $\text{Dir}[i,0] = \leftarrow$ pour tout i de 1 à N
 - $\text{Dir}[0,j] = \uparrow$ pour tout j de 1 à P

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

○ Algorithme

1ere étape: calcul des matrices M et Dir

- La matrice des directions est obtenue par les formules suivantes :

- calcul des directions :

- $$\text{Dir}[i,j] = \text{Union} = \begin{cases} \nwarrow \text{ si } M[i,j] = M[i-1,j-1] + w(x_i, y_j) \\ \uparrow \text{ si } M[i,j] = M[i-1,j] + w(x_i, -) \\ \leftarrow \text{ si } M[i,j] = M[i,j-1] + w(-, y_j) \end{cases}$$

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

◦ Algorithme

2eme étape: création d'un alignement

- une fois les matrices M et Dir obtenues, pour trouver un alignement, on remonte depuis $M[N,P]$ vers $M[0,0]$, en suivant les directions autorisées (Nord, Ouest, Nord-Ouest) selon la matrice des directions.
- Cela génère un ensemble d'alignements, mais choisir une seule direction à chaque étape permet d'obtenir un alignement spécifique..

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

- Exemple d'application

- Dans cet exemple, on examine l'impact de deux matrices de score (identité et son inverse) sur l'alignement global des séquences:

- $S = \{ \text{CATTGC} \}$

- $T = \{ \text{ACAGTC} \}$

- La matrice d'identité:

- $w(a,a) = 1$

- $w(a,b) = w(a,-) = w(-,a) = 0$

- matrice de score: *maximisante*.

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

Exemple d'application

Matrices									
S/T-	A	C	A	G	T	C	i		
-	0	0	0	0	0	0	0	0	
C	0	0	1	1	1	1	1	1	
A	0	1	1	2	2	2	2	2	
T	0	1	1	2	2	2	2	3	
T	0	1	1	2	2	3	3	4	
G	0	1	1	2	3	3	3	5	
C	0	1	2	2	3	3	4	6	
j	0	1	2	3	4	5	6		
Matrice des scores optimaux <i>M</i>									
S/T-	A	C	A	G	T	C	i		
-	x	←	←	←	←	←	←	0	
C	↑↖	↖	↖	←	←	←	↖	1	
A	↑↖	↖	↖	↖	←	←	←	2	
T	↑	↑	↖	↑	↖	↖	←	3	
T	↑	↑	↖	↑	↖	↖	↖	4	
G	↑	↑	↖	↑	↖	↖	↖	5	
C	↑	↑	↖	↖	↑	↖	↖	6	
j	0	1	2	3	4	5	6		
Matrice des directions <i>Dir</i>									

-CATTG-C	-CA-TTGC
ACA--GTC	ACAGT--C
-CAT-TGC	
ACA-GT-C	
-CATTGC	
ACAGT-C	
-CA-TTGC	
ACAG-T-C	

- Le score $M[6,6]=4$ indique que tout alignement optimal comportera 4 appariements.

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

- Exemple d'application
- La matrice inverse de l'identité:
 - $w(a,a) = 0$
 - $w(a,b) = w(a,-) = w(-,a) = 1$
- matrice de score: *minimisante*/

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

◦ Exemple d'application

Matrices

S/T-	A	C	A	G	T	C	
-	0	1	2	3	4	5	6
C	1	1	1	2	3	4	5
A	2	1	2	1	2	3	4
T	3	2	2	2	2	2	3
T	4	3	3	3	3	2	3
G	5	4	4	4	3	3	3
C	6	5	4	5	4	4	3

Matrice des scores optimaux **M**

S/T-	A	C	A	G	T	C
-	x	←	←	←	←	←
C	↑	↖	↖	←	←	↖
A	↑	↖	↖	↖	←	←
T	↑	↑	↖	↑	↖	←
T	↑	↑	↖	↖	↖	↖
G	↑	↑	↖	↖	↖	↑
C	↑	↑	↖	↖	↑	↖

Matrice des directions **Dir**

-CATTGC

ACAGT-C

- Le score $M[6,6]=3$ indique que tout alignement optimal comportera 3 substitutions ou insertions de gaps. .

ALIGNEMENT PAR PAIRES

PROGRAMMATION DYNAMIQUE (PD)

◦ Exemple d'application

- la matrice de score Identité privilégie les appariements mais peut introduire des gaps de manière inopportune, remédiée par un alignement avec gap affine.
- En revanche, la matrice de score Inverse de l'Identité tend à minimiser l'insertion de gaps.

ALIGNEMENT PAR PAIRES

ALIGNEMENT LOCAL

- L'alignement local $\rightarrow$ trouver le meilleur alignement d'un segment S avec une séquence T.
- Il peut être réalisé par Programmation Dynamique
- Soient deux séquences S et T de longueurs respectives N et P (avec $N < P$).
- La recherche d'un alignement local optimal entre S et T suivant une fonction de score w est obtenu par construction d'une *matrice des scores optimaux d'alignement* $M[0..N, 0..P]$ telle que :

ALIGNEMENT PAR PAIRES

ALIGNEMENT LOCAL

- initialisation :

- $M[0,0] = 0$

- $M[i,0] = 0$ pour tout i de 1 à N

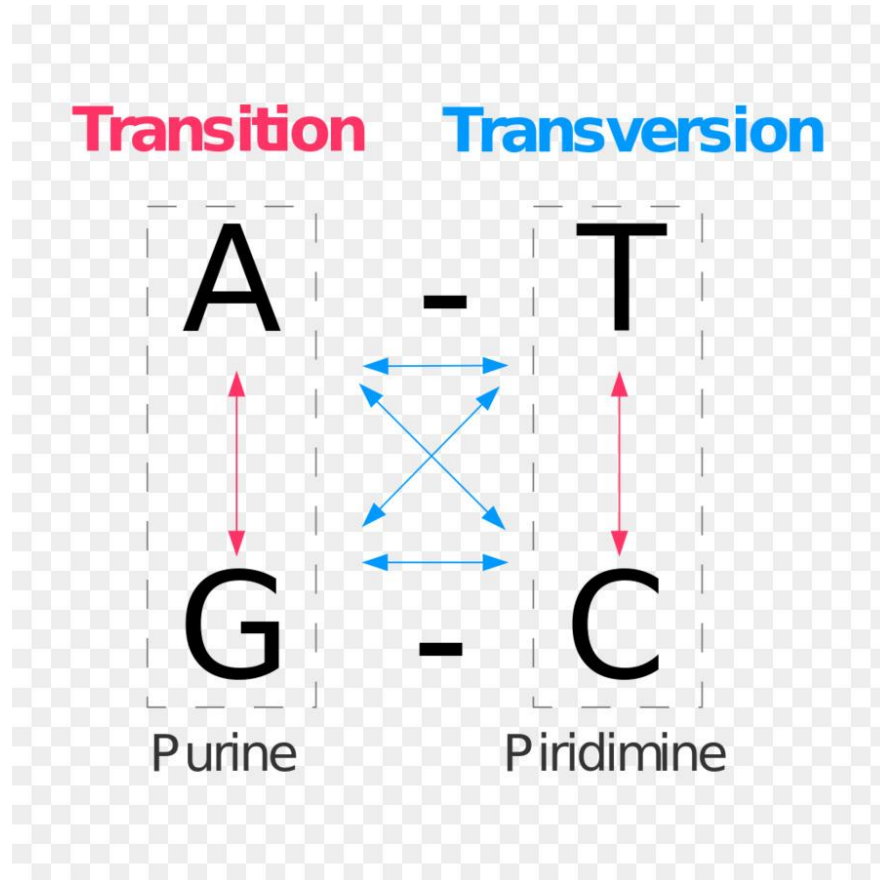
- $M[0,j] = 0$ pour tout j de 1 à P

- calcul du score optimal :

- $$M[i,j] = \begin{cases} M[i-1,j-1] + w(x_i, y_j) \\ M[i-1,j] + w(x_i, -) \\ M[i,j-1] + w(-, y_j) \\ 0 \end{cases}$$

ALIGNEMENT PAR PAIRES

ALIGNEMENT LOCAL



- $w(a,a) = 3$
- $w(a,b) \rightarrow \text{transition} = 1$
- $w(a,b) \rightarrow \text{transver} = 0$
- $w(a,-) = w(-,a) = -1$

ALIGNEMENT PAR PAIRES

ALIGNEMENT LOCAL

- La matrice des directions est obtenue par les mêmes formules que celles définies pour l'alignement global.
- Pour obtenir la position de la séquence T qui correspond au meilleur alignement possible de S avec T, on recherche la position $j_{\max}$ telle que la valeur $M[N, j_{\max}] = \max M[N, j]$ en partant de $j=0$.
- On peut alors créer un alignement local de S avec T en partant de $M[N, j_{\max}]$ jusqu'à obtenir une valeur $M[i, j]=0$.

ALIGNEMENT PAR PAIRES

ALIGNEMENT LOCAL

- Exemple d'application

- Soit S,T deux séquences:

- $S = \{ TGA \}$

- $T = \{ TATGAACTA \}$

- La matrice d'identité:

- $w(a,a) = 1$

- $w(a,b) = w(a,-) = w(-,a) = 0$

ALIGNEMENT PAR PAIRES

ALIGNEMENT LOCAL

Exemple d'application

Matrices											
S/T-	T	A	T	G	C	A	C	T	A	i	
-	0	0	0	0	0	0	0	0	0	0	0
T	0	1	1	1	1	1	1	1	1	1	1
G	0	1	1	1	2	2	2	2	2	2	2
A	0	1	1	1	2	2	3	3	3	3	3
j	0	1	2	3	4	5	6	7	8	9	
Matrice des scores optimaux M											
S/T-	T	A	T	G	C	A	C	T	A	i	
-	x	x	x	x	x	x	x	x	x	x	0
T	x	↖	←	↖	←	←	←	←	↖	←	1
G	x	↑	↗	↗	↖	←	←	←	←	←	2
A	x	↑	↖	←	←	←	↖	←	←	←	3
j	0	1	2	3	4	5	6	7	8	9	
Matrice des directions Dir											

TGA

TCA

TATGA

T--GA

- On cherche la valeur $j_{\max}$ telle que $M[3, j_{\max}] = \max_{j=0} M[3, j]$. On trouve $j_{\max} = 6$ et $M[3, j_{\max}] = 3$