




ClustalLite : Alignement multiple heuristique par la méthode Clustal

Stéphanie Gnanalingam

M2 Bio-informatique parcours biologie informatique

Développement

- <https://github.com/gnanalin/ClustalLite/tree/main>
- CONDA (v.24.7.1)  env_clustallite

Applications

Outil d'alignement
multiple

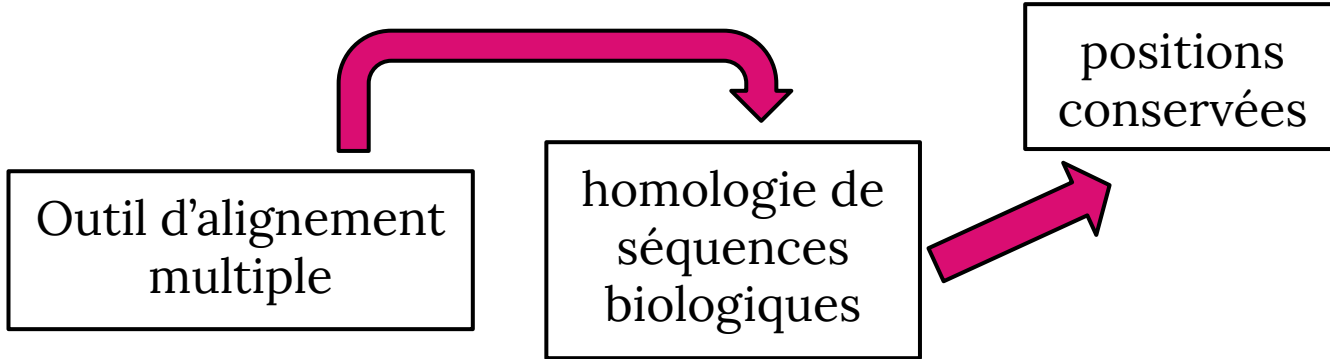
Applications



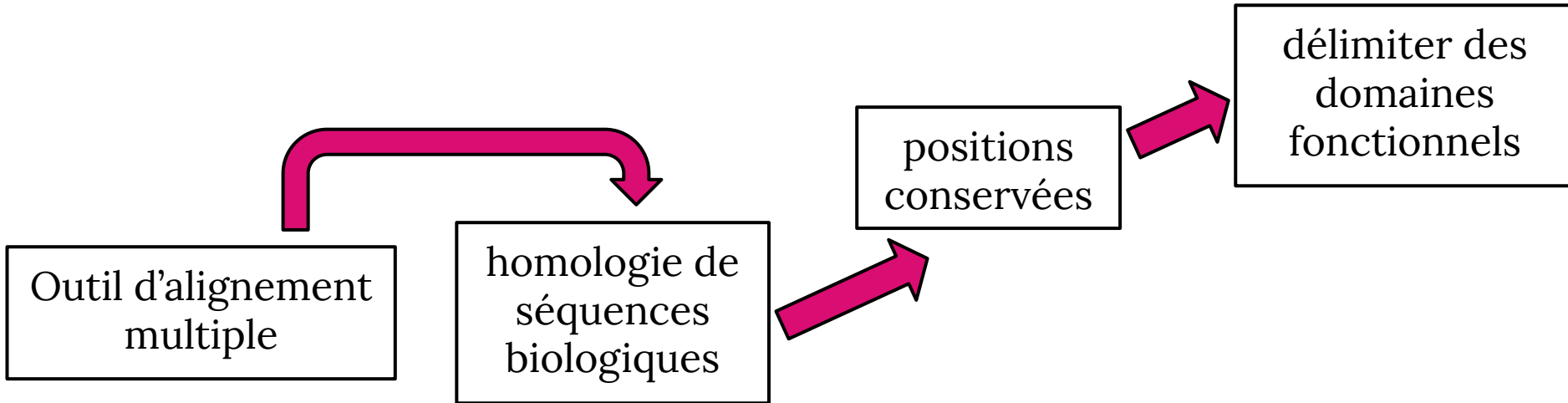
Outil d'alignement
multiple

homologie de
séquences
biologiques

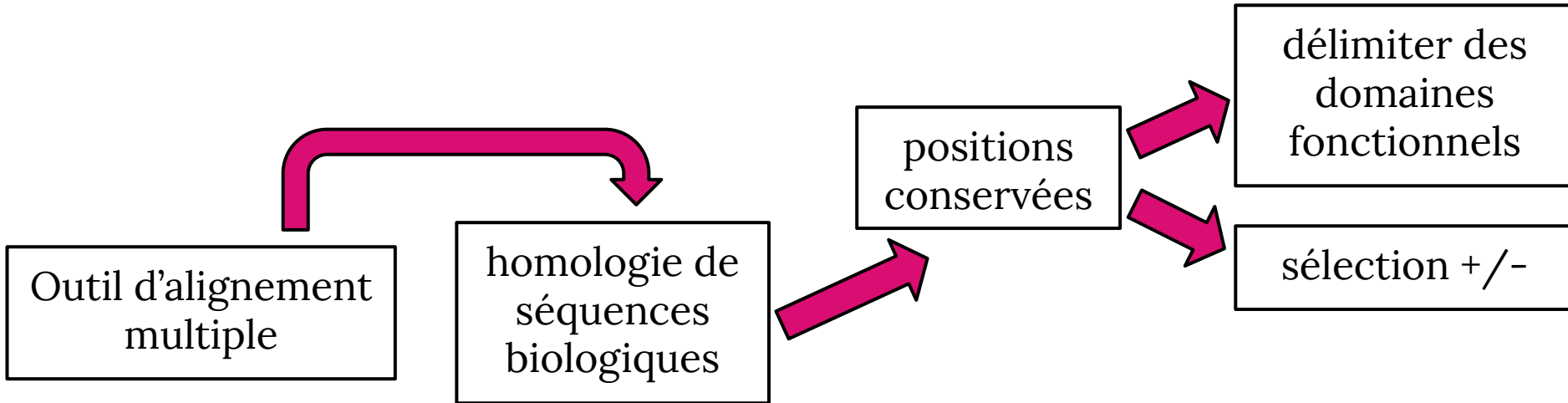
Applications



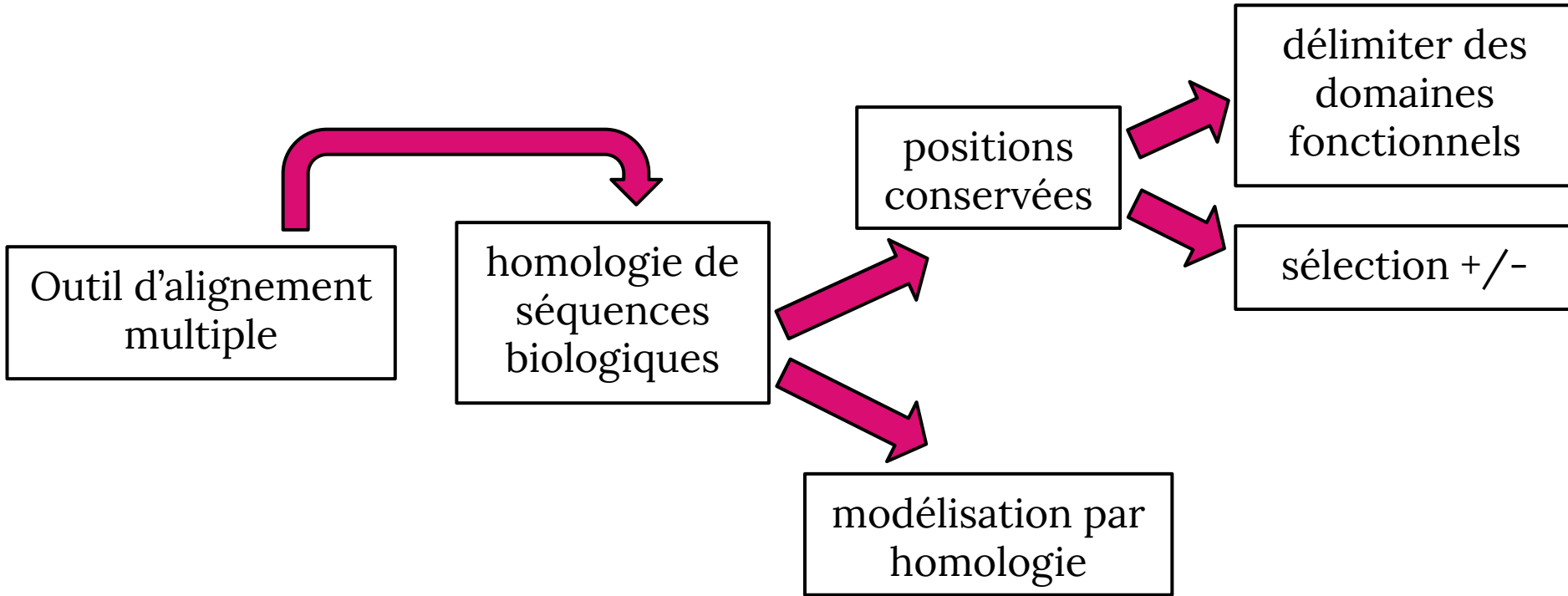
Applications



Applications

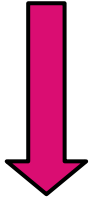


Applications



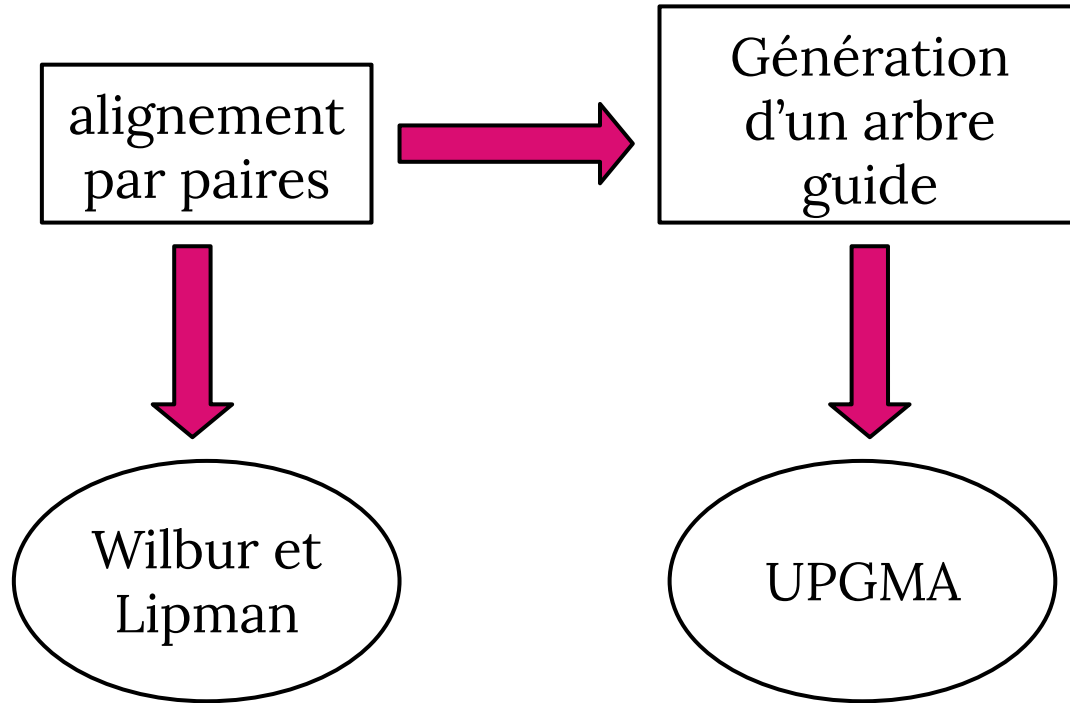
Étapes

alignement
par paires

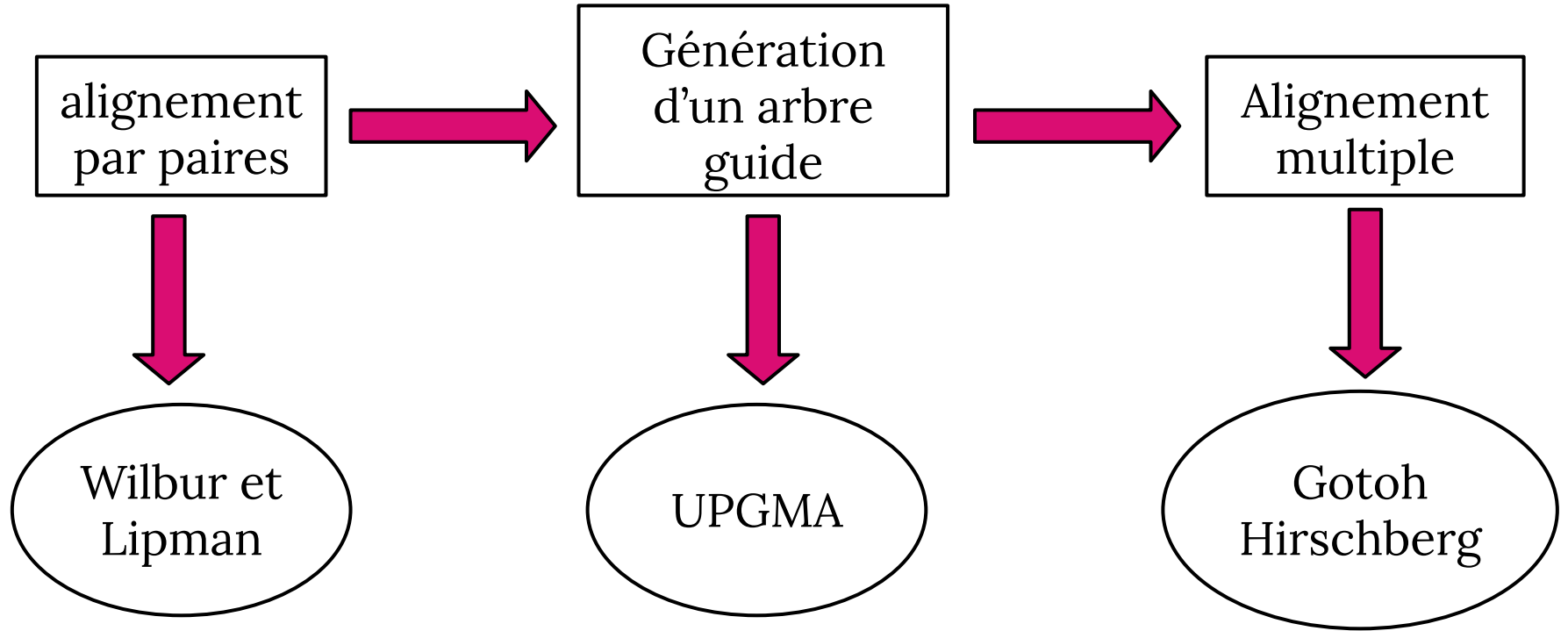


Wilbur et
Lipman

Étapes



Étapes

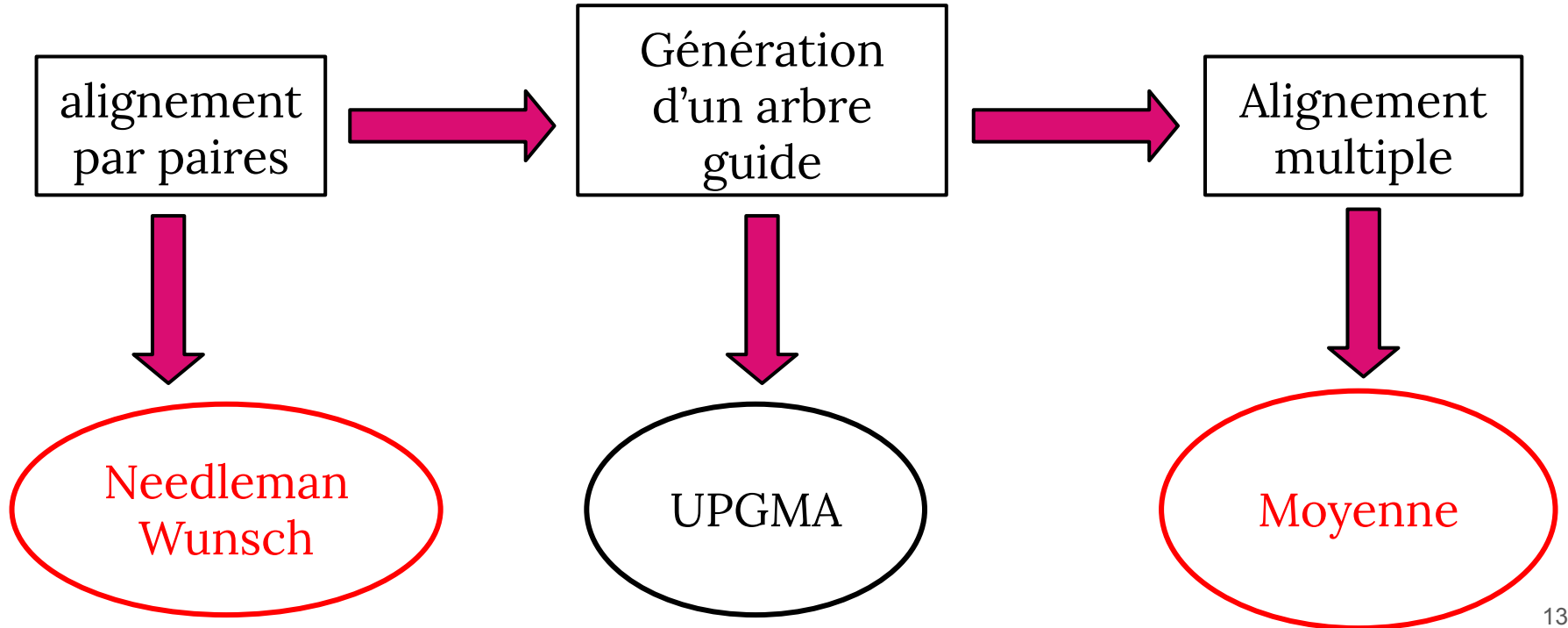


Objectif

- *Réimplémenter une version simplifiée de Clustal*

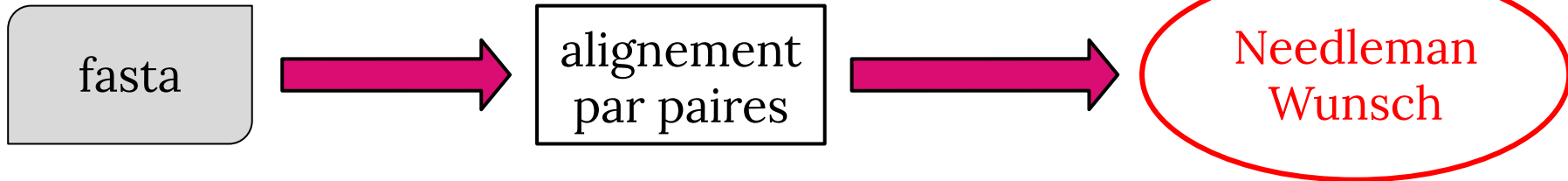
Objectif

- Réimplémenter une version simplifiée de Clustal



Les algorithmes

argparse (v1.1)
biopython (v1.84)



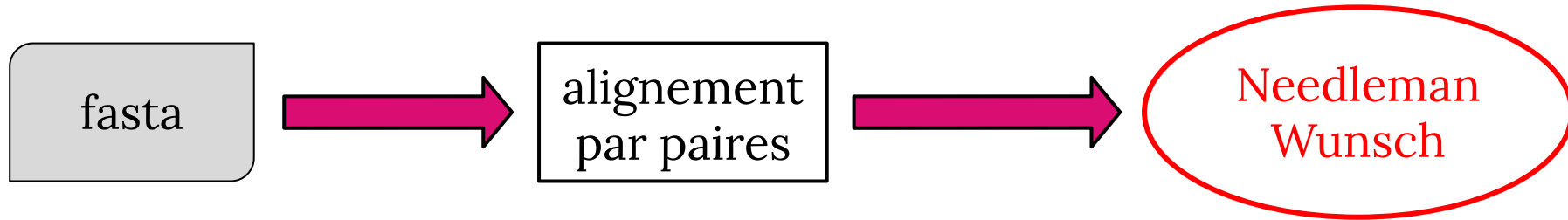
numpy (v2.1.1)

$F(I - 1, J - 1) + S(I, J)$	$F(I - 1, J) - \text{gap}$
$F(I, J - 1) - \text{gap}$	$F(I, J)$

Red arrows indicate dependencies: one from the top-left cell to the bottom-right cell, one from the top-left cell to the bottom-right cell, and one from the bottom-left cell to the bottom-right cell.

- $\text{gap} = 8$
- $S(I, J)$: BLOSUM62
pandas (v2.2.2)

Les algorithmes



$F(I - 1, J - 1) + S(I, J)$	$F(I - 1, J) - \text{gap}$
$F(I, J - 1) - \text{gap}$	$F(I, J)$

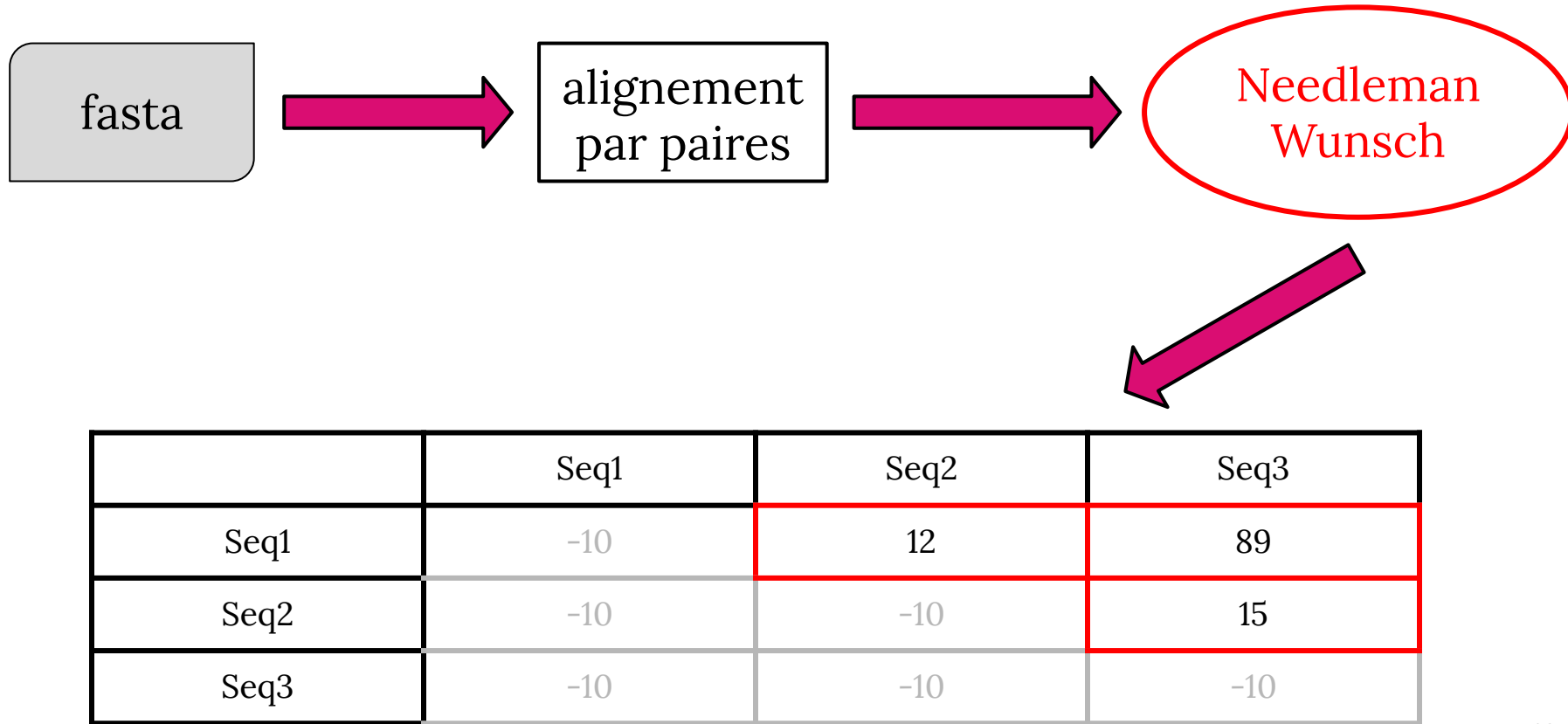
1	1
2	0

0 : diagonale

1: gauche

2: haut

Les algorithmes



Les algorithmes

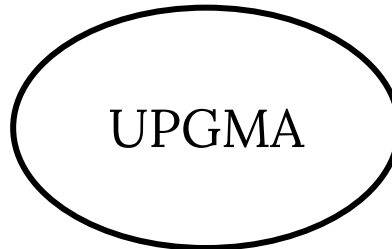
	Seq1	Seq2	Seq3
Seq1	-10	12	89
Seq2	-10	-10	15
Seq3	-10	-10	-10



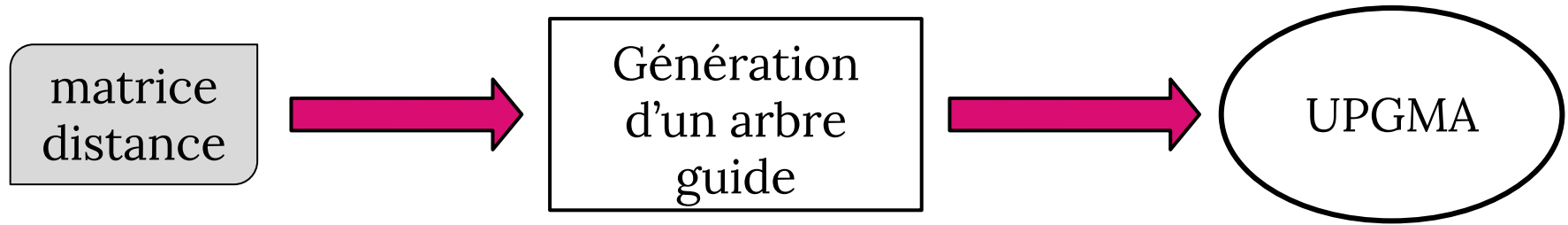
$$\text{matrice distances} = 1 - ((\text{matrice scores} - \text{minimum}) / \text{maximum} - \text{minimum})$$

Les algorithmes

	Seq1	Seq2	Seq3
Seq1	-10	1	0
Seq2	-10	-10	0.96
Seq3	-10	-10	-10



Les algorithmes

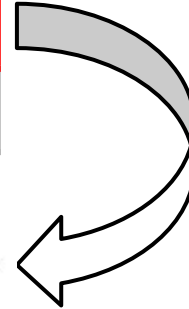


	Seq1	Seq2	Seq3
Seq1	-10	1	0
Seq2	-10	-10	0.96
Seq3	-10	-10	-10

Les algorithmes

	Seq1, Seq3	Seq2
Seq1, Seq3	-10	0.98
Seq2	-10	-10

$$\text{distance}((i, j), k) = \frac{n_i * d_{ki} + n_k + d_{kj}}{n_i + n_j}$$



Les algorithmes

	Seq1, Seq3	Seq2
Seq1, Seq3	-10	0.98
Seq2	-10	-10

$$\text{distance}((i, j), k) = \frac{n_i * d_{ki} + n_k + d_{kj}}{n_i + n_j}$$

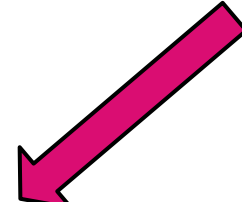
Arbre guide : ((Seq1, Seq3), Seq2)

Les algorithmes

Arbre guide : ((Seq1, Seq3), Seq2)



Alignement multiple



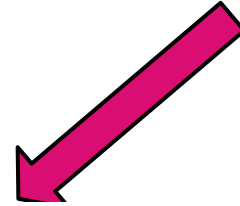
$$\textit{diagonale}(F(i, j)) = F(i - 1, j - 1) + \textit{moyenne}(\textit{cluster1}(i), \textit{cluster2}(j))$$

Les algorithmes

Arbre guide : ((Seq1, Seq3), Seq2)



Alignement multiple



$$\text{diagonale}(F(i, j)) = F(i - 1, j - 1) + \text{moyenne}(\text{cluster1}(i), \text{cluster2}(j))$$

Seq1 : **A**WVED

Seq2 : -TALD

Seq3 : **R**EDAL



$$\text{diagonale} = (S(A, R) - \text{gap}) / 2$$

Résultats

The file exists and is going to be parsed...

Here is the file's content :

Alpha-crystallinBchain: MDIAIHHWPWIRRPFFPFHSPSRLFDQFFGEHLLESDFPTSTSLSPFYLRPPSFLRAPSWFDTGLSEMRLEKDRFSVNLVDVKHFSPEELKVKVLGDVIEVH GKHEERQDEHGFIISREFHRRYRIPADVDPDLTITSSLSDDGLTVNGPRKQVSGPERTIPITREEKPAVTAAPKK

Alpha-crystallinAchain: MDVTIQHPWFKRTLGPFPYSRLFDQFFGEGLFEYDLLPFLSSTISPYRQSLFRTVLDSGISEVRSRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNERQDDHGYISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIPVSREEKPTSAPSS

Heatshockproteinbeta-6: MEIPVPVQPSWLRRASAPLPGLSAPGRLFDQRFGEGLLEAELAALCPTTLAPYYLRAPSVALPVAQVPTDPGHFSVLLDVKHFSPEEIAVKVGEHVEVHA RHEERPDEHGFVAREFHRRYRLPPGVDPAAVTSALSPEGVLSIQAAPASQAAPPPAAAK

Heatshockproteinbeta-1: MTERRVPFSLLRGPSWDPFRDWYPHSRLFDQAFGLPRLPEEWSQWLGGSSWPGYVRPLPPAAIESPAVAAPAYSRLSRQLSSGVSEIRHTADRWVSLDV NHFAPDELTVKTKDGVVEITGKHEERQDEHGYISRCFTRKYTLPPGVDPTQVSSSLSPGTLTVEAPMPKLATQSNEITIPVTFESRAQLGGPEAAKSDETAAK

Heatshockproteinbeta-2: MSGRSVPHAHPATAEYEFANPSRLGEQRFGEGLLPEEILTPTLYHGYVVRPRAAPAGEGSRAGASELRLSEGKFQAFLDVSHFTPDEVTVRTVDNLLVSA RHPQRLDRHGFVSREFCRTYVLPADVDPWRVRAALSHDGILNLEAPRGGRHLDTEVNEVYISLLPAPPDPEEEEEAAIVEP

Heatshockproteinbeta-8: MADGQMPFSCHYPSRLRRDPFRDSPLSSRLDDGFGMDPFPDDLTA SWPDWALPRLSSAWPGTLRSGMVPRGPTATARFGVPAEGRTPPFPGEPPWKVCVN VHSFKPEELMVKTKDGYVEVSGKHEEKQEGGIVSKNFTKKIQLPAEVDPVTVFASLSPEGLLIIEAPQVPPYSTFGESSFNNELPQDSQEVTC

Heatshockproteinbeta-3: MAKIILRHLIEIPVRYQEEFEARGLEDCLRDHALYALPGPTIVDLRKTRAAQSPPVDSAAETPPREGKSHFQILLDVVQFLPEDIIIQTFEGWLLIKAQHGT RMDEHGFISRSFTRQYKLPDGV EIKDLSAVLCHDGILVVEVKDPVGTK

Heatshockproteinbeta-7: MSHRTSSTFRAERSFHSSSSSSSSSTSSSASRALPAQDPPMEKALSMFSDDFGSFMRPHSEPLAFPARPGGAGNIKTLGDAYEFAVDVRDFSPEDIIVTTS NNHIEVRAEKLAADGTMNTFAHKCQLPEDVDPTSVTSALREDGSLTIRARRHPHTEHVQQTFRTEIKI

Heatshockproteinbeta-9: MQRVGNTFSNESRVASRCPVGLAERNRVATMPVRLLRDSPAAQEDNDHARDGFQMKLDAHGFAPPEELVVQVDGQWLMVTGQQQLDVRDPERVSYRMSQKV HRKMLPSNLSPTAMTCCLTPSGQLWVRGQCVALALPEAQTGPSRLGSLGSKASNLTR

Résultats

Here are the Needleman-Wunsch alignments :

Alpha-crystallinBchain: MDIAIHPWIRRPFFPFHSPSRLFDQFFGEHLLESDFP-TSTSLSPFYLRPPSFLRAPS

Alpha-crystallinAchain: MDVTIQHPWFKRTLGPYF-PSRLFDQFFGEGLFEYDLLPFLSSTISPYR-R-QSLFR--T

Alpha-crystallinBchain: WFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHGKHEERQDEHGFISREFH

Alpha-crystallinAchain: VLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNERQDDHGYISREFH

Alpha-crystallinBchain: RKYRIPADVDPLTITSSLSSDGVLTVNGPRKQ--VSG--PERTIPITREEKPAVTAAPKK

Alpha-crystallinAchain: RRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIPVSREEKP--TSAPSS

Alignement score : 466

Alpha-crystallinBchain: MDIAIH-HP-WIRR-PF-FP-FHSPSRLFDQFFGEHLLESDFPTSTSLSPFYLRPPSFL

Heatshockproteinbeta-6: MEIPVPVQPSWLRRASAPLPGLSAPGRLFDQRFGEGLLEAEL--A-ALCPTTL-APYYL

Alpha-crystallinBchain: RAPSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHGKHEERQDEHGFIS

Heatshockproteinbeta-6: RAPS-VALPVAQVPTDPGHFSVLLDVKHFSPEEIAVKVVEHVEVHARHEERPDEHGFVA

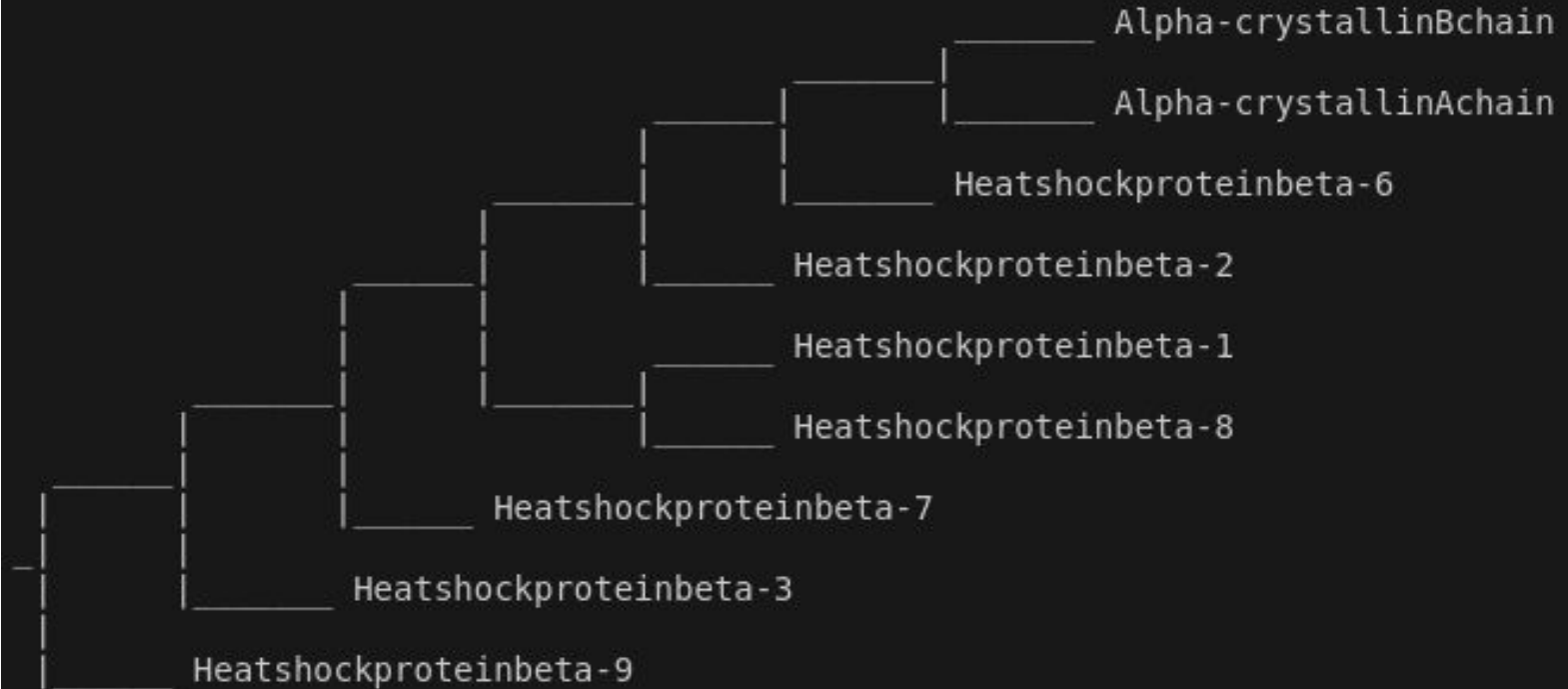
Alpha-crystallinBchain: REFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVSGPERTIPITREEKPAVTAAPKK

Heatshockproteinbeta-6: REFHRRYRLPPGVDPAAVTSALSPEGVLSI-----Q-AAP-AS---AQAPPPA--AA--K

Alignement score : 254

Résultats

Here is the UPGMA algorithm result :



Résultats

Here is the multiple alignment result :

```
Alpha-crystallinBchain: M-DIAIH-HP-WIRR-PFFPF--HS-PSRLFDQFFGEHLLESD-LFP-TSTSLSPFYLRP--P----S-F
Alpha-crystallinAchain: M-DVTIQ-HP-WFKR-TLGP--Y--PSRLFDQFFGEGLFEYD-LLPFLSSTISPYR--Q----S-L
Heatshockproteinbeta-6: M-EIPVPVQPSWLR--ASAPLPGLSAPGRLFDQRFGEGLLEAE-LAALCPTTLAPYYLRA--P----S--
Heatshockproteinbeta-2: M-S-GRS-VP-HAHP-ATAEYE-FANPSRLGEQRFGEGLLEE-IL--TPTLYHGYVVR--P----R-A
Heatshockproteinbeta-1: MTERRVPFSL-LRGP-SWDPRDWYPSRLFDQAFGLPRLPEE-WSQWLGSSWPGYVRPLPPAAIESPA
Heatshockproteinbeta-8: MADGQMPFSCHYPSRLRRDPFRDSPLSRLLDDGFGMDPFPDDLTAWS-PDWALP-RLSSAWPGTLRSGM
Heatshockproteinbeta-7: M-SHRTS-ST-F--R-AERSFHSSSSSSSSSSSSASRALPAQ-DPP-MEKALSMFSDDF--G----S-F
Heatshockproteinbeta-3: M-A-KI--I--LRHL-IEIPVR-Y--QEEFEARGLEDCLRDHA-LYA-LPGPTIV-DLRK--T----R-A
Heatshockproteinbeta-9: M-Q-RVG-NT-FSNE-SR--V-----ASRC-PS-VG--LAERN--R--VAT-M-P--VR-----L
```

```
Alpha-crystallinBchain: LRAP--S-W-FD--T-GLSEMRLEKDRFSVNLVDVKHFSPEELKVKVLGDVIEVHGKHE-E-RQDEH-GF-
Alpha-crystallinAchain: FR---T-V-LD--S-GISEVRSRDRKFIIFLDVKHFSPEDLTVKQVDDFVEIHGKH--E-RQDDH-GY-
Heatshockproteinbeta-6: V----A---L---P--VAQVPTDPGHFSVLLVDVKHFSPEEIAVKVGEHVEVHARHE-E-RPDEH-GF-
Heatshockproteinbeta-2: APAG--E-G-SR--A-GASELRLSEGKFAQLDVSHFTPDEVTVRTVDNLEVSARHP-Q-RLDRH-GF-
Heatshockproteinbeta-1: VAAPAYSRLSRQLSSGVSEIRHTADRWRVSLDVNHFADELTVKTKDGVVEITGKHE-E-RQDEH-GY-
Heatshockproteinbeta-8: VPRGPTATARFGVPAEGRTPPFPPEGPWKVCNVVHSFKPEELMVKTKDGYEVVSGKHE-E-KQQEG-GI-
Heatshockproteinbeta-7: MRPHSEPLA-FPARPGGAGNIKTLDGAYEFAVDVDRFSPEDIIVTTSNNHIEVRA---E-KLAAD-GT-
Heatshockproteinbeta-3: AQSP--P-V-DS--A-AETPPREGKSHFQILLDVVQFLPEDIIITQFEGWLLIKAQH-G-T-RMDEH-GF-
Heatshockproteinbeta-9: LR-D--S---P--A-AQEDNDHARDGFMKLDAGHFAPEELVVQVDGQWLMVTGQQQLDVRDPERVSYR
```

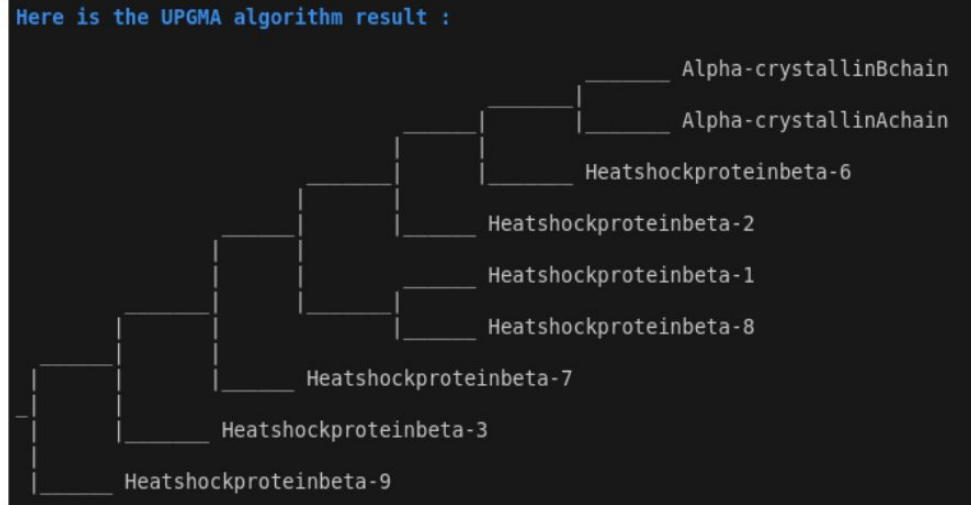
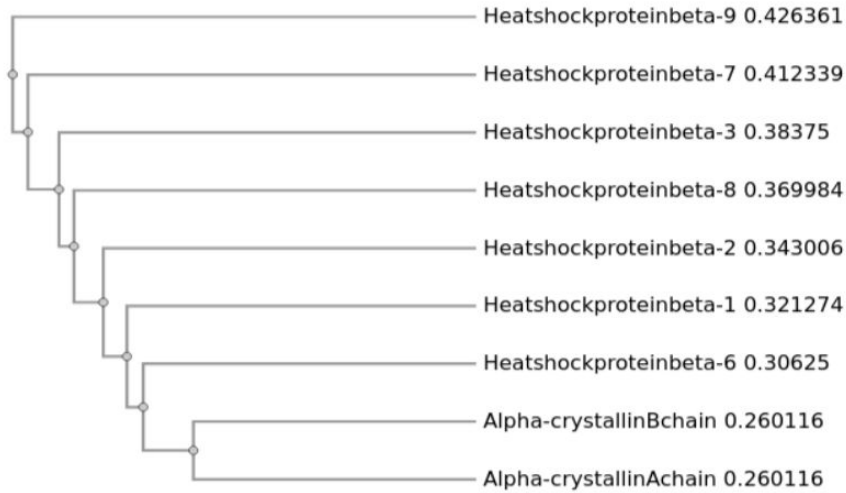
```
Alpha-crystallinBchain: ISREFHRKYRIPADVPLTITSSSSDGLTVNGPRKQ--VSG--PE-R-T-IPITR-EE-KPAV-TAA-
Alpha-crystallinAchain: ISREFHRRYRLPSNVDSALSCSLSDAGMLTFCGPKIQGLDATHAE-R-A-IPVSR-EE-KP---TSA-
Heatshockproteinbeta-6: VAREFHRRYRLPPGVDPAAVTSALSPGVLST----Q---A-----A--PASA-QA-PP---PAA-
Heatshockproteinbeta-2: VSREFCRTYVLPADVDPWRVRAALSHDGILNLEAPRGGRHLDTEVNEVYISLLPAPP-DP-EEEE-EAAI
Heatshockproteinbeta-1: ISRCFTRKYTLPPGVDPQTQVSSSLSPGTLTVEAP-MP-KLATQSNITIPVTTFESRAQLGGPEAAKSDE
Heatshockproteinbeta-8: VSKNFTKKIQLPAEVDPTVTFASLSPEGLLIEAPQVP-PYST-FGE-S---SF-NN-EL--PQ--DSQE
Heatshockproteinbeta-7: VMNTEFAHKCQLPEDVDPTSVTSALREDGSLTIRARRHP-H--T---E-H---V---Q-QT--F--RTE-
Heatshockproteinbeta-3: ISRSFTRQYKLPDGVIEIKDLSAVLCHDGILVV-----E-----V---K-D---P---VG-
Heatshockproteinbeta-9: MSQKVHRKM-LPSNLSPTAMTCLTPSGQLWVRGQCVA--LAL--PEAQ-T-GPSRLGS-LGS--KASN
```

```
Alpha-crystallinBchain: PK-K
Alpha-crystallinAchain: PS-S
Heatshockproteinbeta-6: -A-K
Heatshockproteinbeta-2: VE-P
Heatshockproteinbeta-1: TAAK
Heatshockproteinbeta-8: VTCT
Heatshockproteinbeta-7: IK-I
Heatshockproteinbeta-3: -T-K
Heatshockproteinbeta-9: LT-R
```

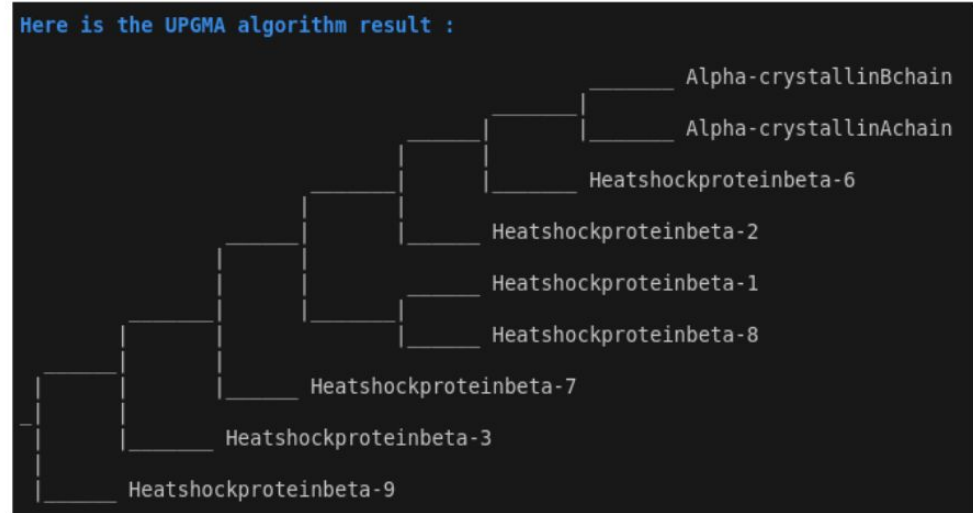
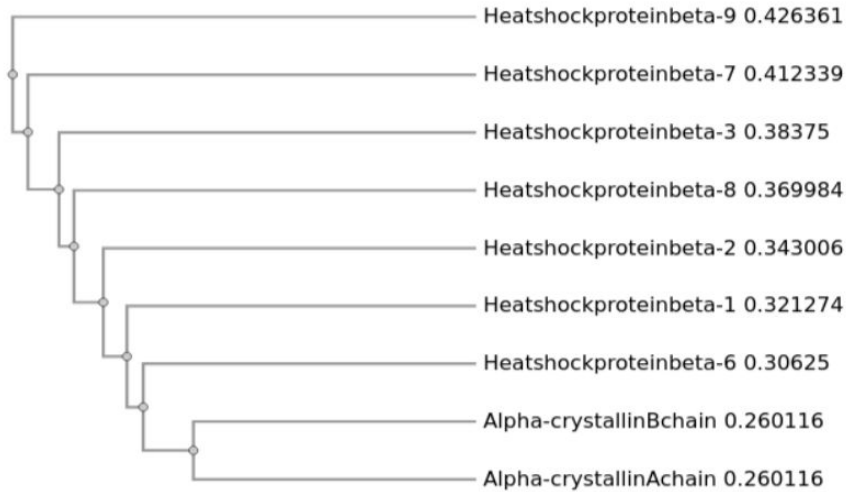
Résultats

9 séquences : COMPLEXITÉ EN TEMPS DE $O(n^2)$	
175 acides aminés	14 secondes
355 acides aminés	37 secondes
860 acides aminés	2 minutes 19 secondes

Résultats : Comparaison

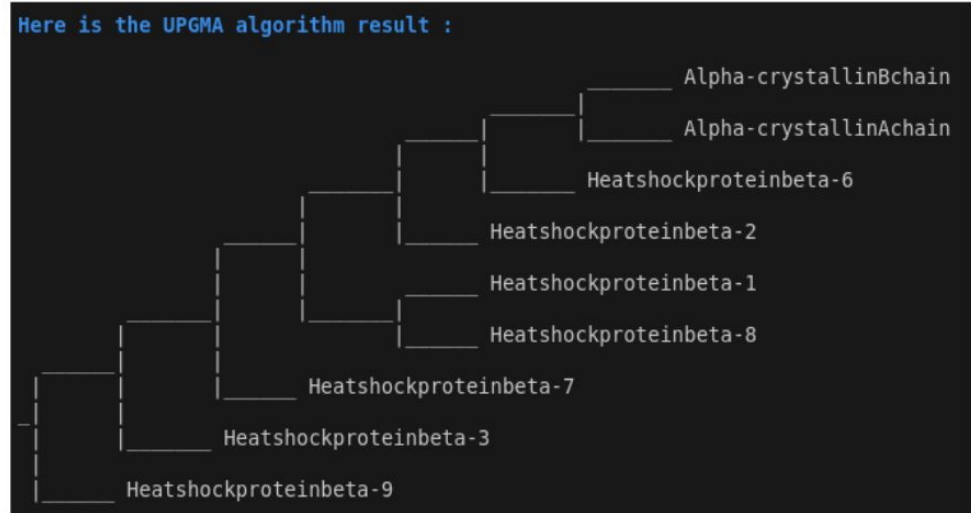
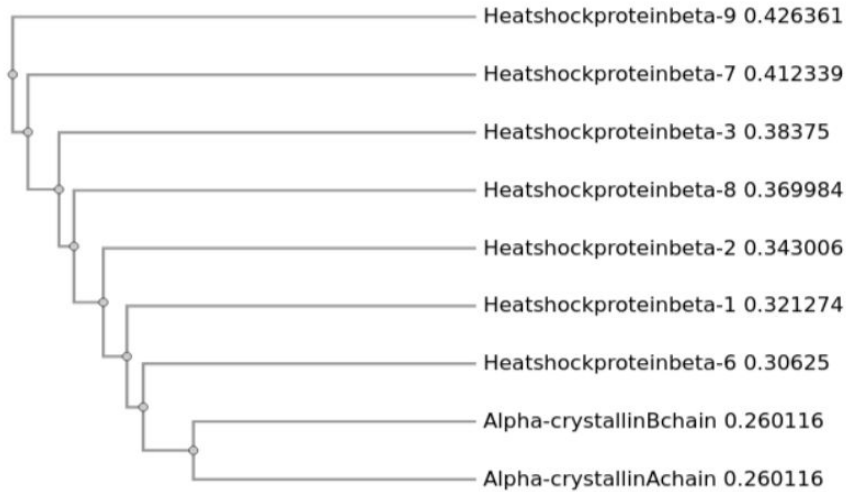


Résultats : Comparaison



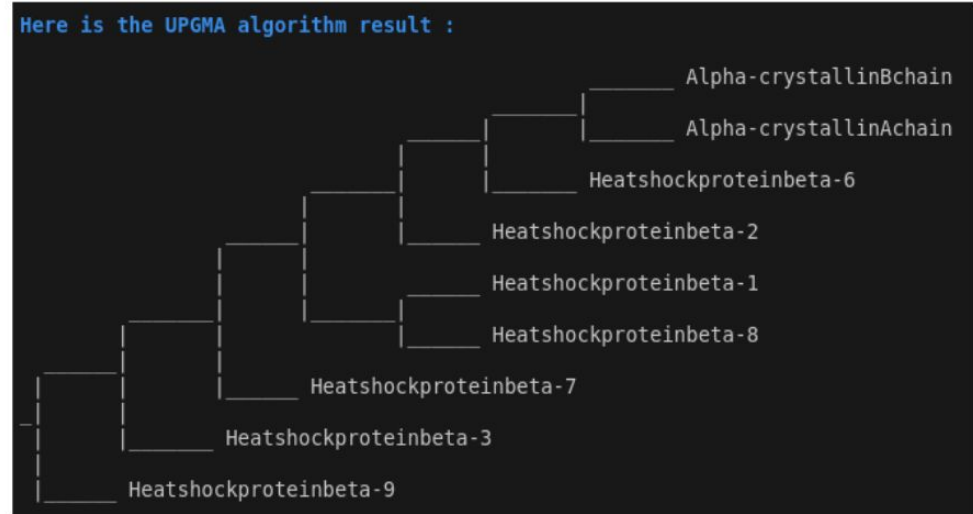
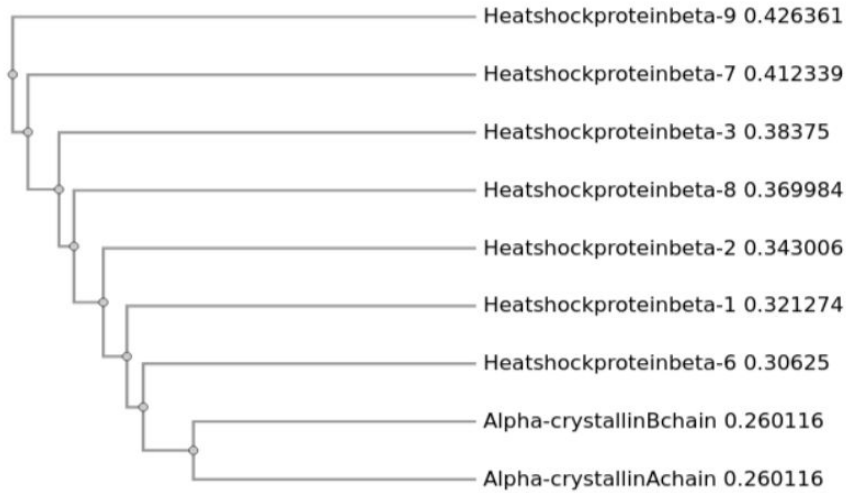
- Needleman-Wunsch VS Wilbur et Lipman

Résultats : Comparaison



- Needleman-Wunsch VS Wilbur et Lipman
- pénalité de gap constante

Résultats : Comparaison



- Needleman-Wunsch VS Wilbur et Lipman
- pénalité de gap constante
- plusieurs méthodes de clusterisation (mBed, k-means, UPGMA)

Résultats : Comparaison

CLUSTAL O(1.2.4) MULTIPLE SEQUENCE ALIGNMENT

HEATSHOCKPROTEINBETA-9	-----MQ-----RVNGTFSNEIR-----VASR	16
HEATSHOCKPROTEINBETA-7	MSHRT-----SSTF-----RAERFSHSSSSSSSSSSSSASRALPAQDPPEKALK	47
HEATSHOCKPROTEINBETA-3	MAKIIILRL-----DEEF-----ETIPVRYQ-----EEG-----EARGLEDPCDL	30
HEATSHOCKPROTEINBETA-8	MADGQMFSSCHYP-SRL--RRDPFRDPSLSSRL--DDF-----FGMDPFDRLTAS	47
HEATSHOCKPROTEINBETA-2	MSGRSVPAHP-----ATAEYEFANPSRLG--EQR-----FGGLPEELIT-	40
HEATSHOCKPROTEINBETA-1	MTRRVRFVSLGRPSW--PDFRWYPSHSLRF--DQA-----FGLPRLPESSQW	45
HEATSHOCKPROTEINBETA-6	-----MEIPVPVQSLRRASAPLGLSAPGRFL--DQR-----FGHGLEAEALAL	45
ALPHA-CRYSTALLINBCHAIN	-----MDIAIHPWTRPPFPF--FHSPSRLF--DQF-----FGHGLESLDPLP	40
ALPHA-CRYSTALLINBCHAIN	-----MDVIQHPFKRTLGF--F-YPSRLF--DQF-----FGHGFEDYLLPF	39

HEATSHOCKPROTEINBETA-9	CPS-	-VGLAEINRVAT-	MPV-	-LLRLDPAAQEGDNHARDGQFMK	57
HEATSHOCKPROTEINBETA-7	-	MPDQFGFSRNPHS-	-	LAPPARPGAGNIGKITLGDAYEFA	85
HEATSHOCKPROTEINBETA-3	-	DHALYALPGPT-	TVDLR-	-KTRAAGSPVPYGSAAETPPREGKSHFOIL	94
HEATSHOCKPROTEINBETA-8	WPDWALPRLSSAMPTGLRSGMVRP-	-	GPTA-	TARFVGPAEGRTPPPFPGEKWKV	79
HEATSHOCKPROTEINBETA-2	-P-	TLHYGVYVPRPAAP-	-	AGEGRGASGLSELRGQFAF	77
HEATSHOCKPROTEINBETA-1	LG-	GSSNPQGVYRPLPAATESPAVAAYASRL	SQSLSGVSEIRTHADNRMVS	98	
HEATSHOCKPROTEINBETA-6	CP-	TTLAPPLYLA-	PSV-	ALPVAQVPTDPGHGSFV	98
ALPHA-CRYSTALLINCHAIN	ST-	SLSPFYLR-RP-	PSFLR-	APSWITGISEMRLEKRGKVFV	78
ALPHA-CRYSTALLINCHAIN	LS-	STISPY-RQ-	SLF-	RTVLDSGISEVRSRDKQFVFV	78

HEATSHOCKPROTEINBETA-9 LDHGFAPPELVVQVGDGVEVGVGKQDLVDRPERSYVSKQVHKMLPSMTATMC 117
HEATSHOCKPROTEINBETA-7 DVDRFDSPEDIIIVTSGMNIIEVRA---EKKLADEG---TVNMTFATKQCPEDVDTSVTS 139
HEATSHOCKPROTEINBETA-3 LDVVQFLPEDIITQTFEGWLLIKAQHGRMDEHG---FISRSFTRQYKLPDGVGKIDLSA 139
HEATSHOCKPROTEINBETA-8 VNVHSPKPEELWLVKLVQGVSEVSGHEEKQEGEG---IVSKNFKTIQLPADVDVPVTFVA 156
HEATSHOCKPROTEINBETA-2 LDVSHFTPOVTVRTVDNLLVFSARHPQKLDRIHG---FVFSREFCTRYVLPADVDPWIVRA 156
HEATSHOCKPROTEINBETA-1 LDVNHGFAPDELTVTKYGDGVGKHEERQDEHG---YISRCFTKYTLPDGVDPQVSS 158
HEATSHOCKPROTEINBETA-6 LDVNHGFSEETAIVKVGVEHIVHARHEERDEHG---FVAFERHRRYRLPQGVDPAAVTS 139
ALPHA-CRYSTALLINBCHAIN LDVKHGFSEELKVKVLGDVIEVHGKHEERQDEHG---FISREFHRRYRIPADVPLITS 139
ALPHA-CRYSTALLINACHAIN LDVKHGFSEDLTVKVGQDFDIEVHGKHNERRQDDHG---YISREFHRRYRLPSNDVQSALSC 131

HEATSHOCKPROTEINBETA-9	CLTPSGQLVMVRGGQVICALPEAQGTG-----SPRLGSLGSKASNLTR----	159
HEATSHOCKPROTEINBETA-7	ALRDGSLDITLRAARPHHTHVHQVQ-TRFTEIKI-----	170
HEATSHOCKPROTEINBETA-3	VLCHDGLTLVVEVDPVGTG-----	150
HEATSHOCKPROTEINBETA-8	SLSPGELLIIAEPQVPSYTFGESSFNNELPDQDSQE-----VTCT-----	196
HEATSHOCKPROTEINBETA-2	ALSHDGLNLVQGGPHGLHLDTEVN--EYVTSLLPA-----PPDPEEEEEAAIVP	182
HEATSHOCKPROTEINBETA-1	SLSPGTLTVAAEPMKPLAT--QSN-EITTPVTFESRQLQGPEAAKSDEATAK----	205
HEATSHOCKPROTEINBETA-6	ALSPGVLTVQAPASQAAPPPA--AK-----	160
ALPHA-CRYSTALLINBCHAIN	SLSDGVLTVGPGPKIQVQ-----SGP-ERTPTVTREREKPAVTAAPKK-----	175
ALPHA-CRYSTALLINBCHAIN	SLSDGMLTFCGPKITQGLDTALDA--ERATPVSEEREKPTSPASS-----	173

Alpha-crystallinBchain: :DIATH-HP-WIRR-PFFPF--HS-PSRLFDQFFGEHLLESD-LFP-TSTSLSPFYLRP--P----S-F
Alpha-crystallinAchain: M-DVTIQ-HP-WFKR-TGLPF--Y--PSRLFDQFFGEGLEFYD-LLPFLSSSTISPYLR--Q----S-F
Heatshockproteinbeta-6: M-EIYPVPQSVLWRR-ASAPGLSPAGRLFDQRFGEGLLEED-LAALCPPTLAPYYLR--P----S--
Heatshockproteinbeta-2: M-S-GRS-VF-HAHP-ATAEY-FANPSRLGEQRFGEGLLPEE-IL--TPTLYGGYYVR--P----R-A
Heatshockproteinbeta-1: MTERRVQFSL-LRGP-SWDPFRDWPHSRLFDQAQFLRLPCE-WSQWLGGSSWPGVYRPLPPAAIESPA
Heatshockproteinbeta-8: MADGQVFFSCHYSLRRRDFRFDPSLLSRLLDGGFMDFPDDLTAS-PDWALP-LSSAWPGTLRSGM
Heatshockproteinbeta-7: M-SHRTS-ST-F--R-AERSFHSVSSSSSSSSSSSSSSSRAALPAQ-DPP-MEKALSMFSDDF--G----S-F
Heatshockproteinbeta-3: M-A-KI-I--LRHL-TEIYPR-Y---QEFEAGLEDRLDHA-LYA-LPGPTIV-DLRK-T----R-A
Heatshockproteinbeta-9: M-Q-RVG-NT-FSNE-SR-V---ASRC-PS-VG-LAERN-R-VAT-M-P-VR-----

Alpha-crystallinBchain: LRAP--S-W-FD--T-GLSEMRLEKDRFSVNLDVKKHFSPEELKVKVLGDVIEVHGKHE-E-RQDEH-GF-
Alpha-crystallinAchain: FR---T-V-LD--S-GISEVSRDQDKFVFLDVKKHFSPEELTKVQVDDFVEIHGKHE-E-RQDDH-GY-
Heatshockproteinbeta-6: V-----A--L--P--VAQSPDTPGHFSVLNDVKHFSPEELIAVVGVGHEVHKHHE-E-RPDEH-GF-
Heatshockproteinbeta-2: APAG--E-G-SR--A-GASELRLSEGKQFAFLDVDSHFTPDVEVTVRTVDNLLVEVSARHP-Q-RLDRH-GF-
Heatshockproteinbeta-1: VAAPAYSRAISRQLSSVSEIRHDAWRKVSVDVNHFPADLETKTKDGVLETSKHE-E-E-RQDEH-GY-
Heatshockproteinbeta-8: VPRGTATARGFVPAEGRTPPPFPFGEKMKCVNHHSPKPEELMKTKDGVVEVSGKHE-E-QQEG-GF-
Heatshockproteinbeta-7: MRPHSEPLA-FPARPGGAGNIKTLGDAYEFAVLDVDFSPEDIIVTSSNNHLEKRA---E-KLAAD-GT-
Heatshockproteinbeta-4: AQPSP--P-V-DS--A-AETPPREGKSHFQLVDVDFPELIIITTFEGWLLIVRA---E-RMDEH-GF-
Heatshockproteinbeta-9: LR-D-S--S---P--A-AEQNDHARDGFQMKLDAHGAPPEELVQVQDQGLVMTGQQLDVRDPERVSY-

Alpha-crystallinBchain: ISREFHRYKYPADVDPLTITSSSSDGLVTVNGPRKQ--VSG--PE-R-T-IPITR-EE-KPAV-TAA-
Alpha-crystallinAchain: VREFHRRYRLPPGVNDQSLASCSLSADGMLTFCGPKITQGLDATHAE-R-A-IPVSR-EE-KP--TSA-
Heatshockproteinbeta-6: VAREFHRRYRLPPGVNDQSPAAVTSQSGELSI---Q---Q---A---PASA-QA-PP---PAA-
Heatshockproteinbeta-2: VSREFCTRYVLPAVDVPRVRAALSHDGLNLNEAPGRGRHLDTEVNEVYSLLPAPD-PP-EEEE-EAAI-
Heatshockproteinbeta-1: ISRCFTRYKLPVGDVPTQVSSLSPEGTITVEAP-MP-KLATQSNIEIPTVFESPAQLGGPEAAKSDE-
Heatshockproteinbeta-8: VSKNFTHKYLPAEVDVPTVFASSPEGLLIEAPQP-PYST-FGE-S--SF-NN-LD-PQ--DSQE-
Heatshockproteinbeta-7: VMNFTAHKQLPEDVDPTSVALSALREDSGITIRARRHP-H-T--E-H--V--Q-Q-T-F--RTE-
Heatshockproteinbeta-9: ISRSFTRQYKLPDGVGEIKLISAVLCHDGLTV---E---E---V--K-D--P--V--V-
Heatshockproteinbeta-3: MSQKHVRKM-LPSNLSPANTCCLTPSGOLVWRGQCA--LAL-PEAQ-T-GPSPLRGS-LGS-KASN-

Alpha-crystallinBchain: PK-K
Alpha-crystallinAchain: PS-S
Heatshockproteinbeta-6: -A-K
Heatshockproteinbeta-2: VE-P
Heatshockproteinbeta-1: TAAK
Heatshockproteinbeta-8: VTCT
Heatshockproteinbeta-7: IK-I
Heatshockproteinbeta-3: -T-K
Heatshockproteinbeta-9: LT-R

Résultats : Comparaison

CLUSTAL O(1.2.4) MULTIPLE SEQUENCE ALIGNMENT

```

HEATSHOCKPROTEINBETA-9  -----MQ-----RVGNTFSNEIR-----VASR      17
HEATSHOCKPROTEINBETA-7  MSHRT-----SSTF-----RAERSFHSSSSSSSTSSASRALPAQDPPEKALS--      46
HEATSHOCKPROTEINBETA-3  MAKILRLHL-----ETPRVQ-----EEF-----EARGLEDCRL--      38
HEATSHOCKPROTEINBETA-8  MADGQMPFSCHYP-SRL--RRDPFRDPSLSRL--DDG-----FGMDPFDOLTAS      47
HEATSHOCKPROTEINBETA-2  MSGRSVPHAHP-----ATAEYEFANPRDG-----EQR-----FGEGGLEPEILT-      40
HEATSHOCKPROTEINBETA-1  MTERRVPFSLRRGSPW-----DPFRDWPYHRGF--DQA-----FGLPRLPEEWSQW      45
HEATSHOCKPROTEINBETA-6  ---MEIPVPVQPSWLRASAPLPLGLSAPRGF--DQR-----FGGLELEAELAAL      45
ALPHA-CRYSTALLINBCHAIN  ---MDIAIHPWIRRRPFPPF--FHSPPRGF--DQF-----FGHELESOLFPT      46
ALPHA-CRYSTALLINACHAIN  -----MDVTIQHWPKRTLGP-----F-YPRRGF--DQF-----FGGLEFEYDLPF      39
    
```

```

HEATSHOCKPROTEINBETA-9  CPS-----VGLAERNRVAT-----MPV--LRLLRSDPAAQEDNDHARDGFMK      57
HEATSHOCKPROTEINBETA-7  ---MFSDDFGSFMRPHSEP-----LAFPARPGGAGNIKTLGDAYEFA      85
HEATSHOCKPROTEINBETA-3  -----DHALYALPGPT--IVDLR--KTRAAQSPVVDAAETPPREGKSHFQIL      74
HEATSHOCKPROTEINBETA-8  WPDNALRLSSAMPTLRSGMVRP-----GPTA-TARFQVPAEGRTPPPFGPEKQVC      99
HEATSHOCKPROTEINBETA-2  -P-----TLYHGYYVPPRAAP-----AGEGSRAGASRLSEKGFQAF      77
HEATSHOCKPROTEINBETA-1  LG-----GSSMPGYRPLPAAAEISPAVAAPAYSRLSRQLSSGVSEIRHTADMRVRS      98
HEATSHOCKPROTEINBETA-6  CP-----TTLAPYYLR-----PSV-----ALPVAQVTPDGHFSVL      77
ALPHA-CRYSTALLINBCHAIN  ST-----SLSPFLY-RR-----PSFLR-APSWIFDGLSEMRLEKGFVSN      78
ALPHA-CRYSTALLINACHAIN  LS-----STISPIY-RQ-----SLF--RTVLDSGISEVRSRDRKQVIF      74
    
```

```

HEATSHOCKPROTEINBETA-9  .DAHGFAPEELVVQVQDGLMVTGQQQLDVRDPEVSYVMSQKVHRKMLPSLSPMTATC      117
HEATSHOCKPROTEINBETA-7  VVDRDQSPEDIIVTTSNNHIEVRA---EKLAADG---TMTNTFAHKCQLPEDVPTSVTS      139
HEATSHOCKPROTEINBETA-3  .DVVQFLPEDIIQTTFEGWLLIKAQHGTRDEHG--FSSRSFTQYKLPDGVIEIKDLSA      131
HEATSHOCKPROTEINBETA-8  VNVHSFKPEELMWKTDGVEVSGVSGHKEEKQEGG--IVSKNFTKKIQLPAEVDVPTVFA      156
HEATSHOCKPROTEINBETA-2  .DVSHFTPDEVTVRTVDNLLVESARHPQLDRHG--FSSREFCRTYVLPADVDPWRVRA      134
HEATSHOCKPROTEINBETA-1  .DVNHFAPELTVKTKDGVVEITGKHEERQDEHG--VYSRCFTKYTLPPGVDPTQVSS      155
HEATSHOCKPROTEINBETA-6  .DVKHFSPEEIAVKVVEGVHVEHARHEERDQHG--FPAREFHRRYRLPPGVDPAVTS      145
ALPHA-CRYSTALLINBCHAIN  .DVKHFSPEELKVYKVGDIIEVHGKHEERQDEHG--FSSREFHRRYRIPADVDPILTTS      135
ALPHA-CRYSTALLINACHAIN  .DVKHFSPEELTVKVGDDFVEIHGKHNERQDQHG--YYSREFHRRYRLPSNDQSQALS      131
    
```

```

HEATSHOCKPROTEINBETA-9  CLTPSGQLWVRGQCVALLPEAQGTG-----SPRLSGSLGKASNLTR-----      159
HEATSHOCKPROTEINBETA-7  ALREDGSLTIRARRHPHTEHVQQTFRTEIKI-----      170
HEATSHOCKPROTEINBETA-3  VLCHDGLVVEVQDPVGTG-----      150
HEATSHOCKPROTEINBETA-8  SLSPEGLLIEAPQVPVPTGTFGESSFNNELPQDSQE-----VTCT--      196
HEATSHOCKPROTEINBETA-2  ALSHDGILNLEAPRGGRHLDETVN--EYISLLPA-----PPDPEEEEEAAIVP      182
HEATSHOCKPROTEINBETA-1  SLSPEGTLVEAPMKLAT--QSN--EITIPVTFSRAQLGGPEAAKSDETAAK--      205
HEATSHOCKPROTEINBETA-6  ALSPEGVLISQAAPASQAAPPAA--AK-----      160
ALPHA-CRYSTALLINBCHAIN  SLSSDGVLTVNGPRKQV--SGP--ERTIPITREKPAVTAAPK--      175
ALPHA-CRYSTALLINACHAIN  SLSADGMLTFCGPKIQTGLDATHA--ERAIPVSREKPKTAPSS-----      173
    
```

```

Alpha-crystallinBchain: M-DIAIH-HP-WIRR-PFFPF--HS-PRLFDQFFGHEHLES-D-LFP-TSTLSLSPFYLRP--P---S-F
Alpha-crystallinAchain: M-DVTIQ-HP-WFKR-TLGP--Y--PRLFDQFFGEGLEFYD-LLPFLSSSTISPIY-R---Q---S-L
Heatshockproteinbeta-6: M-EIPVPVQPSWLR--ASAPLPLGLSAPRG-LDQRFEGGLELEA-LAALCPTTLAPYYLRA--P---S-
Heatshockproteinbeta-2: M-S-GRS-V-PAHPR-ATAEYE-FANPRLGEORFEGGLELPEE-IL--TPTLYHGYYVR--P---R-A
Heatshockproteinbeta-1: MTERRVPFSL-LRGP-SWDPRDWPYHRLFDQAGFLPRLPEE-WSQWLGGSSWPGYVRPLPPAAIESPA
Heatshockproteinbeta-8: MADGQMPFSCHYPSRLRRDPFRDPSLSRLLDGFGMDPFPDDLTA-SW-PDWALP-RLSSAWPGTLRSGM
Heatshockproteinbeta-7: M-SHRTS-ST-F--R-AERSFHSSSSSSSSSTSSASRALPAQ-DPP-MEKALSMFSDDF--G---S-F
Heatshockproteinbeta-3: M-A-KI--I--LRHL-IEIPVR-Y--QEEFEARGLEDCLRDHA-LYA-LPGPTIV-DLRK--T---R-A
Heatshockproteinbeta-9: M-Q-RVG-NT-FSNE-SR--V-----ASRC-PS-VG--LAERN--R--VAT-M-P--VR-----L
    
```

```

Alpha-crystallinBchain: LRAP--S-W-FD--T-GLSEMRLEKDRFSVLDVVKHFSPEELKVKVLGDVIEVHGKHE-E-RQDEH-GF-
Alpha-crystallinAchain: FR---T-V-LD--S-GISEVRSRDRKDFVLDVVKHFSPEELTVKVQDDFVIEHGKHN-E-RQDDH-GY-
Heatshockproteinbeta-6: V----A--L--P--VAQVPTDPGHFSVLDVVKHFSPEEIAVKVVEGVHVEHARHE-E-RPDEH-GF-
Heatshockproteinbeta-2: APAG--E-G-SR--A-GASELRLSEGKFAFLDVSHFTPDEVTVRTVDNLLVESARHP-Q-LDRH-GF-
Heatshockproteinbeta-1: VAAPAYSRLSRQLSSGVSEIRHTADRWRLVDVNHFAPELTVKTKDGVVEITGKHE-E-RQDEH-GY-
Heatshockproteinbeta-8: VPRGPTATARFGVPAEGRTPPPPFGPEKWKVCNVHVSFKPEELMWKTDGVEVSGKHE-E-KQDEH-GI-
Heatshockproteinbeta-7: MRPHSEPLA-FPARPGGAGNIKTLGDAYEFAVDVDRDQSPEDIIVTTSNNHIEVRA---E-KLAAD-GT-
Heatshockproteinbeta-3: AQSP--P-V-DS--A-AETPPREGKSHFQILLDVVQFLPEDIIQTTFEGWLLIKAQHG-T-RMDEH-GF-
Heatshockproteinbeta-9: LR-D--S---P--A-AQEDNDHARDGFMKLDHAHGAPEELVVQVQDGLMVTGQQQLDVRDPERVSYR
    
```

```

Alpha-crystallinBchain: ISREFHRRYRIPADVDPILTTSVSSDGLTVNGPRKQ--VSG--PE-R-T-IPITR-EE-KPAV-TAA-
Alpha-crystallinAchain: ISREFHRRYRLPSNDQSQALSGLSADGMLTFCGPKIQTGLDATHAE-R-A-IPVSR-EE-KP--TSA-
Heatshockproteinbeta-6: VAREFHRRYRLPPGVDPAVTSVLSPEGVLSI-----Q-----A-----A-PASA-QA-PP--PAA-
Heatshockproteinbeta-2: VSREFCRTYVLPADVDPWRVRAVLSHDGILNLEAPRGGRHLDETVNEVYISLLPAPP-DP-EEEE-EAAI-
Heatshockproteinbeta-1: ISRCFTRYKTLPPGVDPTQVSSLSPEGTLTVEAP-MP-KLATQSNIEITPVTFSRAQLGGPEAAKSDE
Heatshockproteinbeta-8: VSKNFTKKIQLPAEVDVPTVFAVLSPEGLLIEAPQV-PYST-FGE-S--SF-NN-EL--PQ--DSQE
Heatshockproteinbeta-7: VMNTFAHKCQLPEDVPTSVTSVLSREDGSLTIRARRHP-H--T---E-H---V---Q-QT--F---RTE-
Heatshockproteinbeta-3: ISRSFTQYKLPDGVIEIKDLSVLDCHDGLLVV-----E-----V---K-D---P---VG-
Heatshockproteinbeta-9: MSQKVHRKM-LPSNLSPMTATCLTPSGQLWVRGQCV--LAL--PEAQ-T-GPSRLGS-LGS--KASN
    
```

```

Alpha-crystallinBchain: PK-K
Alpha-crystallinAchain: PS-S
Heatshockproteinbeta-6: -A-K
Heatshockproteinbeta-2: VE-P
Heatshockproteinbeta-1: TAAK
Heatshockproteinbeta-8: VTCT
Heatshockproteinbeta-7: IK-I
Heatshockproteinbeta-3: -T-K
Heatshockproteinbeta-9: LT-R
    
```


Résultats : Comparaison

CLUSTAL O(1.2.4) MULTIPLE SEQUENCE ALIGNMENT

```

HEATSHOCKPROTEINBETA-9      -----MQ-----RVGNTFSNESR-----VASR      17
HEATSHOCKPROTEINBETA-7      MSHRT-----SSTF-----RAERSFHSSSSSSSSSSSSASRALPAQDPPEKALS--      46
HEATSHOCKPROTEINBETA-3      MAKILRLHL-----ETPVRYQ-----EARGLEDCRL--      38
HEATSHOCKPROTEINBETA-8      MADGQMPFSCHYP-SRL--RRDPFRDPLSSRL--DQG-----FGMDPFPDGLTAS      47
HEATSHOCKPROTEINBETA-2      MSGRSVPHAHP-----ATAEYEFANPSRLG--EQR-----FGEGGLEPEILIT--      46
HEATSHOCKPROTEINBETA-1      MTERRVPFSLRLGPSW-----DPFRDWPYHSRLF--DQA-----FGLPRLPEEWSQW      45
HEATSHOCKPROTEINBETA-6      -----MEIPVPVQPSWLRASAPLGLSAPGRLF--DQR-----FGGLELEAELAA      45
ALPHA-CRYSTALLINBCHAIN      -----MDIAIHPWIRRRPFPF--FHSPSRLF--DQF-----FGHELSDLFPPT      46
ALPHA-CRYSTALLINACHAIN      -----MDVTIQHWPFKRTLGP--F-YPSRLF--DQF-----FGGLEFEYDLPF      39
  
```

```

HEATSHOCKPROTEINBETA-9      CPS-----VGLAERNRVAT-----MPV--LRLLDSAPAAQEDNDHARDGFMK      57
HEATSHOCKPROTEINBETA-7      -----MFSDDFGSFMRPHSEP-----LAFPARPGGAGNITLGDAYEFA      85
HEATSHOCKPROTEINBETA-3      -----DHALYALPGPT--IVDLR--KTRAAQSPVVDAAETPPREGKSHFQIL      74
HEATSHOCKPROTEINBETA-8      WPDNALRLSSAMGPTLRSGMVRP-----GPTA-TARFGVPAAGTSPFPFGPKVC      99
HEATSHOCKPROTEINBETA-2      -P-----TLYHGYYVPPRAAP-----AGEGSRAGASRLSEKGFQAF      97
HEATSHOCKPROTEINBETA-1      LG-----GSSMPGYRPLPAAAEISPAVAAPAYSRLSRQLSSGVSEIRHTADRMVRS      98
HEATSHOCKPROTEINBETA-6      CP-----TTLAPYYLR-----PSV-----ALPVAQPTDPGHFSVL      77
ALPHA-CRYSTALLINBCHAIN      ST-----SLSPFL-RR-----PSFLR-APSWHFLDSEMRLEKQFVSFN      78
ALPHA-CRYSTALLINACHAIN      LS-----STISPHY-RQ-----SLF--RTVDSGISEVRSRDRKQFVF      74
  
```

```

HEATSHOCKPROTEINBETA-9      LDAHGFAPEELVVQDQGLMVTGQQQLDVRDPERVSYRMSQVKHRLMPLSMLPTAMTC      117
HEATSHOCKPROTEINBETA-7      VDVRDQSPEDIIVTTSNNHIEVRA--EKLADG--TVMNTFAHKCOLPEDVPTSVTS      139
HEATSHOCKPROTEINBETA-3      LDVVQFLPEDIIQTFFEGMLLKAQHGTRDMEHG--FISRSFTRQYKLPDGVIEKDLA      131
HEATSHOCKPROTEINBETA-8      VNVHSFKPEELMVKTDGVEVSGKHEEQDEGG--IVSKNFTKKIQLPAEDVPVTVA      156
HEATSHOCKPROTEINBETA-2      LDVSHFTPDEVTVRTVDNLLVESARHPQLDRHG--FVSREFCRTQYVLPADVDPWRVRA      134
HEATSHOCKPROTEINBETA-1      LDVNHFAPELTVKTKDGVVEITGKHEERQDEHG--YISRCFTTKYTLPPGDVPTQSS      155
HEATSHOCKPROTEINBETA-6      LDVKHFSPEEIAVKVVEGHEVHARHEPRDEHG--FVAREFHRRYRLPPGVDPAVTS      146
ALPHA-CRYSTALLINBCHAIN      LDVKHFSPEELKVKVGLGVIEVHGKHEERQDEHG--FISREFHRRYRIPADVDPLTIS      135
ALPHA-CRYSTALLINACHAIN      LDVKHFSPEDLTKVKQDDFVIEVHGKHEERQDHHG--YISREFHRRYRLPSNDQASAL      131
  
```

```

HEATSHOCKPROTEINBETA-9      CLTPSGQLWVRGQCVALLPEAQGTG-----SPRLGSLGSKASNLTR-----      159
HEATSHOCKPROTEINBETA-7      ALREDGSLTIARRHPHTEHVQQTFRTEIKI-----      170
HEATSHOCKPROTEINBETA-3      VLCHDGLLVVEVQDPVGTG-----      150
HEATSHOCKPROTEINBETA-8      SLSPEGLLIEAPQVPVPTGTFGESSFNNELPQDSQE-----VTCT--      196
HEATSHOCKPROTEINBETA-2      ALSHOGILNLEAPRGGRHLDETVN--EYVLSLLPA-----PPDPEEEEEAAIVP      182
HEATSHOCKPROTEINBETA-1      SLSPEGTLVEAPMKLAT--QSN--EITIPVTFESRAQLGGPEAAKSDETAAM--      205
HEATSHOCKPROTEINBETA-6      ALSPEGVLISQAAPASQAAPPAA--AK--      160
ALPHA-CRYSTALLINBCHAIN      SLSSDGVLTVNGPRKQV--SGP--ERTIPITREKPAVTAAPK--      175
ALPHA-CRYSTALLINACHAIN      SLSADGMLTFCGPKIQTLGDATHA--ERAIPVSREEKPTAPSPP--      173
  
```

```

Alpha-crystallinBchain: M-DIAIH-HP-WIRR-PFFPF--HS-PSRLFDQFFGEGHLLSD-LFP-TSTLSLSPFYLRP--P---S-F
Alpha-crystallinAchain: M-DVTIQ-HP-WFKR-TLGP--Y--PSRLFDQFFGEGLEFYD-LLPFLSSSTISPHY-R---Q---S-L
Heatshockproteinbeta-6: M-EIPVPVQPSWLR--ASAPLPLGSAPGRLDQRFEGGLEAE--LAALCPTTLAPYYLR--P---S--
Heatshockproteinbeta-2: M-S-GRS-V-LHP--ATAEYE-FANPSRLGEORFEGGLPEE-IL--TPTLYHGYVVR--P---R-A
Heatshockproteinbeta-1: MTERRVPFSL-LRGP-SWDPRDWPYHSRLFDQAGFLPRLPEE-WSQWLGSSWPGVYRPLPAAIESPA
Heatshockproteinbeta-8: MADGQMPFSCHYPsRLRRDPFRDPLSSRLLDGFGMDPFPDGLTASW-PDWALP-RLSSAWPGTLRSGM
Heatshockproteinbeta-7: M-SHRTS-ST-F--R-AERSFHSSSSSSSSSSSSASRALPAQ-DPP-MEKALSMFSDDF--G---S-F
Heatshockproteinbeta-3: M-A-KI--I--LRHL-IEIPVR-Y--QEEFEARGLEDCLDLHA-LYA-LPGPTIV-DLRK--T---R-A
Heatshockproteinbeta-9: M-Q-RVG-NT-FSNE-SR--V-----ASRC-PS-VG--LAERN--R--VAT-M-P--VR-----L
  
```

```

Alpha-crystallinBchain: LRAP--S-W-FD--T-GLSEMRLEKDRF--VNLDVKHFSPEELKVKVLGDVIEVHGKHE--RQDEH-GF-
Alpha-crystallinAchain: FR---T-V-LD--S-GISEVRSRDRKF--IFLDVKHFSPEELTKVKVQDDFVEIHGKHN--EQDDH-GY-
Heatshockproteinbeta-6: V-----A--L--P--VAQVPTDPGHF--VNLDVKHFSPEEIAVKVVGHEVVEHARHE--RPDEH-GF-
Heatshockproteinbeta-2: APAG--E-G-SR--A-GASELRLSEKGF--VLDVSHFSTPQCVTVRTVDNLLVESARHP--QLDRH-GF-
Heatshockproteinbeta-1: VAAPAYSRLSRQLSSGVSEIRHTADRVRSVDNHNHFAPELTVKTKDGVVEITGKHE--RQDDH-GY-
Heatshockproteinbeta-8: VPRGPTATARFGVPAEGRTPPPPFGPEKVKCVNVHFSKPEELMVKTKDGVVEVSGKHE--EQQEG-GI-
Heatshockproteinbeta-7: MRPHSEPLA-FPARPGGAGNITLGDAYEFAVDVRDPSPEIDIIVTTSNNHIEVRA---E-KLAAD-GT-
Heatshockproteinbeta-3: AQSP--P-V-DS--A-AETPPREGKSHFQILLDVVQFLPEDIITQTFEGWLLKAQHG-T-RMDEH-GF-
Heatshockproteinbeta-9: LR-D--S---P--A-AQEDNDHARDGFMKLDHAGFAPEELVVQDQGLMVTGQQQLDVRDPERVSYR
  
```

```

Alpha-crystallinBchain: ISREFHRRYRIPADVDPLTITSSSLSSDGLTVVGRPKQ--VSG--PE-R-T-IPITR-EE-KPAV-TAA-
Alpha-crystallinAchain: ISREFHRRYRLPSNVDSALSCSLADGMLTF--CGPKIQTLGDATHAE-R-A-IPVSR-EE-KP--TSA-
Heatshockproteinbeta-6: VAREFHRRYRLPPGVDPAVTSALSPEGVLSI-----Q-----A-----PASA-QA-PP--PAA-
Heatshockproteinbeta-2: VSREFCRTYVLPADVDPWRVRAALSHDGLNLLEAPRGGRHLDETVNEVYISLLPAPP-DP-EEEE-EAAI
Heatshockproteinbeta-1: ISRCFTRKYTLPPGVDPDTQVSSSLPEGLTLVEAP-MP-KLATQSNIEITPVTFESRAQLGGPEAAKSDE
Heatshockproteinbeta-8: VSKNFTKILQPAEDVPVTVFASLSPEGLLIEAPQVP-PYST-FGE-S--SF--NN-EL--PQ--DSQE
Heatshockproteinbeta-7: VMNTFAHKQLPEDVPTSVTSALREDGSLTIARRHP-H--T---E-H---V---Q-QT--F---RTE-
Heatshockproteinbeta-3: ISRSFTRQYKLPDGVIEKDLSAVLCHDGLLVV-----E-----V---K-D---P---VG-
Heatshockproteinbeta-9: MSQKVHRKM-LPSNLSPAMTCLTPSGQLWVRGQCV--LAL--PEAQ-T-GPSRLGS-LGS--KASN
  
```

```

Alpha-crystallinBchain: PK-K
Alpha-crystallinAchain: PS-S
Heatshockproteinbeta-6: -A-K
Heatshockproteinbeta-2: VE-P
Heatshockproteinbeta-1: TAAK
Heatshockproteinbeta-8: VTCT
Heatshockproteinbeta-7: IK-I
Heatshockproteinbeta-3: -T-K
Heatshockproteinbeta-9: LT-R
  
```

[illegible]

- 36

Résultats : Comparaison

CLUSTAL O(1.2.4) MULTIPLE SEQUENCE ALIGNMENT	
HEATSHOCKPROTEINBETA-9NQ---RVGNTFSNESR-----VASR
HEATSHOCKPROTEINBETA-7	MSHTT-----SSTF---RAEFSHSSSSSSSSSSASRALPAQDPHEKALS--
HEATSHOCKPROTEINBETA-3	MAKILLNHL-----EIPVYQ---EEF-----EAGLEDCQ---
HEATSHOCKPROTEINBETA-8	NAGQWPFSCRP-SRI--RQFPRDSLSILL---DQG-----KQHPQDQTLAS
HEATSHOCKPROTEINBETA-2	MSGSVPMHP-----ATAEYFANPSQLF---EQD-----FEGGLPEELTL
HEATSHOCKPROTEINBETA-1	NTERVFFSLRGPSS---DPFRDYNHSILF---DQA-----FGLRLPEEKQW
HEATSHOCKPROTEINBETA-6	---NEIPVPQPSMLRRASAPLGLSARILF---DQR-----FEGGLEALAL
ALPHA-CRYSTALLINCHAIN	---MDIAIHMPKRRPFPF---FHSRILF---DQF-----FGEHLSDLPFT
ALPHA-CRYSTALLINCHAIN	---MDVIQHMPKRTLLG---F-YPSRILF---DQF-----FGEGLFYDLPF
HEATSHOCKPROTEINBETA-9	CPS-----VGLAEIRNVAT-----MPV---DLLRSDPAQEDNDHAGDQMK
HEATSHOCKPROTEINBETA-7	-----HFSDFGSPMRHSEP-----LAFPARPGAGNKTLDGAYEFA
HEATSHOCKPROTEINBETA-3	-----DHAYALGPT--IVDLR---KTAAGSPVDSAAETPRGKSHFQIL
HEATSHOCKPROTEINBETA-8	WPMALPRLSSAMFGLRSGVPR-----QPTA-TAIFGVPAEGRTPPPFGEPMKVC
HEATSHOCKPROTEINBETA-2	---P-----TLHYGYVPRPAAP-----AGESSAGASELSEGIGQAF
HEATSHOCKPROTEINBETA-1	LG-----GSDPVRVPRPAALZEPAAVPAVPSMLSRQLSSGSEIDHADRVS
ALPHA-CRYSTALLINCHAIN	LP-----TTLAPYLLA-----PSV---ALPVAQVPTDQHWVSL
ALPHA-CRYSTALLINCHAIN	ST-----SLSPYL-RP-----PSFLR-APSWFDTGLSEMILKDFSVN
ALPHA-CRYSTALLINCHAIN	LS-----STISPY-RQ-----SLF---RTVLDGSEIVSDDDQFVIF
HEATSHOCKPROTEINBETA-9	LDAGFAPEELVQVQGGKLVYTGQQLDVQDFEYVYSRQVHKMLPSNLSPMTAC
HEATSHOCKPROTEINBETA-7	VGVQVDFPEDIIVTSSNNIVEA---ELALAG---TYNFAFKQLPQVDQITVTS
HEATSHOCKPROTEINBETA-3	LDVQVDFPEDIQITFEGALLTAQDGTMDHG---FTSRFTQYQLPDQETDLASA
HEATSHOCKPROTEINBETA-8	VWHSFPEELMWKTKDGYVESSGHEEQGSG---YVSKNFTKIKLPAEVPVTFPA
HEATSHOCKPROTEINBETA-2	LDVSHFTPDQVTVRTVNDLVSARHPQLDQHG---FVSRFCRTYVLPADVPKWVSA
HEATSHOCKPROTEINBETA-1	LDVSHFPADELVTKTGQVETTGQHEEQQHG---YVSRFCRTYVLPQVDQVSS
HEATSHOCKPROTEINBETA-6	LDVSHFSPEEAVVGVGEVHARHEEQDGH---FVAREFHRVYLPQVDPAAVTS
ALPHA-CRYSTALLINCHAIN	LDVSHFSPEELVXVLDGVTEVHGHEEQDGH---FTSRFHRVYLPQVDPDLTTS
ALPHA-CRYSTALLINCHAIN	LDVSHFSPEELTVYQDQFVEIGHHEEQDGH---FTSRFHRVYLPQVDSGLSC
HEATSHOCKPROTEINBETA-9	CLTPSGQWQGVQVALPEAQTP-----SPRLGSLGKASNLTR-----
HEATSHOCKPROTEINBETA-7	ALVDEGSLTIRARRPHTEHVQQ-TFRTEIK-----
HEATSHOCKPROTEINBETA-3	VLCQGLLVVEVPDVKTC-----
HEATSHOCKPROTEINBETA-8	SLSPGELLTAEPQVPPVYTFSGSPNMLPQDGE-----VYCT-----
HEATSHOCKPROTEINBETA-2	ALSHQGLNLNLAEPGGHGLDTEVN--EYVSLPA-----PPDEEEEAALVEP
HEATSHOCKPROTEINBETA-1	SLSPGCTLYEAPMKPLAT--QSN--EITIPYTFESRAQLGGEPAKSDETAAK--
HEATSHOCKPROTEINBETA-6	ALSPGVLSTQAAPASQAAPPAA--AK-----
ALPHA-CRYSTALLINCHAIN	SLSDGVLTVNPKQV---SGP--EETPTITREEKPAVTAAPKK--
ALPHA-CRYSTALLINCHAIN	SLSADGMLTFCGKIQTGLDATHA--EATVPSHEKPTSPAPSS-----
Alpha-crystallinBchain:	M DIAIH-HP-WKRR-PFFPF--HS-PSRLFDQFFGEHLLSD-LFP-TSTLSPEFYLRP--P----S-F
Alpha-crystallinAchain:	M DVTIQ-HP-WFKR-TLQPF--Y--PSRLFDQFFGEGLFEYD-LLPFLSSTISPYR--R--Q---S-L
Heatshockproteinbeta-6:	M EIVPVPQPSMLRR-ASAPLPLGASAPRLFDQFQFGEGLLEAE-LAALCPTLLAPVYLR--P----S-
Heatshockproteinbeta-2:	M S-GRS-VP-MHP-ATAEYE-FANPSRLGEQFEGGLLEF-TL--TPLYHGYVPR--P---R-A
Heatshockproteinbeta-1:	M EIRVPFSL-LRGP-SNDPFRDYNHSRLDDQAFGLRLPEE-MSQWLGSSNGYVRLPPLAATESPA
Heatshockproteinbeta-8:	M DGMQPFSCYPSRLRRDPFRDPSLSRLLDDGFGDPFDOLTASW-PNALP-LRSSAMFTLRSGL
Heatshockproteinbeta-7:	M SHRTS-ST-F-R-AERSFHSSSSSSSSSSSASRALPAQ-DPP-MEKALSMFSDDF--G---S-F
Heatshockproteinbeta-3:	M A-KI--I--LRHL-TEIPVR-Y--QEEFEARGLEDCLDHA-LYA-LPGTIV-DRK--T---R-A
Heatshockproteinbeta-9:	M Q-RVG-NT-FSNE-SR--V-----ASRC-PS-VG--LAERN--R--VAT-M-P--VR-----L
Alpha-crystallinBchain:	LRAP--S-W-FD--T-GLSEMRLEKDRF--VNLDVKHFSPEELKVKVLGDVIEVHGKHE--RQDEF-GF-
Alpha-crystallinAchain:	FR---T-V-LD--S-GISEVRSDRDKRIFLVDVKHFSPEELTVKVKVQDDFVEIHGKHE--RQDDF-GY-
Heatshockproteinbeta-6:	V----A--L--P--VAQVPTDQGHFVLVDVKHFSPEELIAVKVGEHVEVHARHE--RQDEF-GF-
Heatshockproteinbeta-2:	APAG--E-G-SR--A-GASELRLSEGKFAF-----TVRTVNDLVSARHP--KLDRI--GF-
Heatshockproteinbeta-1:	VAAAPYSAALRQLSSGSEIHTADRNVSLDWNHFAPEELTKTKDGYVEVSGHE--RQDEF-GY-
Heatshockproteinbeta-8:	VVRGPTATARGVPAEGRTPPPFGEPMKVCNVHVSFKPEELMWKTKDGYVEVSGHE--KQEGG-GI-
Heatshockproteinbeta-7:	MRHSEPLA-FPARPGAGNKTLDGAYEFAVDVDFSPEDIIVTSSNNHIVERA---E-KLAD-GT-
Heatshockproteinbeta-3:	AOSP--P-V-DS--A-AETPPREGKSHFOILLDVQFDPEDIIQTFEGWLLIAQHG--T-RMDEH-GF-
Heatshockproteinbeta-9:	LR-D-S----P-A-AQEDNDHARGDQFOMLDAHGAPEELVQVQDQWLVMTGQQ--RQDPRVSVYR
Alpha-crystallinBchain:	ISREFHRRYRIPADVPLTITSSLSDDGLTVGPRKQ--VSG--PE-R-T-IPTR-EE-KPAV-TAA-
Alpha-crystallinAchain:	ISREFHRRYRLPSNVDSQSLCSLSADGMLTFGPKIQTGLDATHAE-R-A-IPVSR-EE-KP---TAA-
Heatshockproteinbeta-6:	VAREFHRRYRLPGVDPAVAVTSALSPEGLST----Q---A-----A--PASA-QA-PP---PASA-
Heatshockproteinbeta-2:	VSRFCRTYVLPADVQPMVRAALSHDGLNLAPRGRHLDETVNEVYISLLPAPP-DP-EEEE-EAAT
Heatshockproteinbeta-1:	ISRCFTKTYVLPQVDQVSSLSPEGLTAP-HP-KLATOSMETIIPVTFESRAQLGGEPAKSDG
Heatshockproteinbeta-8:	VSKNFTKIKLPAEVPVTFVFSALSPGELLTAEPQV-PYST-FGE-S--SF-INL-EL--PQ--DQEG
Heatshockproteinbeta-7:	VNMTFAHKQLPQVDQVPTVSALRDEGSLTVARRHP-H-T---E-H---V---Q-OT-F--RTE-
Heatshockproteinbeta-3:	ISRSFTQYKLPDGEVTKLDSAVLCHDGLVW-----E-----E---K-D--P---VG-
Heatshockproteinbeta-9:	MSQKVHRKM-LPSNLSPMTACCLTSPGQWLWQCVAA--LAL--PEAQ-T-GPSRLGS-LGS--KASN
Alpha-crystallinBchain:	PK-K
Alpha-crystallinAchain:	PS-S
Heatshockproteinbeta-6:	A-K
Heatshockproteinbeta-2:	VE-P
Heatshockproteinbeta-1:	TAAK
Heatshockproteinbeta-8:	VTC
Heatshockproteinbeta-7:	IK-I
Heatshockproteinbeta-3:	T-K
Heatshockproteinbeta-9:	LT-R

- Arbre guide différent
- Modèle de Markov caché

Résultats : Comparaison

CLUSTAL O(1.2.4) MULTIPLE SEQUENCE ALIGNMENT		
HEATSHOCKPROTEINBETA-9NQ-----RVGNTFSNESR-----VVASR	17
HEATSHOCKPROTEINBETA-7	MSHTT-----SSTF---RAEKSFHSSSSSSSSSSASRALPAQDQMEKALS--	36
HEATSHOCKPROTEINBETA-3	MAKITLHLLT-----EIPVYQ---EEF-----EAELEDCR---	40
HEATSHOCKPROTEINBETA-8	NAIQQWPSCHRP-SRL--RQDFRDSLSESL--DQF-----VQDHPDQLTAS	47
HEATSHOCKPROTEINBETA-2	MSGSVVMHP-----ATAEYFANPSILG---EQD-----FEGGLPEELTL	44
HEATSHOCKPROTEINBETA-1	NTERVFFSLRGPSS---DFFRDVYSRILF---DQA-----FGLRLPEEKQW	45
HEATSHOCKPROTEINBETA-6	---NEIPVQVPSMLRRASAPGLSAPGILF---DQR-----FEGGLEALAL	45
ALPHA-CRYSTALLINCHAIN	---MDIAIHWMIRRPFPF---FHSRILF---DQF-----FGEHLESDLPFT	40
ALPHA-CRYSTALLINCHAIN	---MDVIQHWPKRITLGP---F-VYSRILF---DQF-----FGEGLFYDLPF	39
HEATSHOCKPROTEINBETA-9	CPS-----VGLAEIRNVAT-----MPV---DLLRDSAPQEDNDHARDGQMK	57
HEATSHOCKPROTEINBETA-7	---HFSDFGSPMRHSEP-----LAFPARPGAGNKLTDGAYEFA	85
HEATSHOCKPROTEINBETA-3	---DHAYALPGPT--IVDLR---KTAAGSPVDSAAETPRGKSHFQIL	74
HEATSHOCKPROTEINBETA-8	WPMALPRLSSAMFGLRQVPR-----QPTA-TAIRGVPAEGRTPPPFGPEKVC	99
HEATSHOCKPROTEINBETA-2	---P-----TLHYGYVVRPAAP-----AGESSAGASLSEGLQDQAF	77
HEATSHOCKPROTEINBETA-1	LG-----GSDPVRVPIPAALPAEPAVAPVPSMLSRQLSSGSSDHTADNRVS	98
HEATSHOCKPROTEINBETA-6	CP-----TTLAPPYLA-----PSV---ALPVAQVPTDQHWVSL	77
ALPHA-CRYSTALLINCHAIN	ST-----SLSPYL-RP-----PSFLR-APSMFDTGLSEMILKDFSVN	78
ALPHA-CRYSTALLINCHAIN	LS-----STISPY-RQ-----SLF---RTVLDGISEVSDDDKVFVF	74
HEATSHOCKPROTEINBETA-9	LDANGFAPEELVQVQDQKLVYTGQQLDVQDFEIVSYRSGVHWKMLPSNLPTAMTC	117
HEATSHOCKPROTEINBETA-7	VQVDFPDEIITVSNNIEVFA---ELALAG---TYMVFANKQLPQVDITDYS	139
HEATSHOCKPROTEINBETA-3	LDVQVFPEDIQITFEGHLLTAQDCTMDHSG---FISRSFQVLPQDVEIDLSA	131
HEATSHOCKPROTEINBETA-8	VWHSFPEELMWKTKDGYVESSGHEEQGSG---VYSKNTFKIQLPAEVPVTFPA	156
HEATSHOCKPROTEINBETA-2	LDVSHFTPDQVTVTDNLLEVSARHPQLDQHG---FVSRFCRTYVLPADVPQWVA	134
HEATSHOCKPROTEINBETA-1	LDVSHFPADELVTKTDGVEITGQHEEQGHEG---VYSIRCFRTYVLPQVDQVSS	155
HEATSHOCKPROTEINBETA-6	LDVSHFSPPEIAVWVGEHVEHARHESPDGSG---FVAREFHRHYVLPQVDPAAVTS	134
ALPHA-CRYSTALLINCHAIN	LDVSHFSPPEELVQVQDGYVEHSGHEEQGHEG---FISREFHRYVLPADVDPLTTS	135
ALPHA-CRYSTALLINCHAIN	LDVSHFSPPEELTVQVQDGYVEHSGHEEQGHEG---FISREFHRYVLPADVDPLTTS	131
HEATSHOCKPROTEINBETA-9	CLTPSGQWQGVQVALPEAQTP-----SPRLGSLGKASNLTR-----	159
HEATSHOCKPROTEINBETA-7	ALVDEGSLTIRARRPHTEHVQQ-TFRTEIKI-----	170
HEATSHOCKPROTEINBETA-3	VLCOGILLVVEVPDVKTC-----	150
HEATSHOCKPROTEINBETA-8	SLSPGELLTAPQVPPVYTFSGSPNMLPDQSG---VYCT-----	196
HEATSHOCKPROTEINBETA-2	ALSHOGELNLAPGGHGLDTEVN--EYVSLPA-----PPDKEEAAEVEP	182
HEATSHOCKPROTEINBETA-1	SLSPGELTVEAPMKPLAT-QSN--EITIPYTFESRAQGLGPEAAKSDETAAN--	205
HEATSHOCKPROTEINBETA-6	ALSPGELVLSQAAPASQAPPPAA--AK-----	160
ALPHA-CRYSTALLINCHAIN	SLSDGVLTVGPKQV---SGP--EETPTITREEKPAVTAAPKK--	175
ALPHA-CRYSTALLINCHAIN	SLSADGNLTFGPKIQIGLDATHA--EATIPVSEKPTSPAPSS-----	173
HEATSHOCKPROTEINBETA-9NQ-----RVGNTFSNESR-----VVASR	17
HEATSHOCKPROTEINBETA-7	MSHTT-----SSTF---RAEKSFHSSSSSSSSASRALPAQDQMEKALS--	36
HEATSHOCKPROTEINBETA-3	MAKITLHLLT-----EIPVYQ---EEF-----EAELEDCR---	40
HEATSHOCKPROTEINBETA-8	NAIQQWPSCHRP-SRL--RQDFRDSLSESL--DQF-----VQDHPDQLTAS	47
HEATSHOCKPROTEINBETA-2	MSGSVVMHP-----ATAEYFANPSILG---EQD-----FEGGLPEELTL	44
HEATSHOCKPROTEINBETA-1	NTERVFFSLRGPSS---DFFRDVYSRILF---DQA-----FGLRLPEEKQW	45
HEATSHOCKPROTEINBETA-6	---NEIPVQVPSMLRRASAPGLSAPGILF---DQR-----FEGGLEALAL	45
ALPHA-CRYSTALLINCHAIN	---MDIAIHWMIRRPFPF---FHSRILF---DQF-----FGEHLESDLPFT	40
ALPHA-CRYSTALLINCHAIN	---MDVIQHWPKRITLGP---F-VYSRILF---DQF-----FGEGLFYDLPF	39
HEATSHOCKPROTEINBETA-9	CPS-----VGLAEIRNVAT-----MPV---DLLRDSAPQEDNDHARDGQMK	57
HEATSHOCKPROTEINBETA-7	---HFSDFGSPMRHSEP-----LAFPARPGAGNKLTDGAYEFA	85
HEATSHOCKPROTEINBETA-3	---DHAYALPGPT--IVDLR---KTAAGSPVDSAAETPRGKSHFQIL	74
HEATSHOCKPROTEINBETA-8	WPMALPRLSSAMFGLRQVPR-----QPTA-TAIRGVPAEGRTPPPFGPEKVC	99
HEATSHOCKPROTEINBETA-2	---P-----TLHYGYVVRPAAP-----AGESSAGASLSEGLQDQAF	77
HEATSHOCKPROTEINBETA-1	LG-----GSDPVRVPIPAALPAEPAVAPVPSMLSRQLSSGSSDHTADNRVS	98
HEATSHOCKPROTEINBETA-6	CP-----TTLAPPYLA-----PSV---ALPVAQVPTDQHWVSL	77
ALPHA-CRYSTALLINCHAIN	ST-----SLSPYL-RP-----PSFLR-APSMFDTGLSEMILKDFSVN	78
ALPHA-CRYSTALLINCHAIN	LS-----STISPY-RQ-----SLF---RTVLDGISEVSDDDKVFVF	74
HEATSHOCKPROTEINBETA-9	LDANGFAPEELVQVQDQKLVYTGQQLDVQDFEIVSYRSGVHWKMLPSNLPTAMTC	117
HEATSHOCKPROTEINBETA-7	VQVDFPDEIITVSNNIEVFA---ELALAG---TYMVFANKQLPQVDITDYS	139
HEATSHOCKPROTEINBETA-3	LDVQVFPEDIQITFEGHLLTAQDCTMDHSG---FISRSFQVLPQDVEIDLSA	131
HEATSHOCKPROTEINBETA-8	VWHSFPEELMWKTKDGYVESSGHEEQGSG---VYSKNTFKIQLPAEVPVTFPA	156
HEATSHOCKPROTEINBETA-2	LDVSHFTPDQVTVTDNLLEVSARHPQLDQHG---FVSRFCRTYVLPADVPQWVA	134
HEATSHOCKPROTEINBETA-1	LDVSHFPADELVTKTDGVEITGQHEEQGHEG---VYSIRCFRTYVLPQVDQVSS	155
HEATSHOCKPROTEINBETA-6	LDVSHFSPPEIAVWVGEHVEHARHESPDGSG---FVAREFHRHYVLPQVDPAAVTS	134
ALPHA-CRYSTALLINCHAIN	LDVSHFSPPEELVQVQDGYVEHSGHEEQGHEG---FISREFHRYVLPADVDPLTTS	135
ALPHA-CRYSTALLINCHAIN	LDVSHFSPPEELTVQVQDGYVEHSGHEEQGHEG---FISREFHRYVLPADVDPLTTS	131
HEATSHOCKPROTEINBETA-9	CLTPSGQWQGVQVALPEAQTP-----SPRLGSLGKASNLTR-----	159
HEATSHOCKPROTEINBETA-7	ALVDEGSLTIRARRPHTEHVQQ-TFRTEIKI-----	170
HEATSHOCKPROTEINBETA-3	VLCOGILLVVEVPDVKTC-----	150
HEATSHOCKPROTEINBETA-8	SLSPGELLTAPQVPPVYTFSGSPNMLPDQSG---VYCT-----	196
HEATSHOCKPROTEINBETA-2	ALSHOGELNLAPGGHGLDTEVN--EYVSLPA-----PPDKEEAAEVEP	182
HEATSHOCKPROTEINBETA-1	SLSPGELTVEAPMKPLAT-QSN--EITIPYTFESRAQGLGPEAAKSDETAAN--	205
HEATSHOCKPROTEINBETA-6	ALSPGELVLSQAAPASQAPPPAA--AK-----	160
ALPHA-CRYSTALLINCHAIN	SLSDGVLTVGPKQV---SGP--EETPTITREEKPAVTAAPKK--	175
ALPHA-CRYSTALLINCHAIN	SLSADGNLTFGPKIQIGLDATHA--EATIPVSEKPTSPAPSS-----	173

- Arbre guide différent
- Modèle de Markov caché
- pénalité de gap constante

Conclusion



- Trois étapes implémentées

Conclusion



- Trois étapes implémentées



- Complexité temporelle

Conclusion



- Trois étapes implémentées



- Complexité temporelle



- Des résultats satisfaisants