

BacTrans²

Une plate-forme logicielle d'inférence de connaissances sur les promoteurs forts des génomes procaryotes

<http://www.sciences.univ-nantes.fr/lina/bioserv/BacTrans2/>

Porteur de projet : Christine Sinoquet

Historique

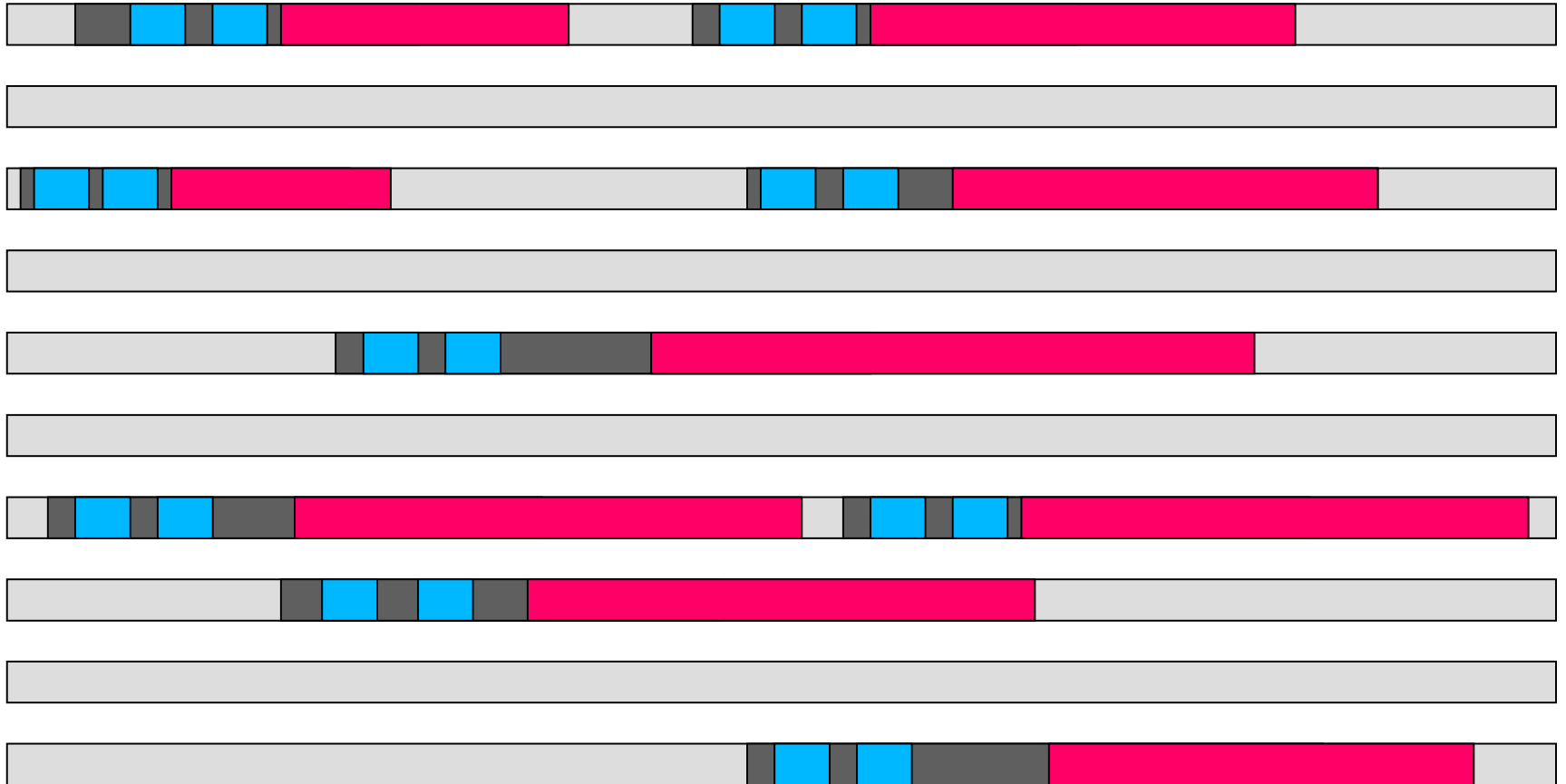
- collaboration entre le Lina et le Laboratoire de Biotechnologie de Nantes
"Biotechnologie, Biocatalyse et Biorégulation", C.N.R.S. UMR 6204,
Pr Vehary Sakanyan, Frédérique Braun
- décembre 2002 – décembre 2006
- contribution au développement
C. Sinoquet, G. Blin, J. Cambefort, G. Herry, A. Drazek, S. Demey
- migration vers nouveau serveur et sécurisation site : en cours - été 2009
- supports financiers :
 - ✓ C.P.E.R Pays de la Loire Innovations technologiques et Post-génomique – CDD 9 mois S. Demey
 - ✓ Ouest-Genopole – CDD 9 mois S. Demey

Motivations

- **public concerné :**
chercheurs impliqués dans l'étude des promoteurs forts des génomes procaryotes
- **objectifs :**
acquisition de connaissances fondamentales,
biotechnologies/génie génétique
- **rôle de la plate-forme :**
à partir du fichier GenBank (NCBI) d'un génome :
 - identification des gènes associés à un promoteur fort putatif
 - description de ces gènes au moyen de critères donnés
 - visualisations graphiques, génomique comparative

Qu'est-ce qu'un promoteur ?

génom



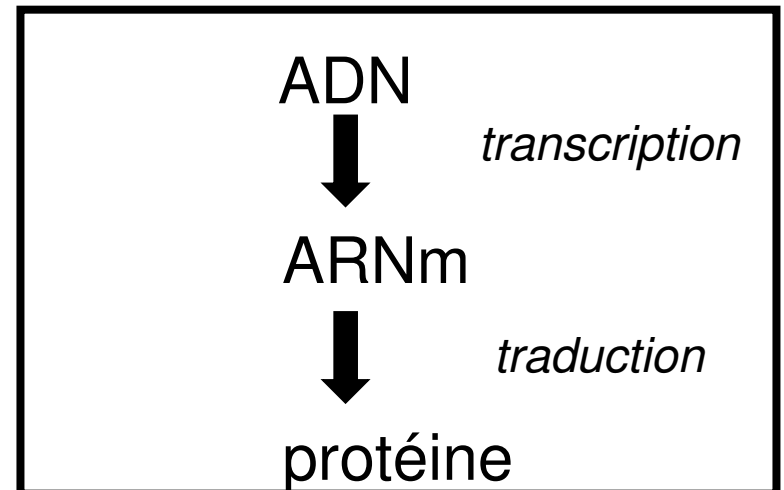
gène



région régulatrice amont

région codante

synthèse protéique



gène



région régulatrice amont

région codante

La région promotrice comporte les sites de fixation d'une molécule nécessaire à la synthèse protéique



boîte -35 boîte -10

promoteur sigma 70

gène



région régulatrice amont

région codante



élément UP

boîte -35

boîte -10

promoteur fort sigma 70

- boîtes à contenus spécifiques
 - boîte -10 TATAAT
 - boîte -35 TTGACA
 - élément UP NNAAAWWTWTTTTNNAAAANN
- distances optimales entre boîtes

gène



région régulatrice amont

région codante



codon start
Shine-Dalgarno

Spécification BacTrans²

- L'utilisateur spécifie un super-motif de son choix, pour identification d'occurrences dans un génome. (par exemple, promoteur sigma 70 du modèle *Escherichia coli*)

UP element

boîte -35

boîte -10

Shine-Dalgarno

codon start

- L'utilisateur spécifie un seuil de similarité minimum et des fourchettes de distances.

UP element

0,15

boîte -35

15, 20

boîte -10

0, 200

Shine-Dalgarno

2, 10

codon start

Structure de BacTrans²

- module d'identification
- module de visualisation
- module de statistiques
- aide en ligne, téléchargement des fichiers bruts et des graphiques

Déroulement d'une session de travail

Déroulement d'une session de travail

- identification des gènes à promoteurs forts, pour un génome donné, selon des contraintes spécifiées

Recherche de promoteur fort

UP element 0,15 boîte -35 15, 20 boîte -10 0, 200 Shine-Dalgarno 2, 10 codon start

BacTrans² - Prediction Files Creation - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

http://www.sciences.univ-nantes.fr/lina/bioserv/BacTrans2/secure/Filter.cgi?identifiant=sinoquet

Démarrage Dernières nouvelles

	Searched pattern	Maximal Hamming distance
UP	nnaaawwtwttttnnaaaenn	4
-35 Box	ttgac	3
-10 Box	tataat	2
Shine Dalgarno	ggagg	0

Expert specifications for score computation

Weighting for Hamming distance contribution:

-10 Box	0.60
-35 Box	0.40
UP element left-hand half	1
UP element right-hand half	1

	Minimal distance between	Maximal distance between
UP and -35 Box	0	15
-35 Box and -10 Box	15	20
-10 Box and Shine Dalgarno	0	200
Shine Dalgarno and Start codon	2	10

Expert specifications for score computation

Weighting for nucleotide distance contribution:

Lower bound for Upper bound for Weighting for

Terminé

démarrer evaluation_cnrs_2007 Microsoft PowerPoint ... [Ne pas fermer] Wifi ... BacTrans² - Predictio... Faire une capture écr... FR 15:48

BacTrans² - Prediction Files Creation - Mozilla Firefox

http://www.sciences.univ-nantes.fr/lina/bioserv/BacTrans2/secure/Filter.cgi?identifiant=sinoquet

	Searched pattern	Maximal Hamming distance
UP	nnaaawwtwttnnaaaenn	4
-35 Box	tgac	3
-10 Box	tataat	2
Shine Dalgarno	ggagg	0

Expert specifications for score computation

Weighting for Hamming distance contribution:

-10 Box 0.60

-35 Box 0.40

UP element left-hand half 1

UP element right-hand half 1

	Minimal distance between	Maximal distance between
UP and -35 Box	0	15
-35 Box and -10 Box	15	20
-10 Box and Shine Dalgarno	0	200
Shine Dalgarno and Start codon	2	10

Expert specifications for score computation

Weighting for nucleotide distance contribution:

Lower bound for Upper bound for Weighting for

Terminé

démarrer evaluation_cnrs_2007 Microsoft PowerPoint ... [Ne pas fermer] Wifi ... BacTrans² - Predictio... Faire une capture écr... FR 15:48

Searched pattern Maximal Hamming distance

UP nnaaawwtwttnnaaaannn

4

-35 Box ttgac

3

-10 Box tataat

2

Shine Dalgarno ggagg

0

	Minimal distance between	Maximal distance between
UP and -35 Box	0	15
-35 Box and -10 Box	15	20
-10 Box and Shine Dalgarno	0	200
Shine Dalgarno and Start codon	2	10

Déroulement d'une session de travail

- liste cliquable des gènes obtenus

BacTrans²...genes identified with a strong promoter - Mozilla Firefox

Echier Edition Affichage Aller à Marque-pages Outils ?

http://www.sciences.univ-nantes.fr/lina/bioserv/BacTrans2/work/genesList.html?id=sinoquet&file=../../work/sinoquet_8_80_080222_thermotoga_maritima_UP_4_2_3_2_0_0_ OK

Démarrage Dernières nouvelles

oO~ BacTrans² ~Oo

List of genes identified with a strong promoter

- [TM1237](#)
- [TM1408](#)
- [TM0060](#)
- [TM1782](#)
- [TM0372](#)
- [TM1815](#)
- [TM1850](#)
- [TM0295](#)
- [TM1016](#)
- [TM0626](#)
- [TM1780](#)
- [TM0364](#)
- [TM1667](#)
- [TM1119](#)
- [TM0339](#)
- [TM1429](#)
- [TM0786](#)
- [TM1605](#)
- [TM1607](#)
- [TM1803](#)
- [TM0716](#)

Terminé

démarrer evaluation_cnrs_2007 Microsoft PowerPoin... [Ne pas fermer] Wifi... BacTrans²...genes i... Faire une capture éc... specification_requet... FR 16:09

Déroulement d'une session de travail

- obtention des caractéristiques des promoteurs forts identifiés

/distance UP,-35 box/ 1
/distance -35 box,-10 box/ 18
/distance -10 box, Shine-Dalgarno/ 32
/distance Shine-Dalgarno, start codon/ 8
/positions for UP w.r.t. +1 transcript./ -60 -39
/positions for -35 box w.r.t. +1 transcript./ -37 -33
/positions for -10 box w.r.t. +1 transcript./ -14 -9
/positions for Shine-Dalgarno box w.r.t. +1 transcript./ 25 29
/position for start codon w.r.t. +1 transcript./ 38

Déroulement d'une session de travail

- description des gènes à promoteurs forts (23 caractéristiques)

Déroulement d'une session de travail

- visualisation graphique d'une caractéristique, pour un génome donné

exemple de caractéristique distribuée en 4 classes

classes

no_unpaired_region

Shine-Dalgarno_in_unpaired_region

start_codon_in_unpaired_region

start_codon_and_Shine-Dalgarno_in_unpaired_region

classes



effectif

Déroulement d'une session de travail

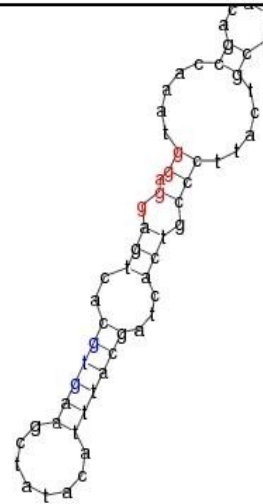
- visualisation des meilleurs repliements d'une molécule nécessaire à la synthèse protéique (ARN messenger)

Folding

- + ygcK
- + codA
- + yjhI
- + argI
- + ybjD
- + ansB
- + b3022
- + deoA
- + cydC
- + ycaK
- + grxB

Best Predicted Version

- + flgB
- + ycfJ
- + bioC
- + ymfO
- + ybgK
- + narI



Save

Rechercher : TM0477 Occurrence suivante Occurrence précédente Surligner tout Respecter la casse Bas de la page atteint, poursuite au début

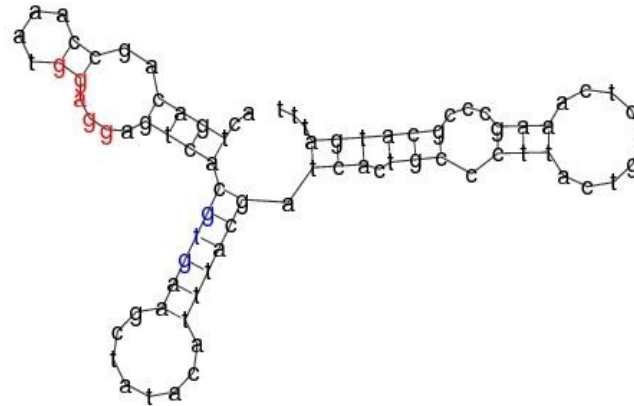
Terminé

Folding

- + ygcK
- + codA
- + yjhI
- + argI
- + ybjD
- + ansB
- + b3022
- + deoA
- + cydC
- + ycaK
- + grxB

Best Predicted Version

- + flgB
- + ycfJ
- + bioC
- + ymfO
- + ybgK
- + narI



Save

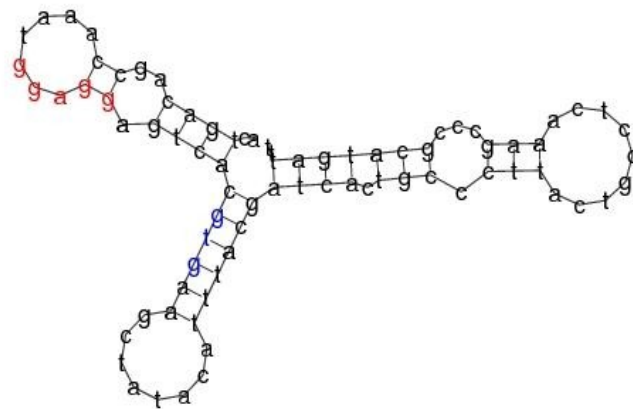
oO~ BacTrans² ~Oo

Rechercher : TM0477 Occurrence suivante Occurrence précédente Surligner tout Respecter la casse Bas de la page atteint, poursuite au début

Terminé

Folding

```
+ grxB  
  Best Predicted Version  
+ flgB  
+ ycfJ  
+ bioC  
+ ymfO  
+ ybgK  
+ narI  
+ oppF  
+ yghD  
+ yjbA  
+ thiE  
+ cpxR  
+ rhaS  
+ yigN  
+ wecF  
+ ydaC  
+ tra8 2
```



Save

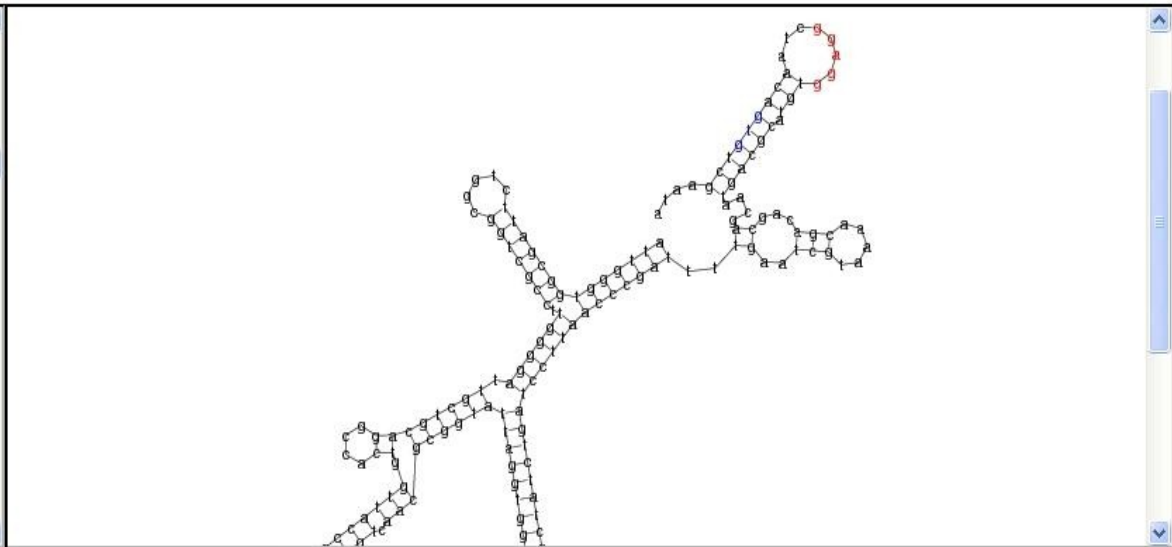
oO~ BacTrans² ~Oo

Rechercher : grxB Occurrence suivante Occurrence précédente Surligner tout Respecter la casse

Terminé

Folding

- + ldcC
 - + yfbM
 - + ydcF
 - + cchB
 - + b2506
 - + yphF
 - + csgF
 - + speE
 - + ygaG
 - + ynhA
 - + ygbA
 - + tra8_3
 - + b0836
 - + b2756
 - + ygcK
 - + codA
- Best Predicted Version

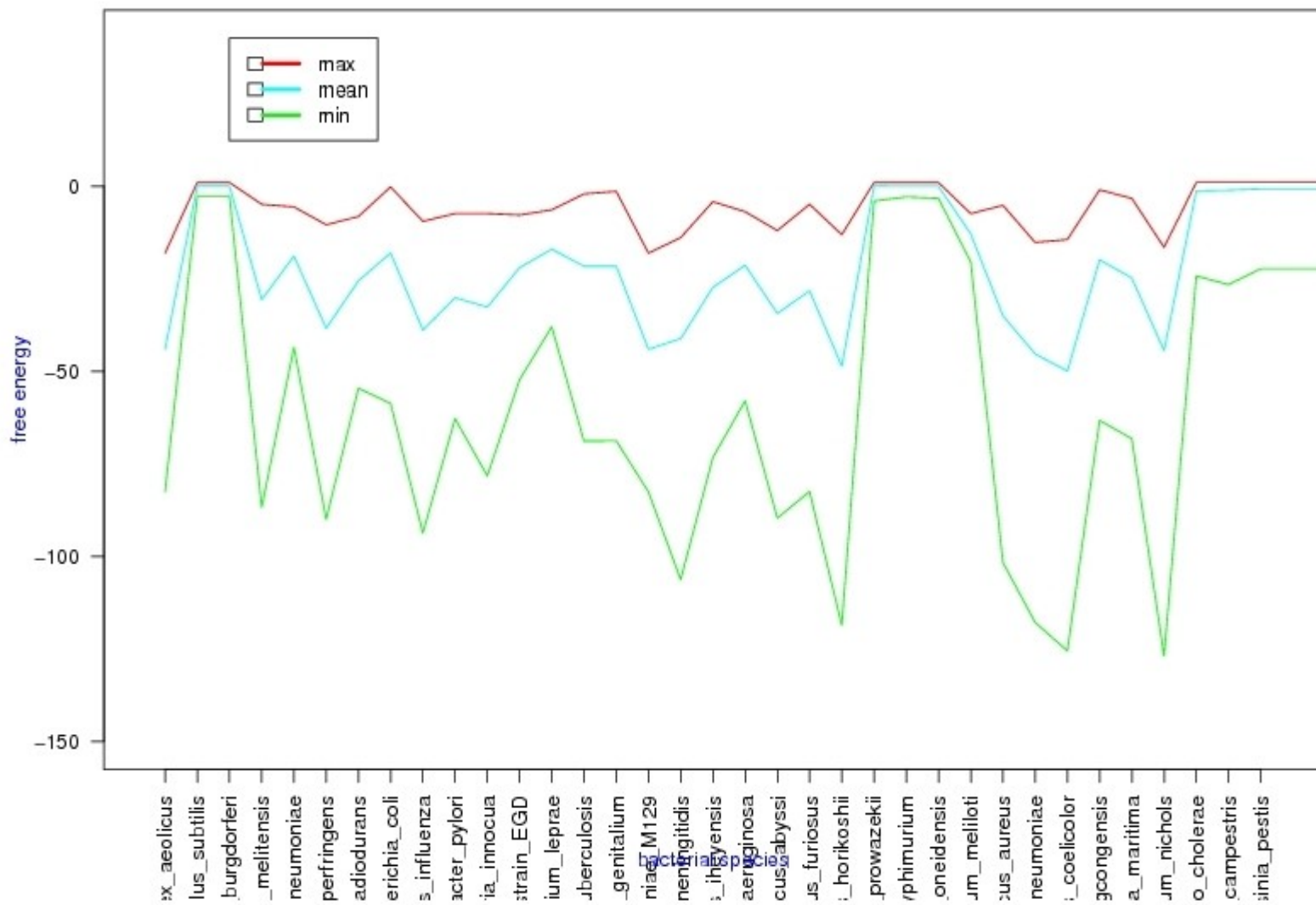


oO~ BacTrans² ~Oo

Déroulement d'une session de travail

- visualisation graphique d'une caractéristique, pour plusieurs génomes

free energy variation graph



Résultats

- changement d'échelle dans l'analyse
- plate-forme générique
(contraintes 3 à 4 boîtes, dédiée tous génomes procaryotes)
- originalité
prise en compte de l'élément UP, pour le promoteur sigma 70
- projet abouti
- présentation JOBIM 2006
- utilisation intensive dans le cadre de l'étude comparative de 32 génomes (collaboration V.Sakanyan, F.Braun - fréquence des promoteurs forts, comparaison avec génomes aléatoires; mise en évidence de constantes ou de différences pour diverses caractéristiques)