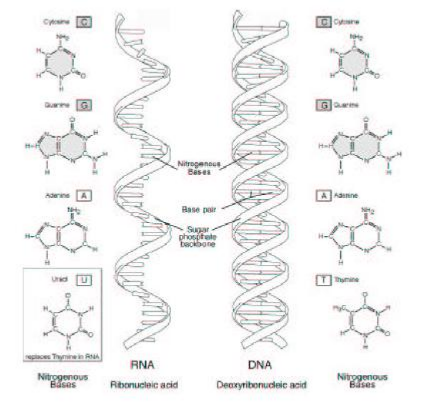
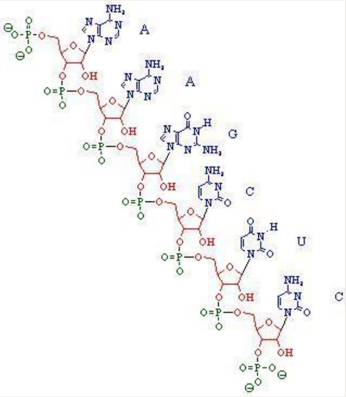
Un aperçu des méthodes de prédiction de structure secondaire d’ARN.

Alain Denise 24/01/2011
Bioinformatique
IUT Informatique, Orsay

Généralités

L'ARN



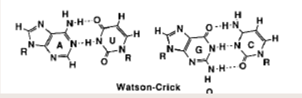


…AAGCUC…

Pourquoi l'ARN ?

- La seule molécule qui peut être génome aussi bien que catalyseur
- Origine de la vie : le monde à ARN
- Présente dans tous les processus cellulaires
- Cible très fréquente des antibiotiques

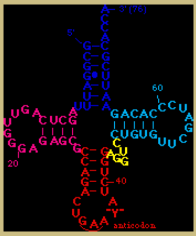
Structure de l’



- Structure primaire

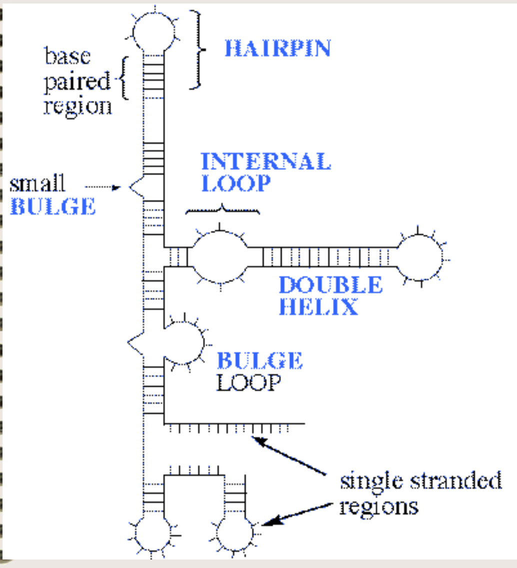
GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAUAUCUGGAGGUCCUGUGUUCGAUCCCACAGAAUUCGCACCA

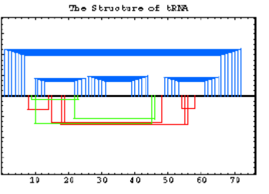
- Structure secondaire

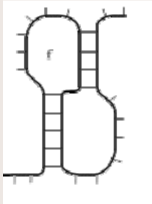


- Structure tertiaire

Structures secondaires

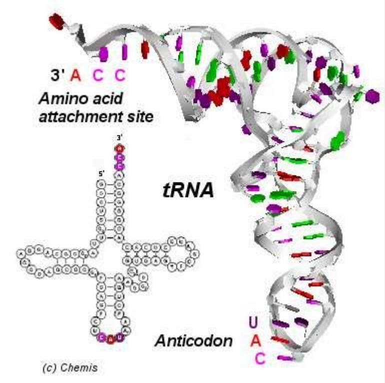


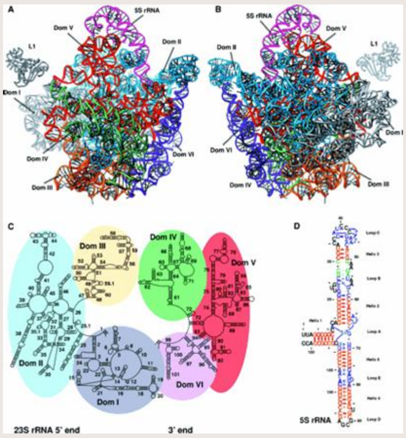




Pseudo-nœud

Structure de l’ARN





Problèmes informatiques

- Prédiction de structure en fonction de la séquence
- Détermination d’une séquence en fonction de la structure
- Détection de motifs structurels dans une séquence ou dans une structure
- Comparaison de deux ou plusieurs structures
- Recherche de sous-structures communes à deux ou plusieurs structures

Niveaux de structure de l’ARN

- primaire
- secondaire
- secondaire avec pseudo-noeuds
- tertiaire
- 3D

Deux types d’approches

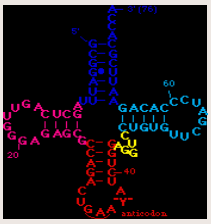
- Approches ab initio : une séquence
- Approches comparatives : plusieurs séquences
  - avec alignement préalable
  - sans alignement préalable

Approche ab initio

Prédiction de structure d’ARN

Séquence → Structure de plus faible énergie libre
(selon un modèle d’énergie donné)

GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAUAUCUGGAGGUCCUGUGUUCGAUCCCACAGAAUUCGCACCA



Modèle thermodynamique

10.2 RNA secondary structure prediction

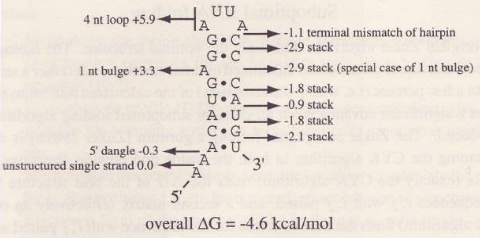


Figure 10.10 *An example ΔG calculation for an RNA stem loop (the wild type R17 coat protein binding site).*

Pour des structures secondaires sans pseudo-noeuds

Biological sequence analysis
Durbin, Eddy, Krogh, Mitchison
Cambridge Univ. Press 1998

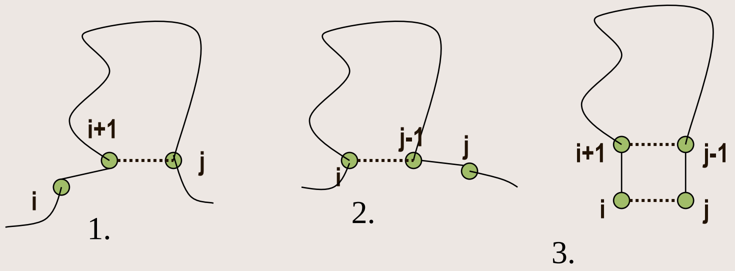
Algorithme pour un modèle plus simple

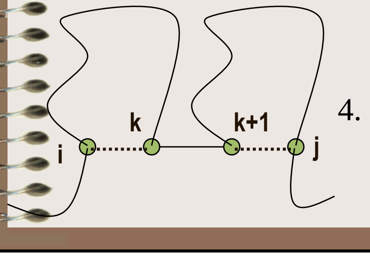
Prédiction de structure
- sans pseudo-nœud,
- en maximisant le nombre de nucléotides appariés.

$\gamma(i,j)$ = nombre d’appariements dans la structure entre les $i^{\text{ème}}$ et $j^{\text{ème}}$ nucléotides.

$\delta(i,j)$ = 1 si i et j sont appariables, 0 sinon.

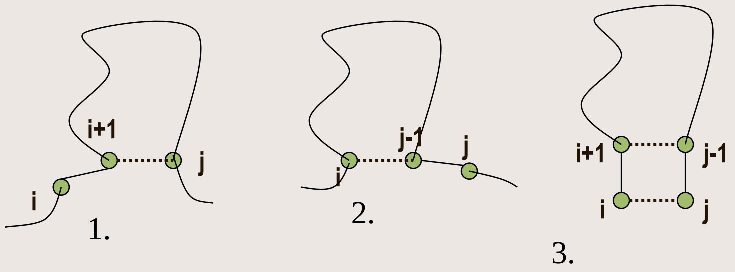
Algorithme de Nussinov (1978)

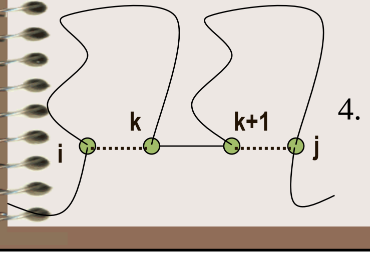




Les 4 façons possibles de construire la meilleure structure entre i et j, connaissant les meilleures sous-structures entre i+1 et j-1.

Algorithme de Nussinov (1978)

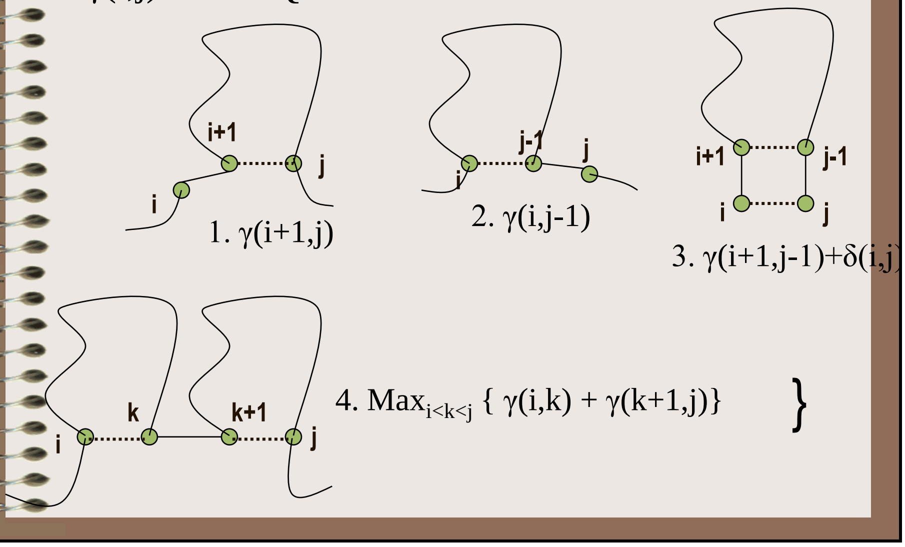




Attention, cette décomposition est ambiguë : plusieurs cas différents peuvent être valables simultanément. Il existe des décompositions non ambiguës pour le même problème

Algorithme de Nussinov (1978)

$\gamma(i,j) = \text{Max} \{$



1. $\gamma(i+1,j)$

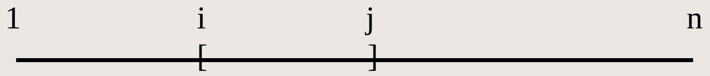
2. $\gamma(i,j-1)$

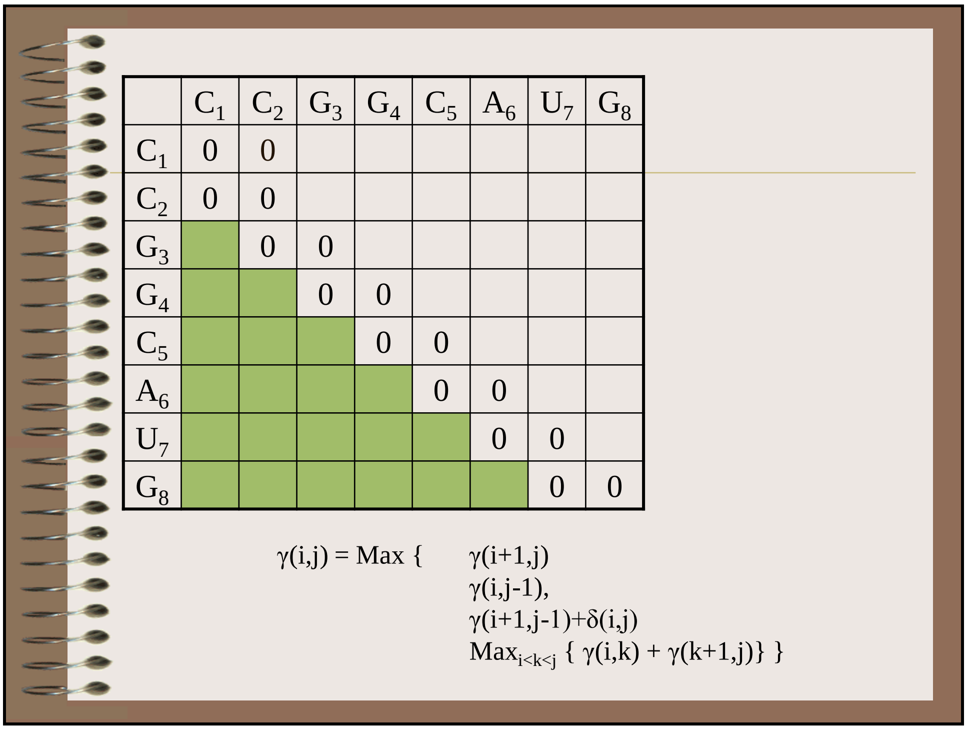
3. $\gamma(i+1,j-1)+\delta(i,j)$

4. $\text{Max}_{i<k<j} \{ \gamma(i,k) + \gamma(k+1,j)\}$ $\}$

Principe : programmation dynamique

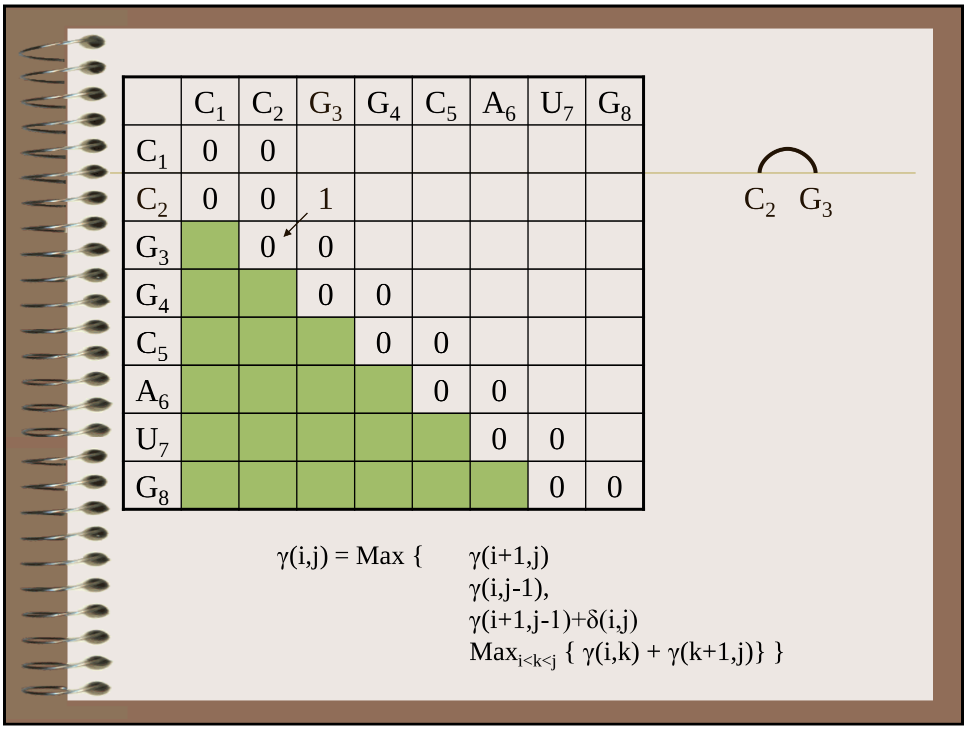
$\gamma(i,j) = \text{Max} \{$ $\gamma(i+1,j)$
$\gamma(i,j-1),$
$\gamma(i+1,j-1)+\delta(i,j)$
$\text{Max}_{i<k<j} \{ \gamma(i,k) + \gamma(k+1,j)\} \}$





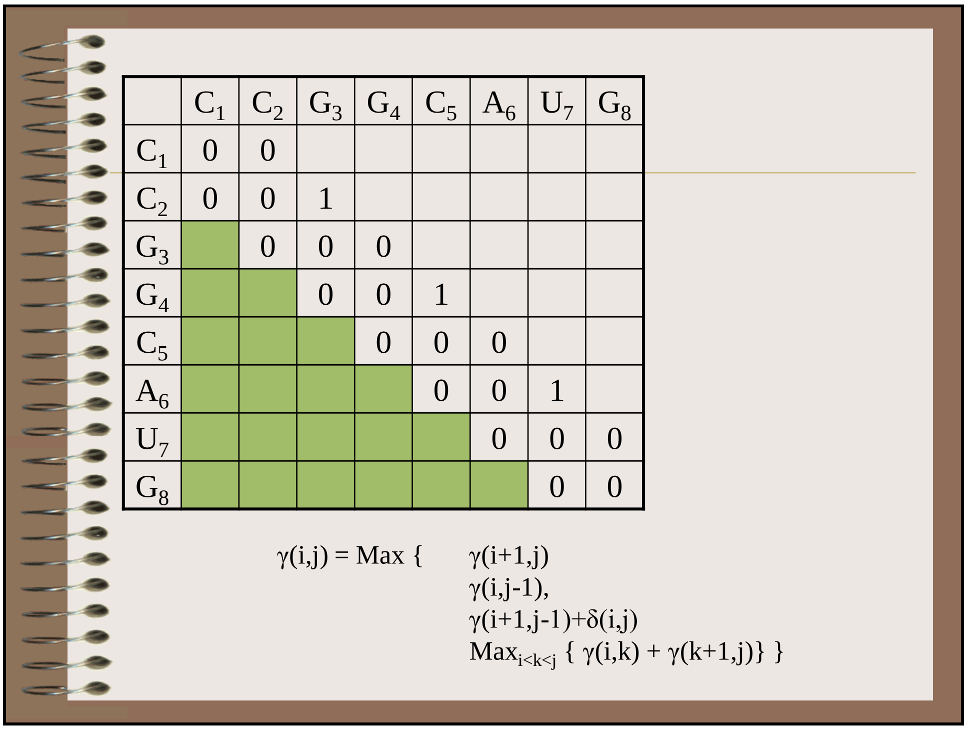
	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0						
C_2	0	0						
G_3	🟩	0	0					
G_4	🟩	🟩	0	0				
C_5	🟩	🟩	🟩	0	0			
A_6	🟩	🟩	🟩	🟩	0	0		
U_7	🟩	🟩	🟩	🟩	🟩	0	0	
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

$$\gamma(i,j) = \text{Max} \{ \quad \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \text{Max}_{i<k<j} \{ \gamma(i,k) + \gamma(k+1,j)\} \} \end{array}$$



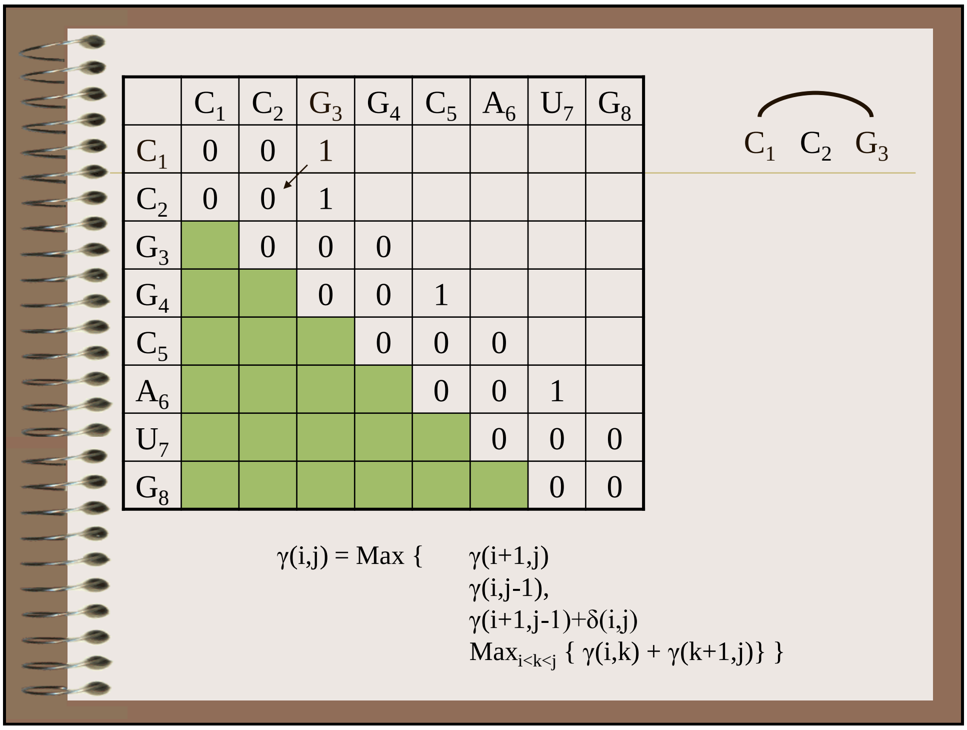
	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0						
C_2	0	0	1					
G_3	🟩	0	0					
G_4	🟩	🟩	0	0				
C_5	🟩	🟩	🟩	0	0			
A_6	🟩	🟩	🟩	🟩	0	0		
U_7	🟩	🟩	🟩	🟩	🟩	0	0	
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

$$\gamma(i,j) = \text{Max} \{ \quad \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \text{Max}_{i<k<j} \{ \gamma(i,k) + \gamma(k+1,j)\} \} \end{array}$$



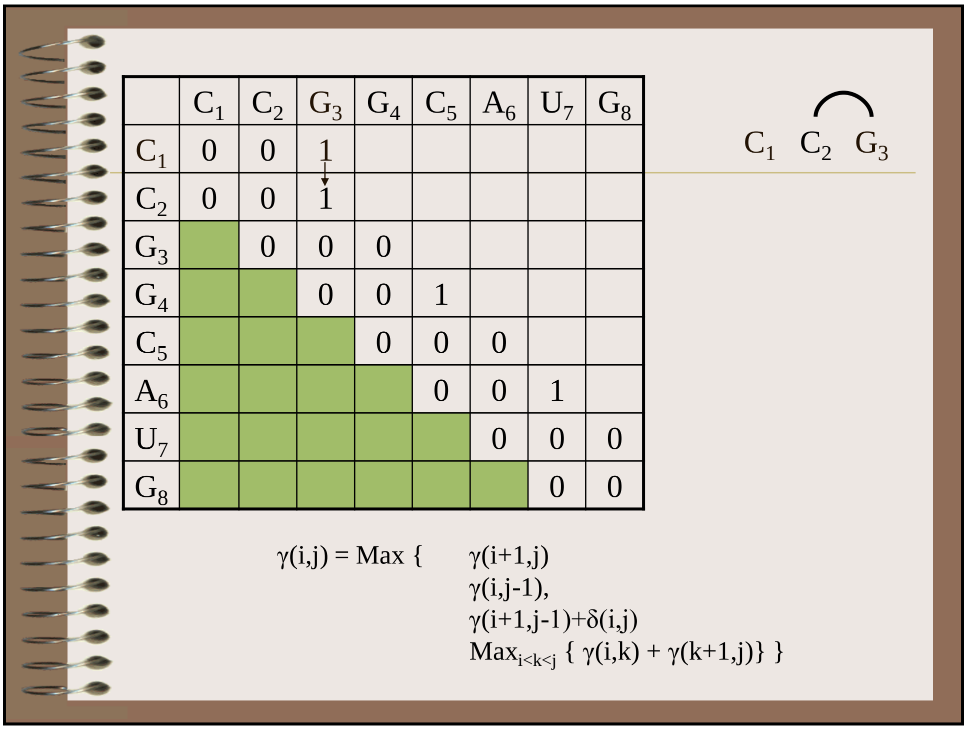
	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0						
C_2	0	0	1					
G_3	🟩	0	0	0				
G_4	🟩	🟩	0	0	1			
C_5	🟩	🟩	🟩	0	0	0		
A_6	🟩	🟩	🟩	🟩	0	0	1	
U_7	🟩	🟩	🟩	🟩	🟩	0	0	0
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

$$\gamma(i,j) = \mathrm{Max}\ \{ \quad \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \mathrm{Max}_{i<k<j}\ \{ \gamma(i,k) + \gamma(k+1,j)\}\ \} \end{array}$$



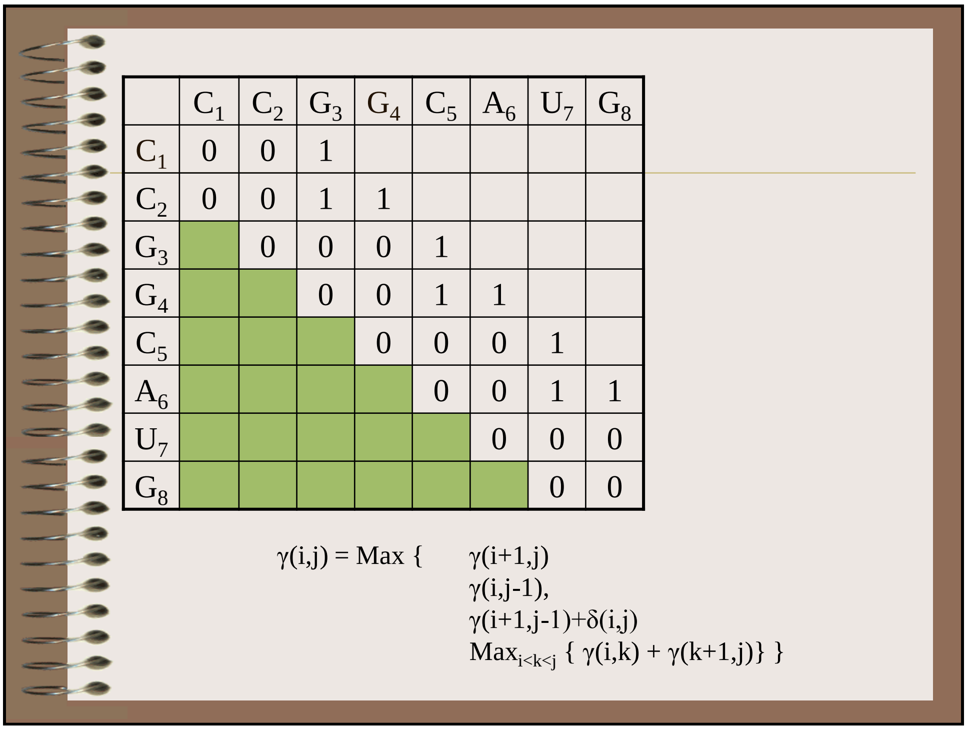
	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0	1					
C_2	0	0	1					
G_3	🟩	0	0	0				
G_4	🟩	🟩	0	0	1			
C_5	🟩	🟩	🟩	0	0	0		
A_6	🟩	🟩	🟩	🟩	0	0	1	
U_7	🟩	🟩	🟩	🟩	🟩	0	0	0
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

$$\gamma(i,j) = \mathrm{Max}\ \{ \quad \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \mathrm{Max}_{i<k<j}\ \{ \gamma(i,k) + \gamma(k+1,j)\}\ \} \end{array}$$



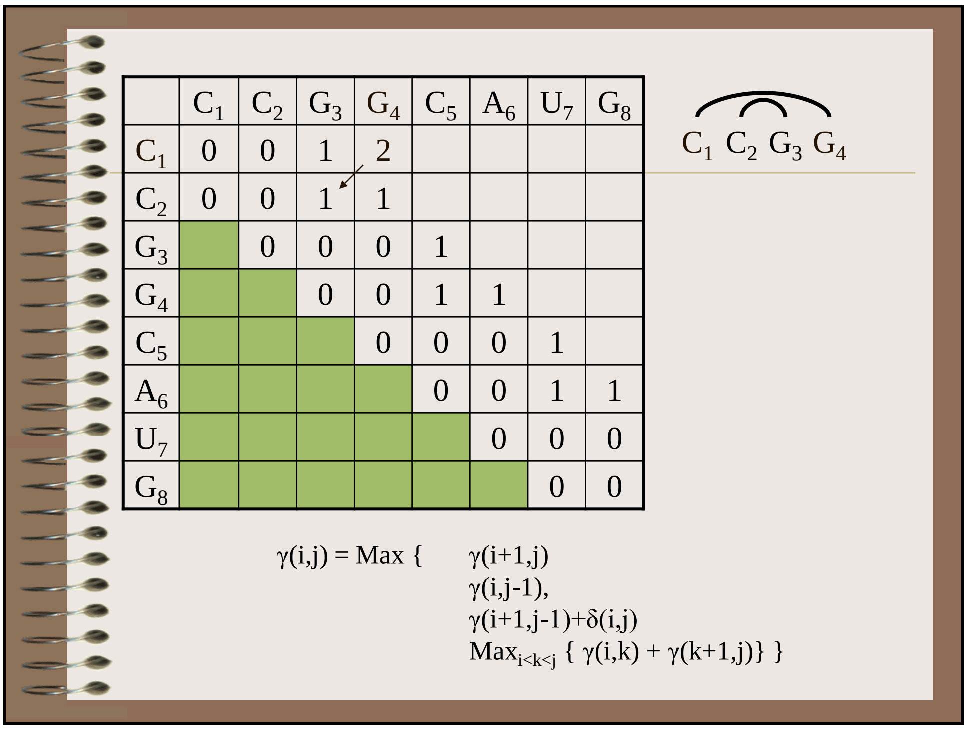
	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0	1					
C_2	0	0	1					
G_3	🟩	0	0	0				
G_4	🟩	🟩	0	0	1			
C_5	🟩	🟩	🟩	0	0	0		
A_6	🟩	🟩	🟩	🟩	0	0	1	
U_7	🟩	🟩	🟩	🟩	🟩	0	0	0
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

$$\gamma(i,j) = \mathrm{Max}\ \{ \quad \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \mathrm{Max}_{i<k<j}\ \{ \gamma(i,k) + \gamma(k+1,j)\}\ \} \end{array}$$



	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0	1					
C_2	0	0	1	1				
G_3	🟩	0	0	0	1			
G_4	🟩	🟩	0	0	1	1		
C_5	🟩	🟩	🟩	0	0	0	1	
A_6	🟩	🟩	🟩	🟩	0	0	1	1
U_7	🟩	🟩	🟩	🟩	🟩	0	0	0
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

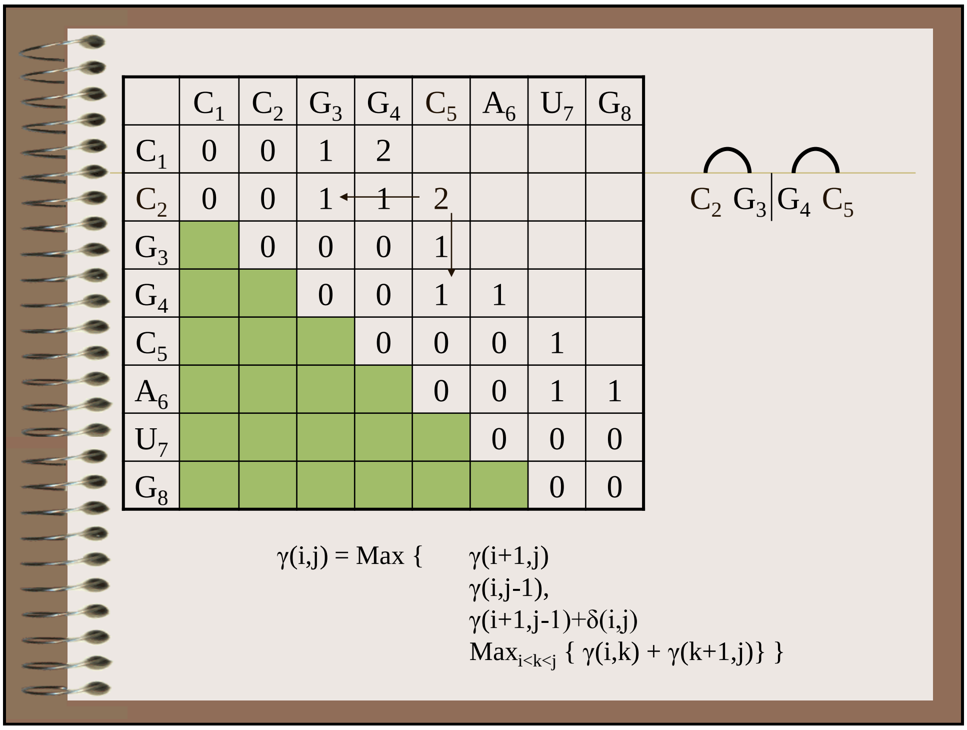
$$\gamma(i,j) = \mathrm{Max}\ \{ \quad \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \mathrm{Max}_{i<k<j}\ \{ \gamma(i,k) + \gamma(k+1,j)\}\ \} \end{array}$$



	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0	1	2				
C_2	0	0	1	1				
G_3	🟩	0	0	0	1			
G_4	🟩	🟩	0	0	1	1		
C_5	🟩	🟩	🟩	0	0	0	1	
A_6	🟩	🟩	🟩	🟩	0	0	1	1
U_7	🟩	🟩	🟩	🟩	🟩	0	0	0
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

C_1 C_2 G_3 G_4

γ(i,j) = Max { γ(i+1,j)
γ(i,j-1),
γ(i+1,j-1)+δ(i,j)
$\mathrm{Max}_{i<k<j}$ { γ(i,k) + γ(k+1,j)} }



	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0	1	2				
C_2	0	0	1	1	2			
G_3	🟩	0	0	0	1			
G_4	🟩	🟩	0	0	1	1		
C_5	🟩	🟩	🟩	0	0	0	1	
A_6	🟩	🟩	🟩	🟩	0	0	1	1
U_7	🟩	🟩	🟩	🟩	🟩	0	0	0
G_8	🟩	🟩	🟩	🟩	🟩	🟩	0	0

C_2 G_3 | G_4 C_5

γ(i,j) = Max { γ(i+1,j)
γ(i,j-1),
γ(i+1,j-1)+δ(i,j)
$\mathrm{Max}_{i<k<j}$ { γ(i,k) + γ(k+1,j)} }

	C_1	C_2	G_3	G_4	C_5	A_6	U_7	G_8
C_1	0	0	1	2	2	2	3	4
C_2	0	0	1	1	2	2	3	3
G_3		0	0	0	1	1	2	2
G_4			0	0	1	1	2	2
C_5				0	0	0	1	2
A_6					0	0	1	1
U_7						0	0	0
G_8							0	0

$$\gamma(i,j) = \text{Max}\ \{\ \begin{array}{l} \gamma(i+1,j) \\ \gamma(i,j-1), \\ \gamma(i+1,j-1)+\delta(i,j) \\ \text{Max}_{i<k<j}\ \{\ \gamma(i,k) + \gamma(k+1,j)\}\ \} \end{array}$$

Algorithme de Zuker-Stiegler (1981)

- Modèle d’énergie plus réaliste (paramètres de Turner *et al.*)
- Même principe (programmation dynamique), avec deux matrices au lieu d’une.
- Structures secondaires toujours sans pseudonoeuds.

Un essai : RNAfold (algo Zuker-Stiegler)

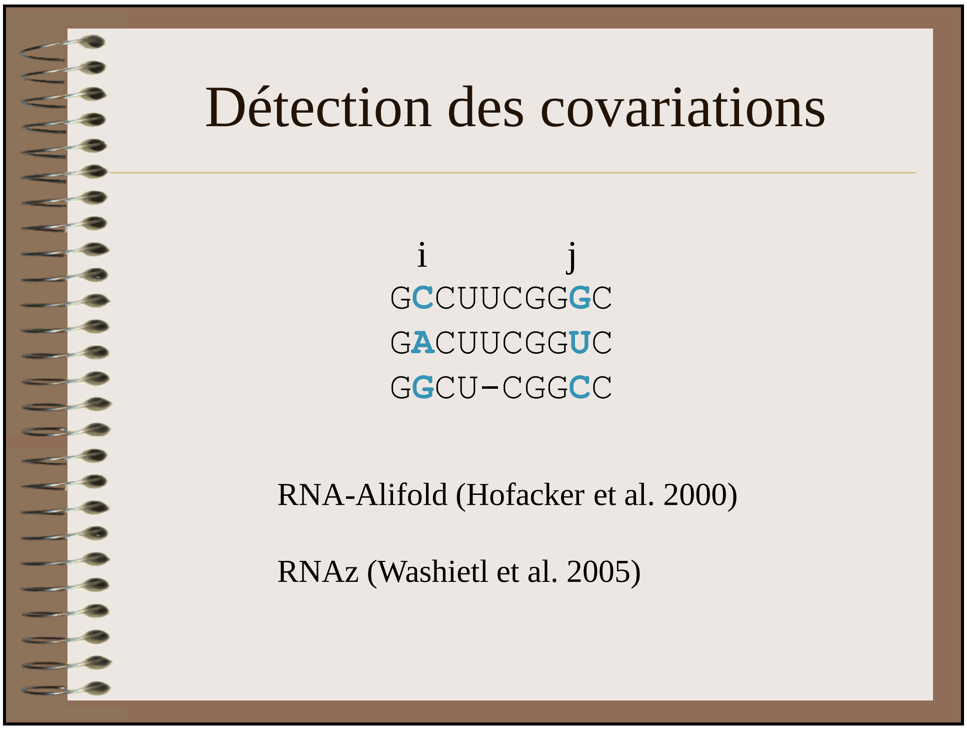
```
>Artibeus jamaicensis, True Structure, tRNA Alanine
AAGGGCTTAGCTTAATTAAAGTAGTTGATTTGCATTCAGCAGCTGTAGGATAAAGTCTTGCAGTCCTTA
(((((((..((((.....)))).(((((.......))))).....(((((.......)))))))))))).
>Artibeus jamaicensis, RNAFold
AAGGGCTTAGCTTAATTAAAGTAGTTGATTTGCATTCAGCAGCTGTAGGATAAAGTCTTGCAGTCCTTA
(((((((.(((((.(((..(((.(((((.......))))).)))....))).))).))...))))))).

>Balaenoptera musculus, True Structure
GAGGATTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATATAGTCTTGCAGTCCTTA
(((((((..((((.....))))..((((.......)))).....(((((.......)))))))))))).
>Balaenoptera musculus, RNAFold
GAGGATTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATATAGTCTTGCAGTCCTTA
((((((...((..((((((.....)))))).))..........((((((((...)))))))))))))).

>Bos taurus, True Structure
GAGGATTTAGCTTAATTAAAGTGGTTGATTTGCATTCAATTGATGTAAGGTGTAGTCTTGCAATCCTTA
(((((((..((((.....)))).(((((.......))))).....(((((.......)))))))))))).
>Bos taurus, RNAFold
GAGGATTTAGCTTAATTAAAGTGGTTGATTTGCATTCAATTGATGTAAGGTGTAGTCTTGCAATCCTTA
(((((....((..((((((.....)))))).))..........(((((((.....))))))).))))).
```

Approche comparative

Avec alignement préalable des séquences



Détection des covariations

i j

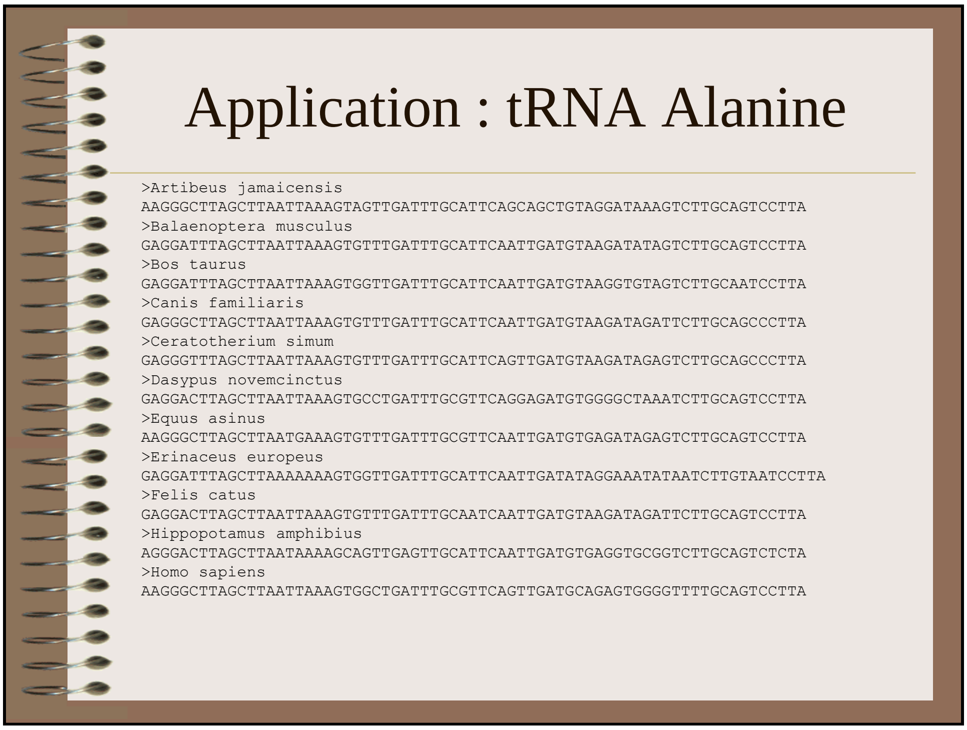
G**C**CUUCGG**G**C

G**A**CUUCGG**U**C

G**G**CU-CGG**C**C

RNA-Alifold (Hofacker et al. 2000)

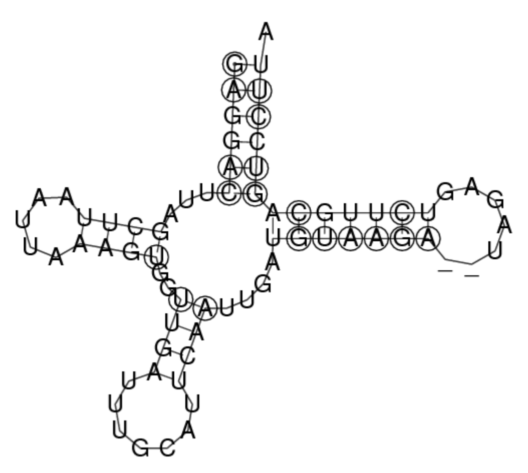
RNAz (Washietl et al. 2005)



Application : tRNA Alanine

```
>Artibeus jamaicensis
AAGGGCTTAGCTTAATTAAAGTAGTTGATTTGCATTCAGCAGCTGTAGGATAAAGTCTTGCAGTCCTTA
>Balaenoptera musculus
GAGGATTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATATAGTCTTGCAGTCCTTA
>Bos taurus
GAGGATTTAGCTTAATTAAAGTGGTTGATTTGCATTCAATTGATGTAAGGTGTAGTCTTGCAATCCTTA
>Canis familiaris
GAGGGCTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATAGATTCTTGCAGCCCTTA
>Ceratotherium simum
GAGGGTTTAGCTTAATTAAAGTGTTTGATTTGCATTCAGTTGATGTAAGATAGAGTCTTGCAGCCCTTA
>Dasypus novemcinctus
GAGGACTTAGCTTAATTAAAGTGCCTGATTTGCGTTCAGGAGATGTGGGGCTAAATCTTGCAGTCCTTA
>Equus asinus
AAGGGCTTAGCTTAATGAAAGTGTTTGATTTGCGTTCAATTGATGTGAGATAGAGTCTTGCAGTCCTTA
>Erinaceus europeus
GAGGATTTAGCTTAAAAAAAGTGGTTGATTTGCATTCAATTGATATAGGAAATATAATCTTGTAATCCTTA
>Felis catus
GAGGACTTAGCTTAATTAAAGTGTTTGATTTGCAATCAATTGATGTAAGATAGATTCTTGCAGTCCTTA
>Hippopotamus amphibius
AGGGACTTAGCTTAATAAAAGCAGTTGAGTTGCATTCAATTGATGTGAGGTGCGGTCTTGCAGTCTCTA
>Homo sapiens
AAGGGCTTAGCTTAATTAAAGTGGCTGATTTGCGTTCAGTTGATGCAGAGTGGGGTTTTGCAGTCCTTA
```

RNA-alifold



Application : tRNA *H.sapiens*

```
>Homo sapiens Arg, True Structure
TGGTATATAGTTTAAACAAAACGAATGATTTCGACTCATTAAATTATGATAATCATATTTACCAA
(((((.(..((((.....)))).(((((.......))))).....(((((...)))))).))))).

>Homo sapiensArg
TGGTATATAGTTTAAACAAAACGAATGATTTCGACTCATTAAATTATGATAATCATATTTACCAA
>Homo sapiensAsn
TAGATTGAAGCCAGTTGATTAGGGTGCTTAGCTGTTAACTAAGTGTTTGTGGGTTTAAGTCCCATTGGTCTAG
>Homo sapiensAsp
AAGGTATTAGAAAAACCATTTCATAACTTTGTCAAAGTTAAATTATAGGCTAAATCCTATATATCTTA
>Homo sapiensCys
AGCTCCGAGGTGATTTTCATATTGAATTGCAAATTCGAAGAAGCAGCTTCAAACCTGCCGGGGCTT
>Homo sapiensGln
TAGGATGGGGTGTGATAGGTGGCACGGAGAATTTTGGATTCTCAGGGATGGGTTCGATTCTCATAGTCCTAG
>Homo sapiensGlu
GTTCTTGTAGTTGAAATACAACGATGGTTTTTCATATCATTGGTCGTGGTTGTAGTCCGTGCGAGAATA
>Homo sapiensGly
ACTCTTTTAGTATAAATAGTACCGTTAACTTCCAATTAACTAGTTTTGACAACATTCAAAAAAGAGTA
>Homo sapiensHis
GTAAATATAGTTTAACCAAAACATCAGATTGTGAATCTGACAACAGAGGCTTACGACCCCTTATTTACC
>Homo sapiensIso
AGAAATATGTCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAGGAGCTTAAACCCCCTTATTTCTA
>Homo sapiensLeuCun
ACTTTTAAAGGATAACAGCTATCCATTGGTCTTAGGCCCCAAAAATTTTGGTGCAACTCCAAATAAAAGTA
```

Alignement ClustalW

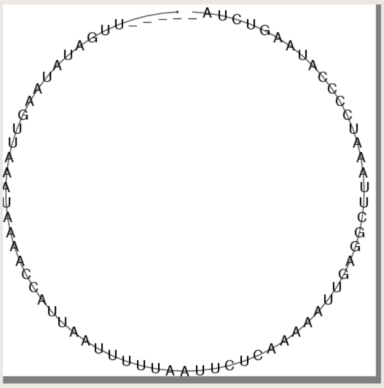
```
CLUSTAL 2.0.1 multiple sequence alignment

Homo1 ---TGGTATATAGTTTAAACAAAACGAATGATTTCGACTCATTAAATTATGATAATCATA 57
Homo7 ---ACTCTTTTAGTATAAATAGTACCGTTAACTTCCAATTAACTAGTTTTGACAAC---A 54
Homo3 ----AAGGTATTAGAAAAACCATTTCATAACTTTGTCAAAGTTAA-ATTATAGGCTAA--  53
Homo8 ---GTAAATAT-AGTTTAACCAAAACATCAGATTGTGAATCTGAC-AACAGAGGCTTACG 55
Homo9 -----AGAAATATGTCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAGGAGCTTAAA 55
Homo6 ---GTTCTTGTAGTTGAAATACAACGATGGTTTTTCA--TATCATTGGTCGTGGTTGTAG 55
Homo5 TAGGATGGGGTGTGATAGGTGGCACGGAGAATTTTGGATTCTCAGGGAT---GGGTTCGA 57
Homo2 TAGATTGAAGCCAGTTGATTAGGGTGCTTAGCTGTTAA--CTAAGTGTTTGTGGGTTTAA 58
Homo10 ---ACTTTTAAAGGATAACAGCTATCCATTGGTCTTAGGCCCCAAAAATTTTGGTGCAAC 57
Homo4 -----AGCTCCGAGGTGATTTTCATATTGAATTGCAAATTCGAAGAAGC---AGCTTCAA 52

*


Homo1 TTTACCAA-------  65
Homo7 TTCAAAAAAGAGTA-  68
Homo3 ATCCTATATATCTTA  68
Homo8 ACCCCTTATTTACC-  69
Homo9 CCCCCTTATTTCTA-  69
Homo6 TCCGTGCGAGAATA-  69
Homo5 TTCTCATAGTCCTAG  72
Homo2 GTCCCATTGGTCTAG  73
Homo10 TCCAAATAAAAGTA-  71
Homo4 ACCTGCCGGGGCTT-  66
```

RNA-Alifold



Approche comparative

Sans alignement préalable des séquences

Approches

- Algorithmes « à la Sankoff » (Sankoff 1985)
  - Foldalign (Gorodkin, Heyer, Stormo 1997, Havgaard, Lyngso, Stormo, Gorodkin 2005)
  - Dynalign (Mathews, Turner 2002)
- Approches heuristiques :
  - Exemple : CARNAC (Perriquet, Touzet, Dauchet 2003)

Pincipe de l’algo de Sankoff : prog. dynamique (encore!)

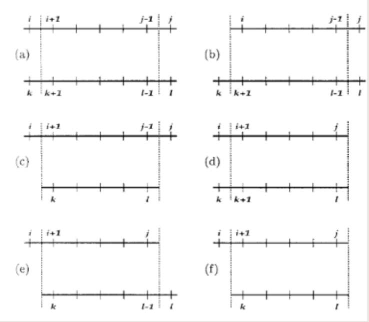
- Entrée : plusieurs séquences
- Paramètres : matrice de score d’alignement de couples de nucléotides, $S_{ij,kl}$
- Sortie : un alignement et une structure secondaire commune

Algo simplifié : Foldalign 1997

- Pas de structures multibranchées
- Le score de la structure dépend du nombre de paires de bases.

Algo simplifié : Foldalign 1997

$$D_{ij,kl} = \max \begin{cases} D_{(i+1)(j-1),(k+1)(l-1)} + S_{ij,kl}, & (a) \\ \\ D_{i(j-1),(k+1)(l-1)} + S_{-j,kl}, & \\ D_{(i+1)j,(k+1)(l-1)} + S_{i-,kl}, & (b) \\ D_{(i+1)(j-1),k(l-1)} + S_{ij,-l}, & \\ D_{(i+1)(j-1),(k+1)l} + S_{ij,k-}, & \\ \\ D_{(i+1)(j-1),kl} + S_{ij,--}, & (c) \\ D_{ij,(k+1)(l-1)} + S_{--,kl}, & \\ \\ D_{(i+1)j,(k+1)l} + S_{i-,k-}, & (d) \\ D_{i(j-1),k(l-1)} + S_{-j,-l}, & \\ \\ D_{(i+1)j,k(l-1)} + S_{i-,-l}, & (e) \\ D_{i(j-1),(k+1)l} + S_{-j,k-}, & \\ \\ D_{(i+1)j,kl} + S_{i-,--}, & \\ D_{i(j-1),kl} + S_{-j,--}, & (f) \\ D_{ij,(k+1)l} + S_{--,k-}, & \\ D_{ij,k(l-1)} + S_{--,-l} & \end{cases}$$



Une approche heuristique : Carnac

- Recherche des tiges-boucles candidates dans chaque séquence, indépendamment
- Recherche de « points d'ancrage » : régions très conservées entre les 2 séquences
- Sélection des tiges « alignables »
- Repliement simultané « à la Sankoff » de chaque paire de tiges alignables

Application : tRNA Alanine

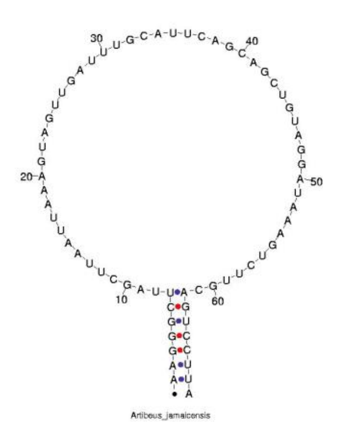
```
>Artibeus jamaicensis
AAGGGCTTAGCTTAATTAAAGTAGTTGATTTGCATTCAGCAGCTGTAGGATAAAGTCTTGCAGTCCTTA
>Balaenoptera musculus
GAGGATTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATATAGTCTTGCAGTCCTTA
>Bos taurus
GAGGATTTAGCTTAATTAAAGTGGTTGATTTGCATTCAATTGATGTAAGGTGTAGTCTTGCAATCCTTA
>Canis familiaris
GAGGGCTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATAGATTCTTGCAGCCCTTA
>Ceratotherium simum
GAGGGTTTAGCTTAATTAAAGTGTTTGATTTGCATTCAGTTGATGTAAGATAGAGTCTTGCAGCCCTTA
>Dasypus novemcinctus
GAGGACTTAGCTTAATTAAAGTGCCTGATTTGCGTTCAGGAGATGTGGGGCTAAATCTTGCAGTCCTTA
>Equus asinus
AAGGGCTTAGCTTAATGAAAGTGTTTGATTTGCGTTCAATTGATGTGAGATAGAGTCTTGCAGTCCTTA
>Erinaceus europeus
GAGGATTTAGCTTAAAAAAAGTGGTTGATTTGCATTCAATTGATATAGGAAATATAATCTTGTAATCCTTA
>Felis catus
GAGGACTTAGCTTAATTAAAGTGTTTGATTTGCAATCAATTGATGTAAGATAGATTCTTGCAGTCCTTA
>Hippopotamus amphibius
AGGGACTTAGCTTAATAAAAGCAGTTGAGTTGCATTCAATTGATGTGAGGTGCGGTCTTGCAGTCTCTA
>Homo sapiens
AAGGGCTTAGCTTAATTAAAGTGGCTGATTTGCGTTCAGTTGATGCAGAGTGGGGTTTTGCAGTCCTTA
```

Foldalign : pas plus de 2 séquences

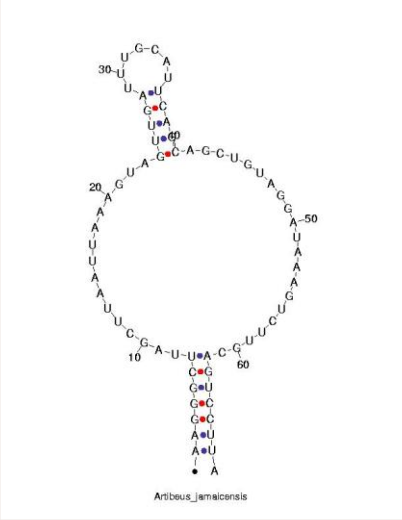
```
>Artibeus jamaicensis
AAGGGCTTAGCTTAATTAAAGTAGTTGATTTGCATTCAGCAGCTGTAGGATAAAGTCTTGCAGTCCTTA
>Balaenoptera musculus
GAGGATTTAGCTTAATTAAAGTGTTTGATTTGCATTCAATTGATGTAAGATATAGTCTTGCAGTCCTTA
```

```
Artibeus
AGGGCUUAGCUUAAUUAAAGUAGUUGAUUUGCAUUCAGCAGCUGUAGGAUAAAGUCUUGCAGUCCUUA
Structure
(((((...((((.....))))...(((.......)))))...(((((((.....)))))))..)))..
Balaenoptera
AGGAUUUAGCUUAAUUAAAGUGUUUGAUUUGCAUUCAAUUGAUGUAAGAUAUAGUCUUGCAGUCCUUA
```

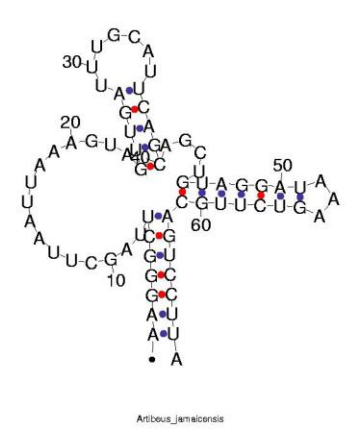
Carnac - 3 séquences



Carnac - 6 séquences



Carnac - 11 séquences

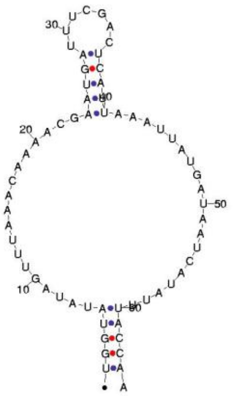


Application : tRNA *H.sapiens Arg*

```
>Homo sapiens Arg, True Structure
TGGTATATAGTTTAAACAAAACGAATGATTTCGACTCATTAAATTATGATAATCATATTTACCAA
(((((.(..((((.....)))).(((((.......)))))....(((((...)))))).))))).

>Homo sapiensArg
TGGTATATAGTTTAAACAAAACGAATGATTTCGACTCATTAAATTATGATAATCATATTTACCAA
>Homo sapiensAsn
TAGATTGAAGCCAGTTGATTAGGGTGCTTAGCTGTTAACTAAGTGTTTGTGGGTTTAAGTCCCATTGGTCTAG
>Homo sapiensAsp
AAGGTATTAGAAAAACCATTTCATAACTTTGTCAAAGTTAAATTATAGGCTAAATCCTATATATCTTA
>Homo sapiensCys
AGCTCCGAGGTGATTTTCATATTGAATTGCAAATTCGAAGAAGCAGCTTCAAACCTGCCGGGGCTT
>Homo sapiensGln
TAGGATGGGGTGTGATAGGTGGCACGGAGAATTTTGGATTCTCAGGGATGGGTTCGATTCTCATAGTCCTAG
>Homo sapiensGlu
GTTCTTGTAGTTGAAATACAACGATGGTTTTTCATATCATTGGTCGTGGTTGTAGTCCGTGCGAGAATA
>Homo sapiensGly
ACTCTTTTAGTATAAATAGTACCGTTAACTTCCAATTAACTAGTTTTGACAACATTCAAAAAAGAGTA
>Homo sapiensHis
GTAAATATAGTTTAACCAAAACATCAGATTGTGAATCTGACAACAGAGGCTTACGACCCCTTATTTACC
>Homo sapiensIso
AGAAATATGTCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAGGAGCTTAAACCCCCTTATTTCTA
>Homo sapiensLeuCun
ACTTTTAAAGGATAACAGCTATCCATTGGTCTTAGGCCCCAAAAATTTTGGTGCAACTCCAAATAAAAGTA
```

Carnac - 10 séquences



Homo_sapiens_arg._RNAFold