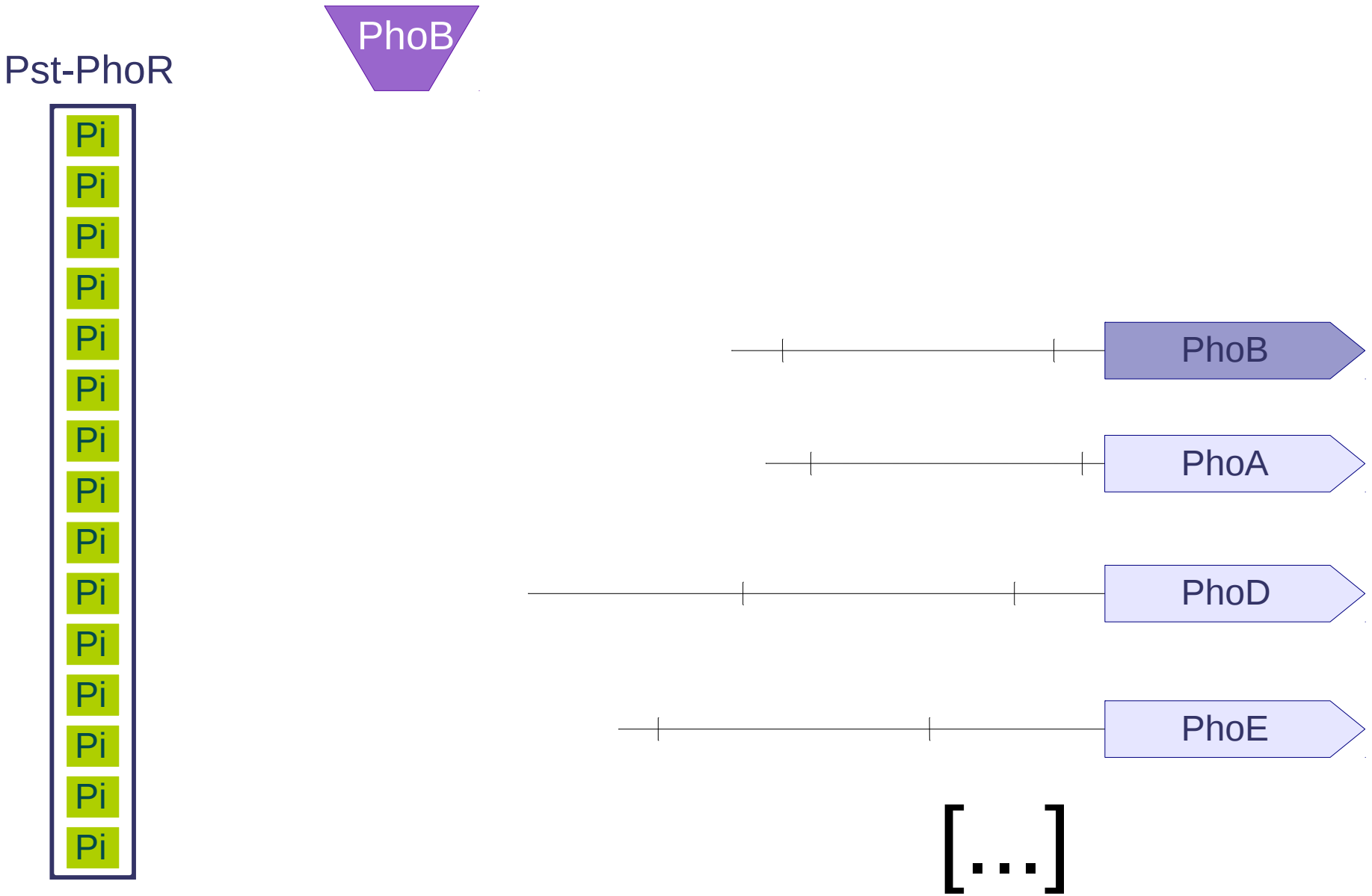
Projet REGULON

Mélissa Cardon – Nika Abdollahi

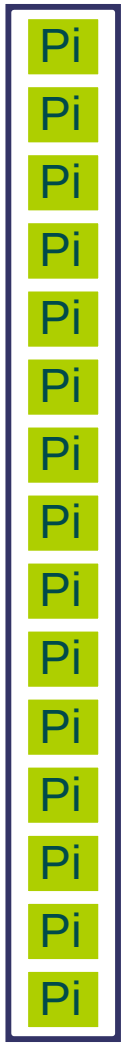
14 décembre 2016
M2 BIM – GENOM
UPMC

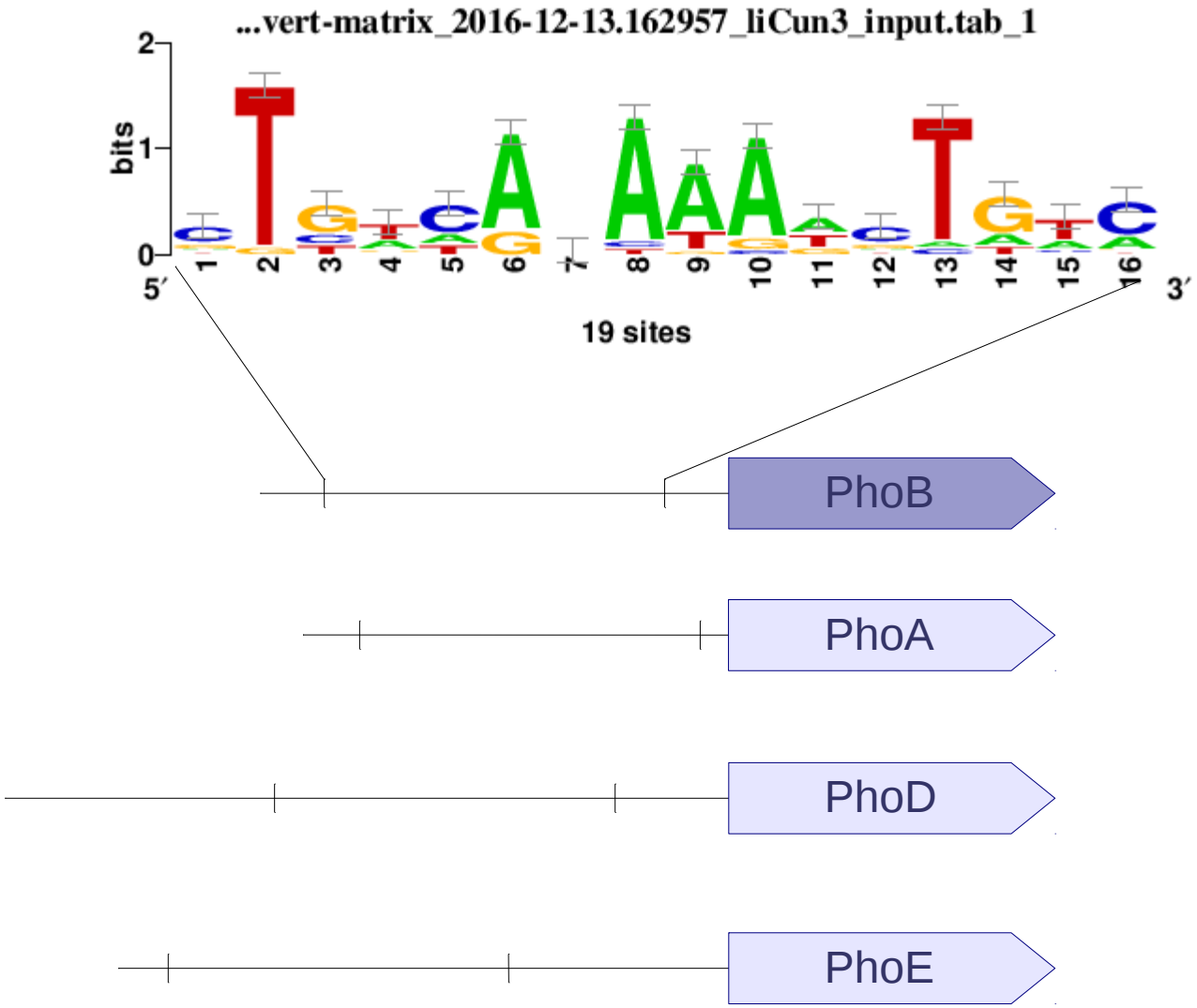
Régulon Pho


PhoB
Pst-PhoR
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
Pi
PhoB
PhoA
PhoD
PhoE
[...]

Régulon Pho

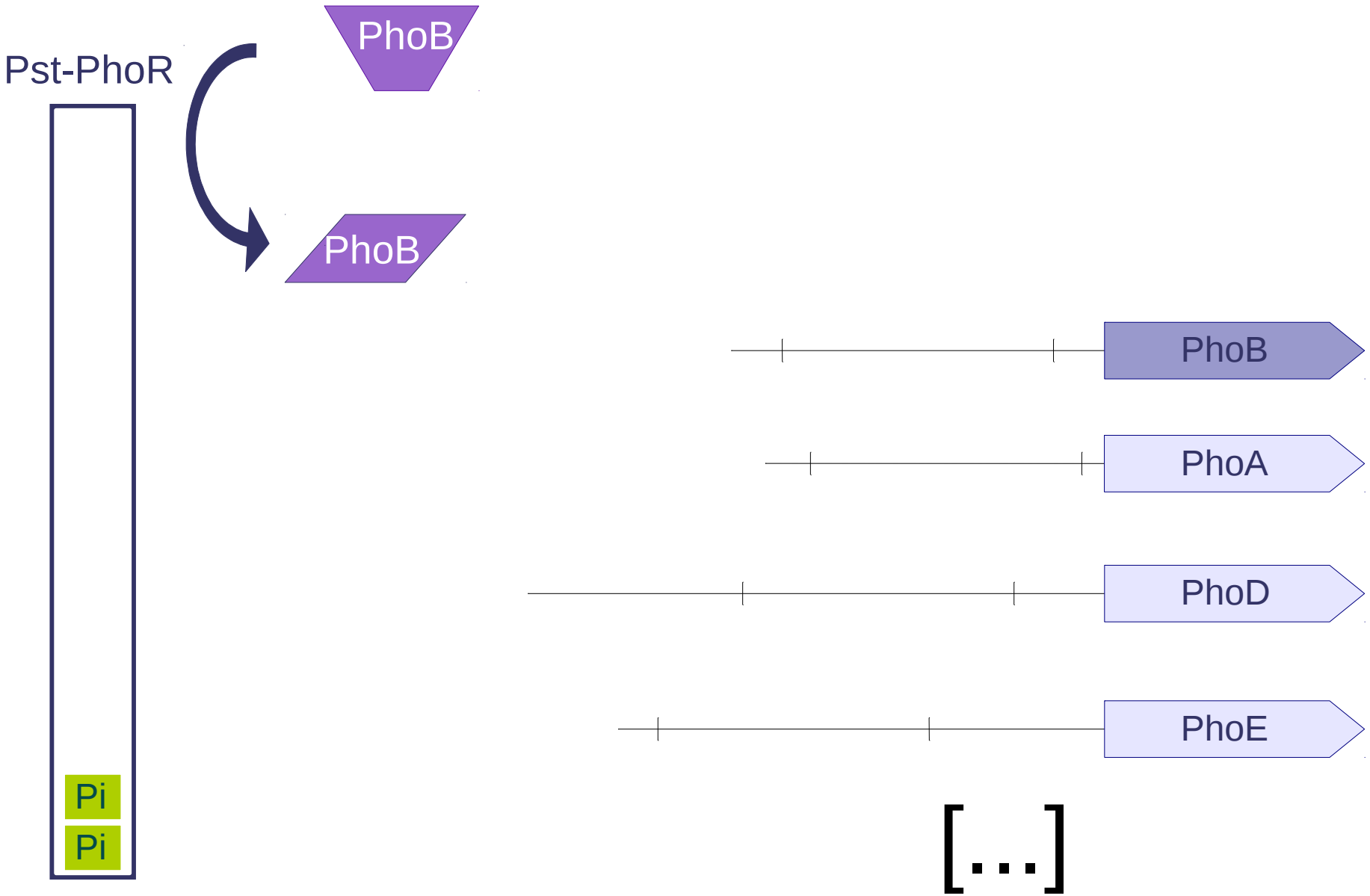
Pst-PhoR



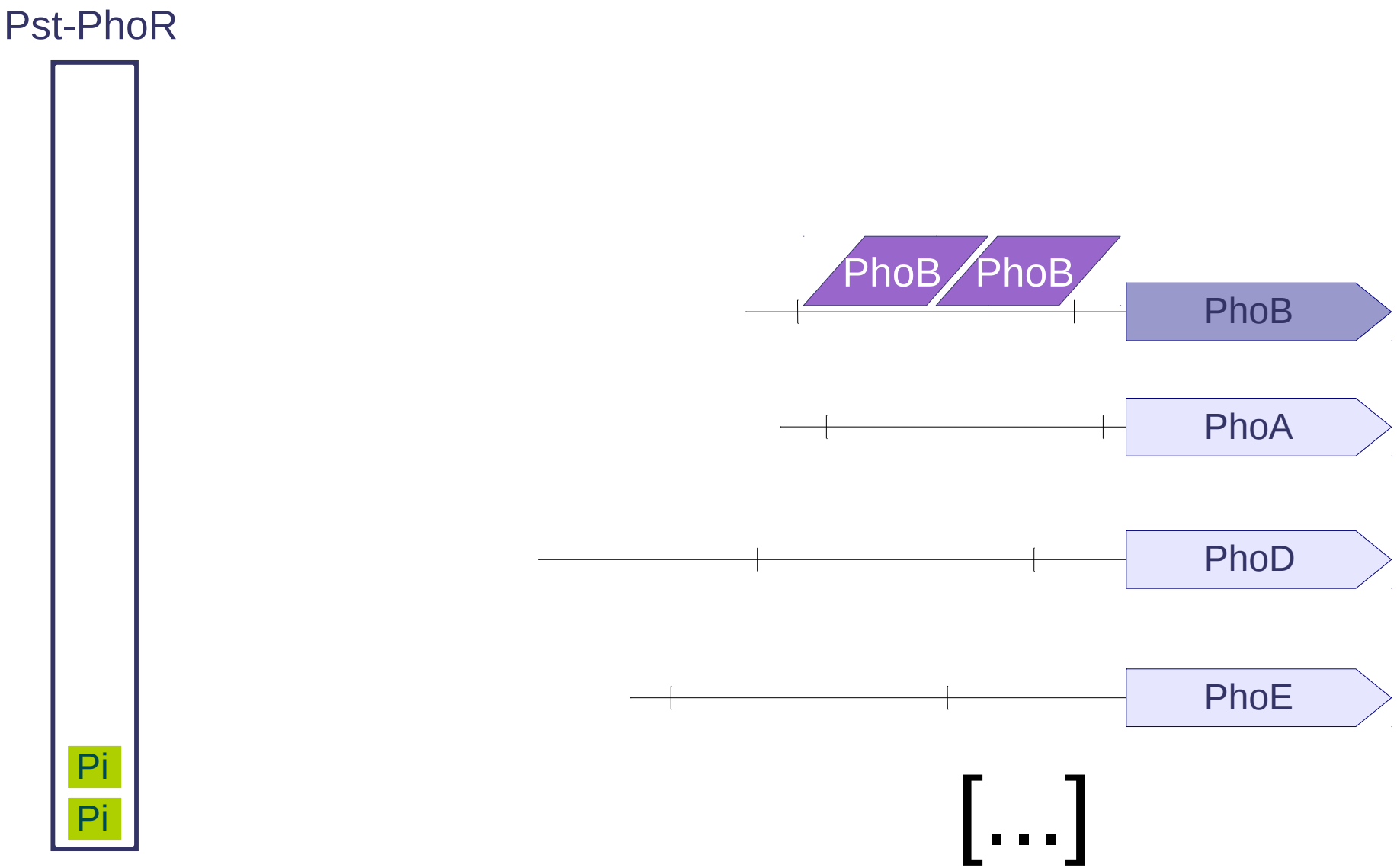


[...]

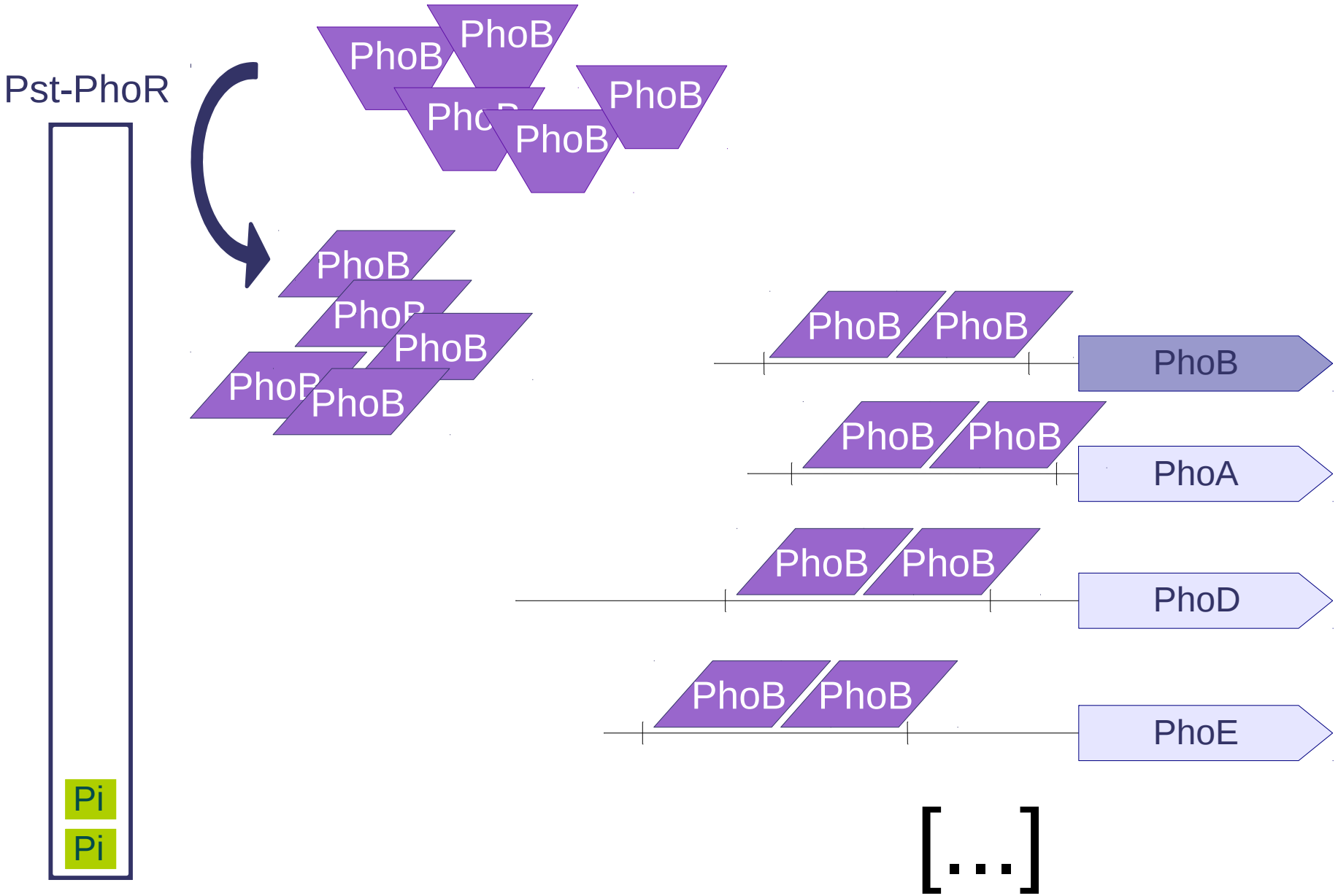
Régulon Pho



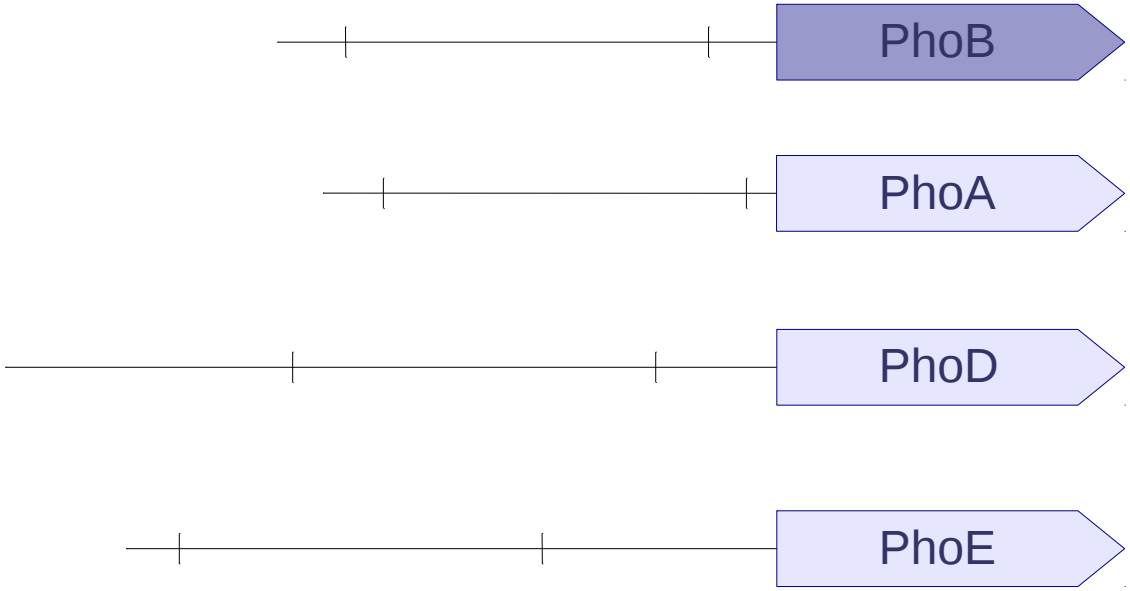
Régulon Pho



Régulon Pho

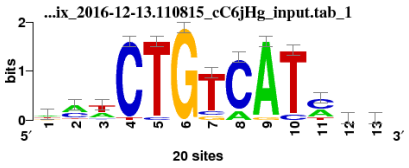

Pst-PhoR
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoB
PhoA
PhoB
PhoB
PhoD
PhoB
PhoB
PhoE
Pi
Pi
[...]

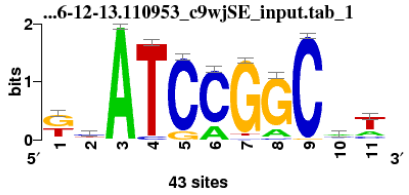
Régulon Pho

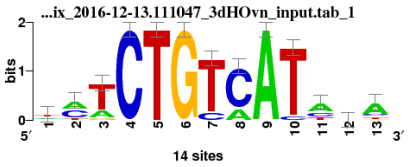


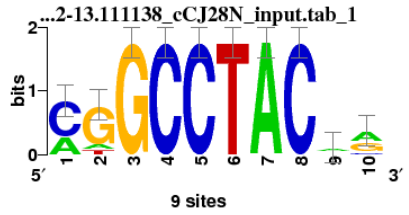
[...]

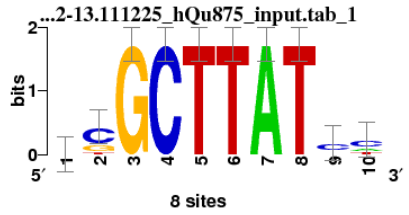
Base de données (86 FT)

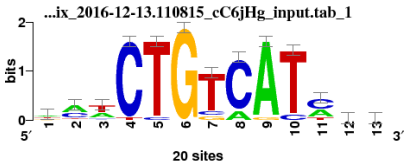


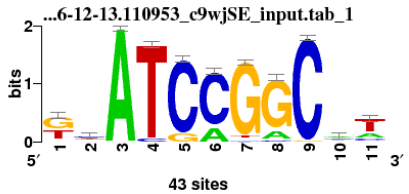


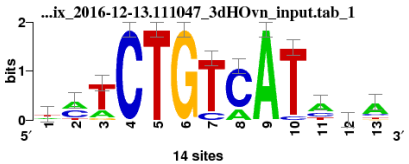


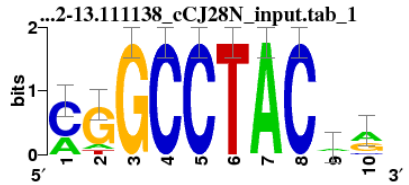


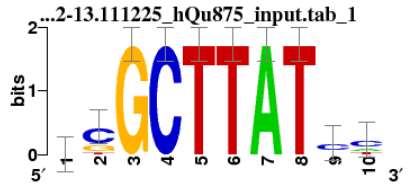










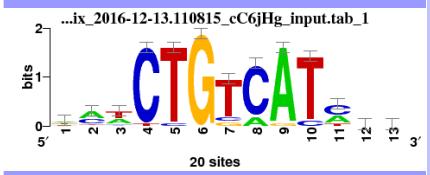
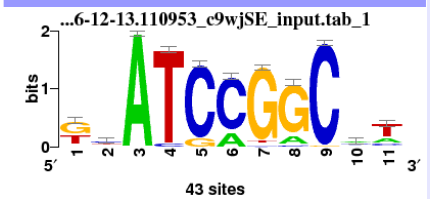
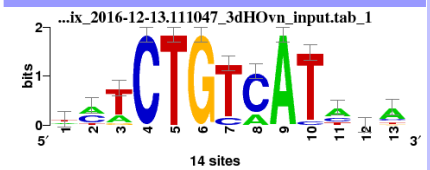
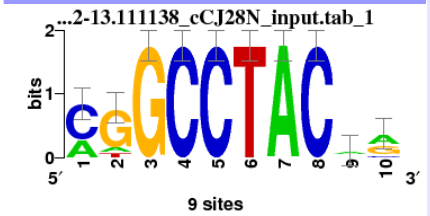
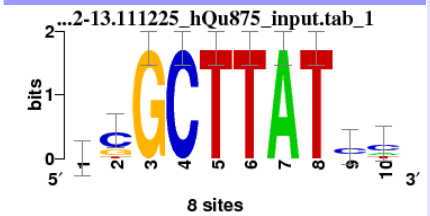


Base de données (86 FT)

- Alignement Local Smith-Waterman

- Plusieurs métriques :

Pearson correlation coefficient (PCC)	$PCC(X,Y) = \frac{\sum_{b=A}^{T}(f_X(b)-\bar{f}_X)\cdot(f_Y(b)-\bar{f}_Y)}{\sqrt{\sum_{b=A}^{T}(f_X(b)-\bar{f}_X)^2\cdot\sum_{b=A}^{T}(f_Y(b)-\bar{f}_Y)^2}}$
Chi-square (pCS) (1 − *p-value of*)	$\chi_3^2(X,Y) = \sum_{K=\{X,Y\}}\sum_{b=A}^{T}\frac{(n_K(b)-n_K^e(b))^2}{n_K^e(b)}$
Average Kullback–Leibler (AKL)	$AKL(X,Y) = 10 - \frac{\sum_{b=A}^{T} f_X(b)\cdot log\frac{f_X(b)}{f_Y(b)} + \sum_{b=A}^{T} f_Y(b)\cdot log\frac{f_Y(b)}{f_X(b)}}{2}$
Sum of squared distances (SSD)	$SSD(X,Y) = 2 - \sum_{b=A}^{T}(f_X(b)-f_Y(b))^2$
Average log-likelihood ratio (ALLR)	$ALLR(X,Y) = \frac{\sum_{b=A}^{T} n_X(b)\cdot log\frac{f_Y(b)}{p_{ref}(b)} + \sum_{b=A}^{T} n_Y(b)\cdot log\frac{f_X(b)}{p_{ref}(b)}}{\sum_{b=A}^{T}(n_X(b)+n_Y(b))}$
ALLR with lower limit (ALLR_LL)	Same as above, but a lower limit of −2 is imposed on the score (see text)

	PCC	SSD	AKL
	**PhoB** SlyA NarP RscB TorR ...	**CysB** OmpR ArgP CRP ArcA ... PhoB : 57/86	**OmpR** XylR ArcA Fis PhoP ... PhoB : 42/86
	**SlyA** RelBRelE HipB MetR IclR ... PhoB : 38/86	**FlhDC** CysB AraC HipB Dan ... PhoB : 52/86	**PhoP** AscG Lrp RcsAB HNS ... PhoB : 27/86
	**PhoB** RcsB TorR HNS RelBRelE ...	**PhoB** CysB OmpR ArgP ArcA ...	**OmpR** CysB CRP ArcA XylR ... PhoB : 73/86
	**IHF** AsnC RstA NarP NanR ... PhoB : 62/86	**NanR** EvgA NarP ModE NtrC ... PhoB : 38/86	**EvgA** DeoR NtrC GadW SoxS ... PhoB : 53/86
	**ExuR** **NanR** RelBRelE DgsA HipB ... PhoB : 50/86	**Nac** ModE ArgP GntR DnaA ... PhoB : 57/86	**GntR** ModE Fur Nac Fis ... PhoB : 75/86

Analyse de signaux

Base de données (304 PSSMs)

Matrice d`affinite

Methode de clustering : Affinity propagation

Analyse de signaux

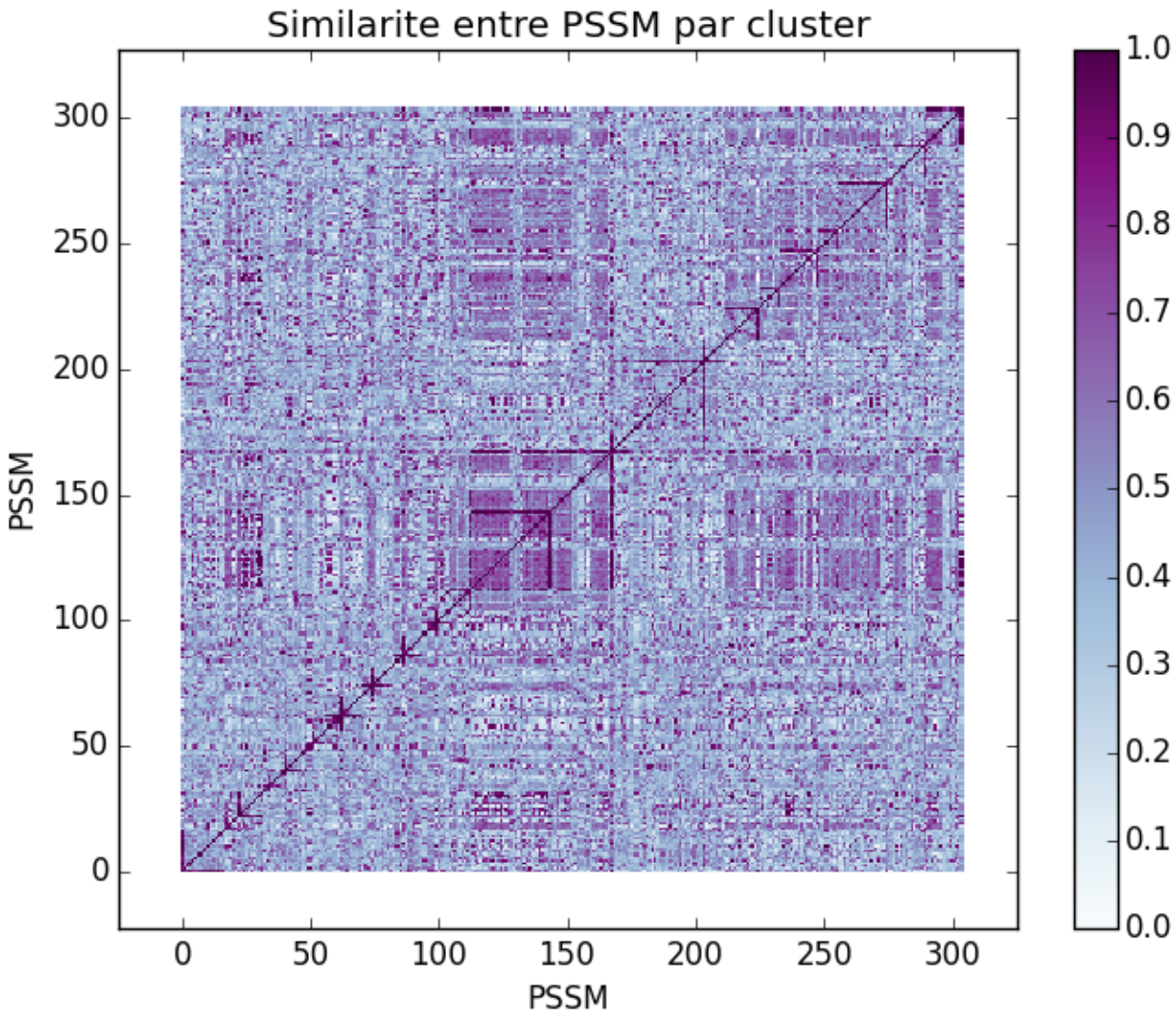
16 clusters

PhoB dans le meme cluster que
PhoA, PhoX, PhoD, ArgR, Lrp, IHF, GadE, GcvA

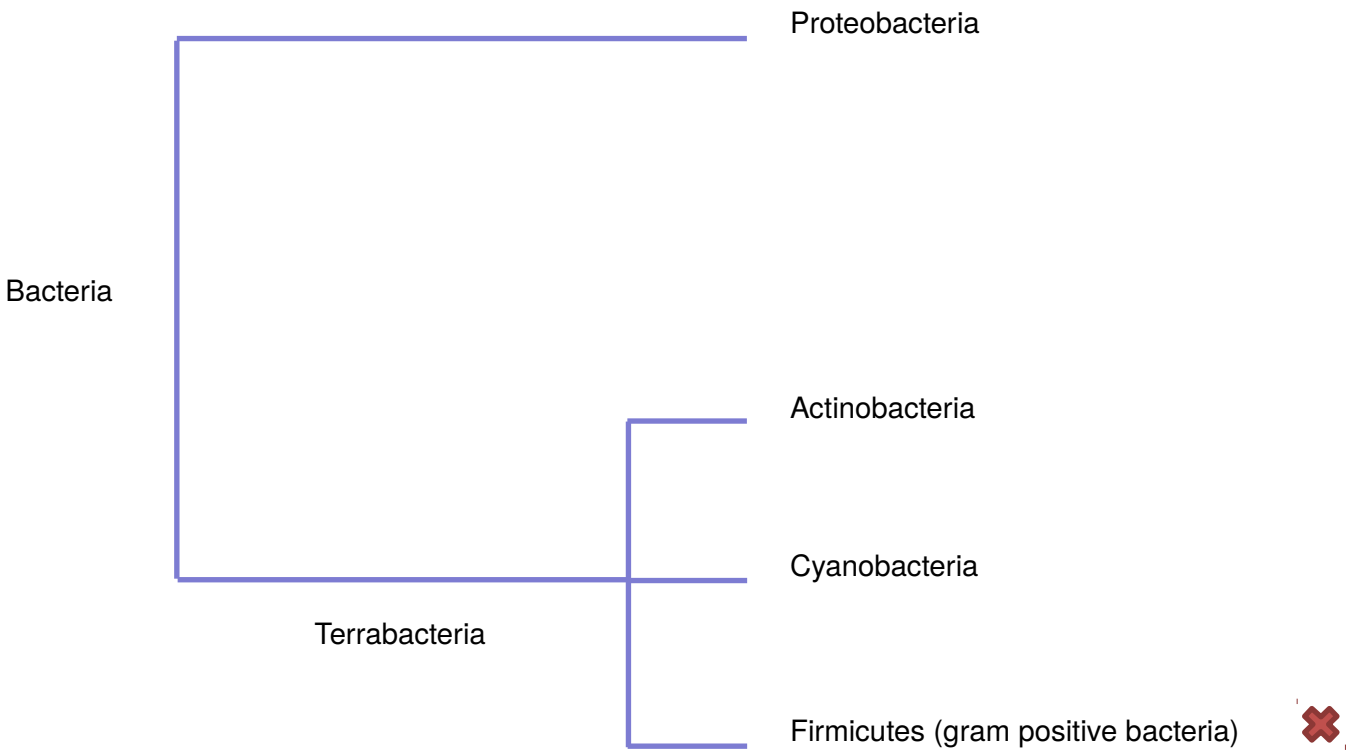
Séquences régulatrices
-Actinobacteria
-Cyanobacteria
-Proteobacteria

Motifs facteurs de transcription *E. Coli*

Similarite de signaux



Conservation de signaux



Merci de votre attention