

INRAE

➤ WP3 : bilan & focus sur Genoscapist

Cyprien & Sandra

Bilan du *workpackage* 3

Veille sur les outils de visualisation de graphes

→ document répertoriant les outils de visualisation de graphes :

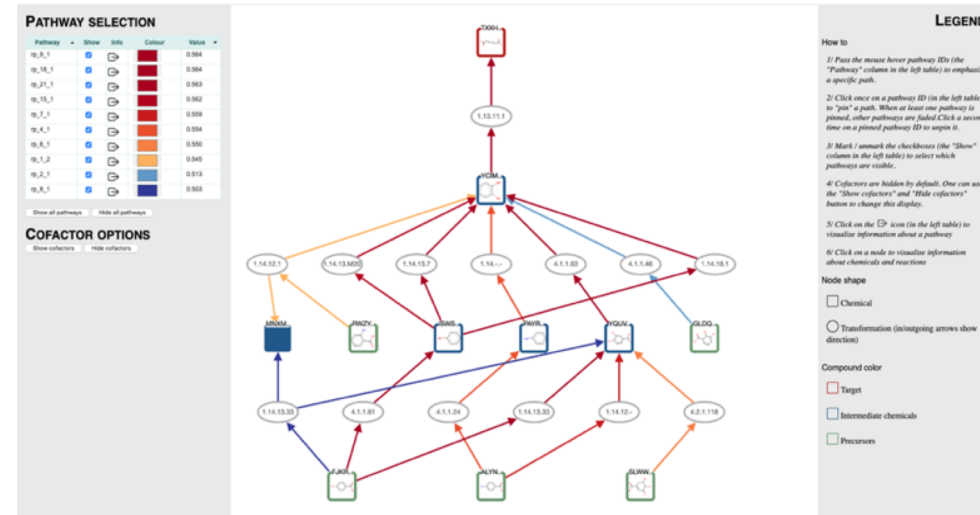
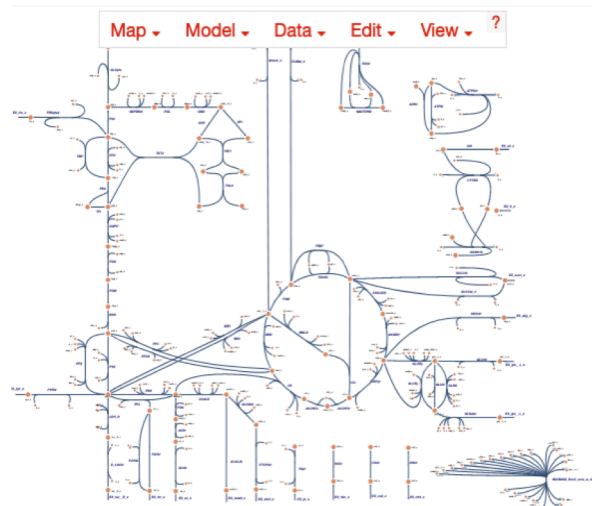
https://sysmics.cati.inrae.fr/sites/default/files/wp3/201906_outil_visualisation_processus.pdf

Investigation plus approfondie des outils :

→ [rpViz](#) : visualisation de voies métaboliques

→ [Escher](#) : visualisation de flux dans un réseau métabolique

Escher



rpViz

INRAE

Assemblée Générale du CATI SysMics

11 mai 2021

<https://sysmics.cati.inrae.fr/workpackage3>

Genoscapist

Visualisation de grandes quantités de données hétérogènes le long du génome

Genoscapist About

Genoscapist

Genoscapist

Genoscapist is a web-tool generating high-quality images for interactive visualization of hundreds of quantitative profiles along a reference genome together with various annotations.

Genomic View

1231001 1232000 1234000 1236000 1238000 1240000 1241000

log2(rp)

log2(rp)

log2(rp)

log2(rp)

Applications

- *B. subtilis*: <http://genoscapist.migale.inrae.fr/seb>
 - using transcriptome-based reannotation studie of *Bacillus subtilis* (Nicolas et al., 2012)
- *S. aureus*: <http://genoscapist.migale.inrae.fr/aeb>
 - using transcriptome-based reannotation studie of *Staphylococcus aureus* (Mäder et al., 2016)

MaiAGE

Copyright © 2020 INRAE - MaiAGE
genoscapist@inrae.fr

INRAE

Contexte

Besoins en visualisation de données des collaborateurs microbiologistes

- données transcriptomiques non agrégées sur de grands ensembles de données (permettant la découverte de mécanismes biologiques non canoniques)
- navigation basée sur la corrélation d'expression entre les gènes
- vues personnalisables : sélection des échantillons et coloration
- accès facile en ligne

Versions web statiques basées sur [MuGeN](#) utilisées par la communauté de *B. subtilis*¹ et de *S. aureus*²

Les outils existants ([JBrowse](#), [IGV](#), [Panoptes](#), ...) ne permettent pas d'avoir toutes ces fonctionnalités dans un seul navigateur

¹ Nicolas, P. et al. (2012). DOI: [10.1126/sciences.1206848](https://doi.org/10.1126/sciences.1206848)

² Mäder, U. et al. (2016). DOI: [10.1371/journal.pgen.1005962](https://doi.org/10.1371/journal.pgen.1005962)



Architecture

Technologies utilisées

→ côté serveur :



→ côté client :



→ stockage des données :

Code source

→ dépôt :



→ licence GNU GPL



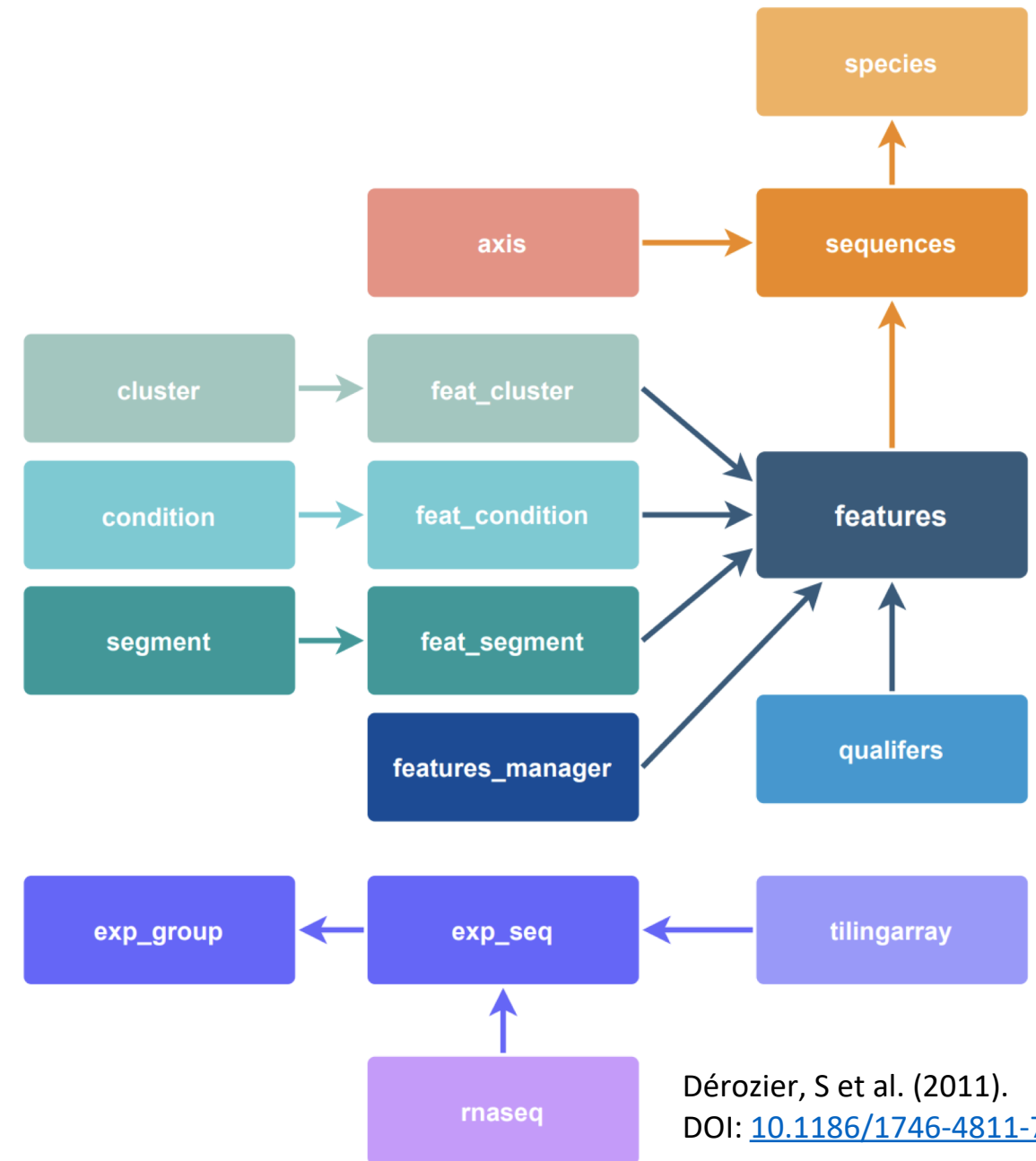
Architecture

Données *Tiling-array*

- 20 conditions
- 269 échantillons
- 191 000 mesures ponctuelles d'expression par échantillon

Données RNA-Seq


- 3 conditions
- 36 échantillons
- 1 300 000 mesures ponctuelles d'expression par échantillon



Dérozier, S et al. (2011).
DOI: [10.1186/1746-4811-7-8](https://doi.org/10.1186/1746-4811-7-8)






Fonctionnalités

B. subtilis Expression Data Browser About ▾ 

Identifier Position (kb) Submit Cluster

Show entries

Name ↑↓	Locus tag ↑↓	Genomic location ↑↓	Expression clusters ? ↑↓
S1	new_1_148	[1, 148] (strand 1)	A1 B1 C1
S2	new_24_297_c	[24, 297] (strand -1)	A2 B2 C2
S3	new_150_409	[150, 409] (strand 1)	A3 B3 C3
S4	new_299_1067_c	[299, 1067] (strand -1)	A4 B4 C2
dnaA	BSU00010	[410, 1750] (strand 1)	A5 B5 C3
S5	new_1751_1938	[1751, 1938] (strand 1)	A5 B5 C3
dnaN	BSU00020	[1939, 3075] (strand 1)	A5 B5 C3
yaaA	BSU00030	[3206, 3421] (strand 1)	A6 B5 C3
recF	BSU00040	[3437, 4549] (strand 1)	A7 B6 C3
yaaB	BSU00050	[4567, 4812] (strand 1)	A7 B6 C3

 Copyright © 2020 INRAE - MaIAGE
Last update: 16 December 2020  

Fonctionnalités

 *B. subtilis* Expression Data Browser View options ▾ Export ▾ About ▾



Details about expression for BSU11590

Expression clusters ⓘ : [A39](#) [B5](#) [C3](#)

Highest Expression Conditions

dia5	LBGtran	dia0	dia15	Lbexp	LBGexp	G180	Mt0	G150	Etha
13.7	13.6	13.6	13.6	13.6	13.5	13.3	13.3	13.2	13.2

Lowest Expression Conditions

BT	Sw	B36	B60	S8	S6	BC	S7	Lbstat	LPhT
9.16	9.22	9.47	9.81	9.91	10.1	10.3	10.6	11.3	11.3

Most Positively Correlated Segments ⓘ

ywiB	S24	S756	yvzF	ydiL	S1381	S886	gatC	S859	pfkA
0.91	0.9	0.89	0.88	0.88	0.87	0.87	0.87	0.87	0.87

Most Negatively Correlated Segments ⓘ

S978	ispA	fadN	S471	S1205	fadE	fadA	ybbC	S472	yycO
-0.77	-0.77	-0.76	-0.76	-0.75	-0.75	-0.74	-0.74	-0.74	-0.73



Fonctionnalités



INRAE

Assemblée Générale du CATI SysMics

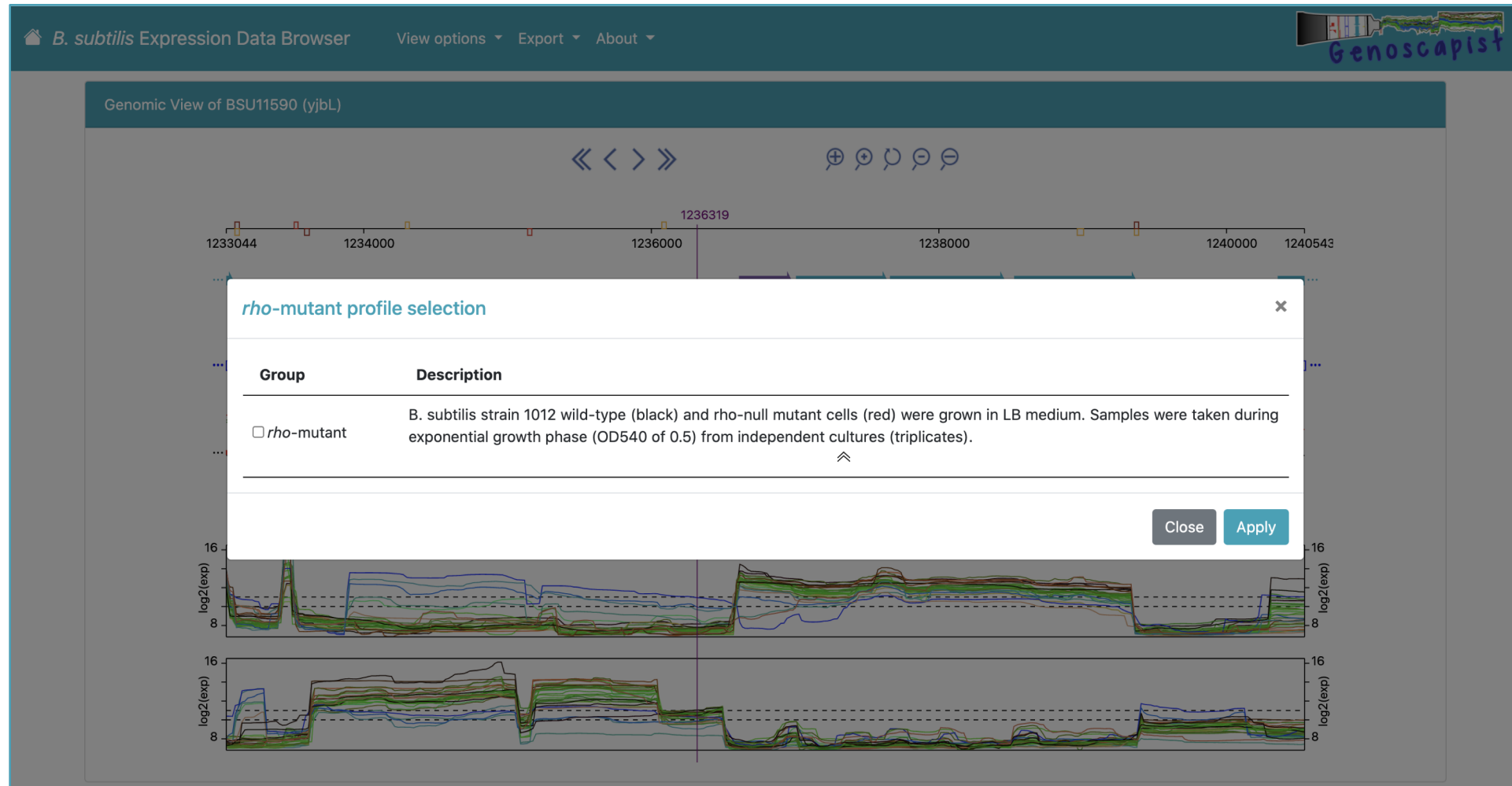
11 mai 2021

<https://genoscapist.migale.inrae.fr/>

Fonctionnalités



Fonctionnalités



Fonctionnalités



Fonctionnalités



Fonctionnalités

Condition-dependent profile selection

List Grid

<input checked="" type="checkbox"/> aero_1	<input type="checkbox"/> aero_2	<input type="checkbox"/> aero_3	<input checked="" type="checkbox"/> B36_1	<input type="checkbox"/> B36_2	<input type="checkbox"/> B60_1
<input type="checkbox"/> B60_2	<input type="checkbox"/> BC_1	<input type="checkbox"/> BC_2	<input type="checkbox"/> BC/t_1	<input type="checkbox"/> BC/t_2	<input type="checkbox"/> BI_1
<input type="checkbox"/> BL_2	<input type="checkbox"/> BMM_1	<input type="checkbox"/> BMM_2	<input checked="" type="checkbox"/> BMM_3	<input type="checkbox"/> BT_1	<input type="checkbox"/> BT_2
<input type="checkbox"/> C30_1	<input type="checkbox"/> C30_2	<input type="checkbox"/> C90_1	<input checked="" type="checkbox"/> C90_2	<input checked="" type="checkbox"/> Cold_1	<input type="checkbox"/> Cold_2
<input type="checkbox"/> Cold_3	<input type="checkbox"/> dia0_1	<input type="checkbox"/> dia0_2	<input type="checkbox"/> dia0_3	<input type="checkbox"/> dia15_1	<input checked="" type="checkbox"/> dia15_2
<input type="checkbox"/> dia15_3	<input type="checkbox"/> dia5_1	<input type="checkbox"/> dia5_2	<input type="checkbox"/> dia5_3	<input type="checkbox"/> Diami_1	<input type="checkbox"/> Diami_2
<input type="checkbox"/> Diami_3	<input type="checkbox"/> Etha_1	<input type="checkbox"/> Etha_2	<input type="checkbox"/> Etha_3	<input type="checkbox"/> ferm_1	<input checked="" type="checkbox"/> ferm_2
<input type="checkbox"/> ferm_3	<input type="checkbox"/> Fru_1	<input type="checkbox"/> Fru_2	<input type="checkbox"/> Fru_3	<input type="checkbox"/> G135_1	<input type="checkbox"/> G135_3
<input checked="" type="checkbox"/> G150_1	<input type="checkbox"/> G150_2	<input type="checkbox"/> G150_3	<input type="checkbox"/> G180_1	<input type="checkbox"/> G180_2	<input type="checkbox"/> G180_3
<input type="checkbox"/> Glu_1	<input type="checkbox"/> Glu_2	<input type="checkbox"/> Glu_3	<input type="checkbox"/> Glucon_1	<input type="checkbox"/> Glucon_2	<input type="checkbox"/> Glucon_3
<input type="checkbox"/> Gly_1	<input type="checkbox"/> Gly_2	<input type="checkbox"/> Gly_3	<input type="checkbox"/> GM-0.1_1	<input type="checkbox"/> GM-0.1_2	<input type="checkbox"/> GM-0.1_3
<input checked="" type="checkbox"/> GM-0.2_1	<input type="checkbox"/> GM-0.2_2	<input type="checkbox"/> GM-0.2_3	<input type="checkbox"/> GM+10_2	<input type="checkbox"/> GM+10_3	<input type="checkbox"/> GM+120_1
<input type="checkbox"/> GM+120_3	<input type="checkbox"/> GM+150_1	<input type="checkbox"/> GM+150_2	<input type="checkbox"/> GM+150_3	<input type="checkbox"/> GM+15_1	<input type="checkbox"/> GM+15_2
<input type="checkbox"/> GM+15_3	<input type="checkbox"/> GM+25_1	<input type="checkbox"/> GM+25_2	<input type="checkbox"/> GM+25_3	<input type="checkbox"/> GM+45_1	<input type="checkbox"/> GM+45_2
<input type="checkbox"/> GM+45_3	<input type="checkbox"/> GM+5_1	<input type="checkbox"/> GM+5_2	<input checked="" type="checkbox"/> GM+5_3	<input type="checkbox"/> GM+60_1	<input type="checkbox"/> GM+60_2
<input type="checkbox"/> GM+60_3	<input checked="" type="checkbox"/> GM+90_1	<input type="checkbox"/> GM+90_2	<input type="checkbox"/> GM+90_3	<input type="checkbox"/> G+S_1	<input type="checkbox"/> G+S_2
<input type="checkbox"/> G+S_3	<input type="checkbox"/> H2O2_1	<input type="checkbox"/> H2O2_2	<input type="checkbox"/> H2O2_3	<input type="checkbox"/> Heat_1	<input type="checkbox"/> Heat_2
<input type="checkbox"/> Heat_3	<input type="checkbox"/> HiOs_1	<input type="checkbox"/> HiOs_2	<input type="checkbox"/> HiOs_3	<input type="checkbox"/> HiTm_1	<input type="checkbox"/> HiTm_2
<input checked="" type="checkbox"/> HiTm_3	<input type="checkbox"/> HPh_1	<input type="checkbox"/> HPh_2	<input type="checkbox"/> HPh_3	<input type="checkbox"/> LBexp_1	<input type="checkbox"/> LBexp_2
<input type="checkbox"/> LBexp_3	<input type="checkbox"/> LBExp_1	<input type="checkbox"/> LBExp_2	<input type="checkbox"/> LBExp_3	<input type="checkbox"/> LBExp_1	<input type="checkbox"/> LBExp_2

Discard non-applied selection Default None All Apply

Fonctionnalités

Condition-dependent profile selection

List Grid

Show 10 entries

Search:

Group	Description
<input type="checkbox"/> Anaerobic growth	Cells were grown in a synthetic medium (E. Härtig, A. Hartmann, M. Schätzle, A. M. Albertini, D. Jahn, Appl Environ Microbiol 72, 5260, 2006) at 37 °C. For aerobic growth, an overnight culture was used to
<input checked="" type="checkbox"/> Biofilm formation	A fresh colony grown on an LB plate was used to inoculate 10 ml of LB and grown for 10 hours at 30°C. This culture was used to inoculate 10 ml of MSgg medium (S.S. Branda et al., J Bacteriol 186, 3970, Jun, 2004) and incubated with vigorous shaking. The cultures in MSgg were diluted to the same extent in 96 wells microtiterplates (5 µl for 1.5 ml of medium) and incubated without shaking at 30°C. Cells from the control cultures were harvested after 24 hours of incubation [BT]. Biofilms were harvested from 96 well plates after incubation for 36 hours [B36] and 60 hours [B60].
<input type="checkbox"/> Carbon sources	A 5 ml aliquot of LB medium was inoculated using frozen culture stocks. After a few hours growth at 37°C, precultures were prepared by inoculating 5 ml of M9 with this LB culture at several different

B36 B36_1 B36_2 B60 B60_1 B60_2 BT BT_1 BT_2

Discard non-applied selection Default None All Apply



Fonctionnalités

Condition-dependent profile selection

List Grid

Show 10 entries

Search:

Group	Description
<input type="checkbox"/> Anaerobic growth	Cells were grown in a synthetic medium (E. Härtig, A. Hartmann, M. Schätzle, A. M. Albertini, D. Jahn, Appl Environ Microbiol 72, 5260, 2006) at 37 °C. For aerobic growth, an overnight culture was used to
<input checked="" type="checkbox"/> Biofilm formation	A fresh colony grown on an LB plate was used to inoculate 10 ml of LB and grown for 10 hours at 30°C. This culture was used to inoculate 10 ml of MSgg medium (S.S. Branda et al., J Bacteriol 186, 3970, Jun, 2004) and incubated with vigorous shaking. The cultures in MSgg were diluted to the same extent in 96 wells microtiterplates (5 µl for 1.5 ml of medium) and incubated without shaking at 30°C. Cells from the control cultures were harvested after 24 hours of incubation [BT]. Biofilms were harvested from 96 well plates after incubation for 36 hours [B36] and 60 hours [B60].
<input type="checkbox"/> Carbon sources	A 5 ml aliquot of LB medium was inoculated using frozen culture stocks. After a few hours growth at 37°C, precultures were prepared by inoculating 5 ml of M9 with this LB culture at several different

B36 B60 BT

B36_1 B60_1 BT_1

B36_2 B60_2 BT_2

Discard non-applied selection Default None All Apply

Fonctionnalités



Fonctionnalités

Condition-dependent profile selection

List Grid

Show 10 entries

Search:

Group	Description
<input type="checkbox"/> Anaerobic growth	Cells were grown in a synthetic medium (E. Härtig, A. Hartmann, M. Schätzle, A. M. Albertini, D. Jahn, Appl Environ Microbiol 72, 5260, 2006) at 37 °C. For aerobic growth, an overnight culture was
<input checked="" type="checkbox"/> Biofilm formation	A fresh colony grown on an LB plate was used to inoculate 10 ml of LB and grown for 10 hours at 30°C. This culture was used to inoculate 10 ml of MSgg medium (S.S. Branda et al., J Bacteriol 186,
<input checked="" type="checkbox"/> B36 <input type="checkbox"/> B36_1 <input type="checkbox"/> B36_2	<input checked="" type="checkbox"/> B60 <input type="checkbox"/> B60_1 <input type="checkbox"/> B60_2
<input checked="" type="checkbox"/> Carbon sources	A 5 ml aliquot of L 37°C, precultures
<input checked="" type="checkbox"/> Competence	Cells grown overn minimal medium a

Discard non-applied selection Default None All Apply

Fonctionnalités



INRAE

Assemblée Générale du CATI SysMics

11 mai 2021

<https://genoscapist.migale.inrae.fr/>

Fonctionnalités



Evolutions & Perspectives

Visualisation d'autres types de données

- RNA-Seq et 5'-RNA-Seq
- utilisation de glyphes personnalisés en SVG

Personnalisation et gestion des « pistes de données » côté utilisateur

Chargement et visualisation de données à la volée



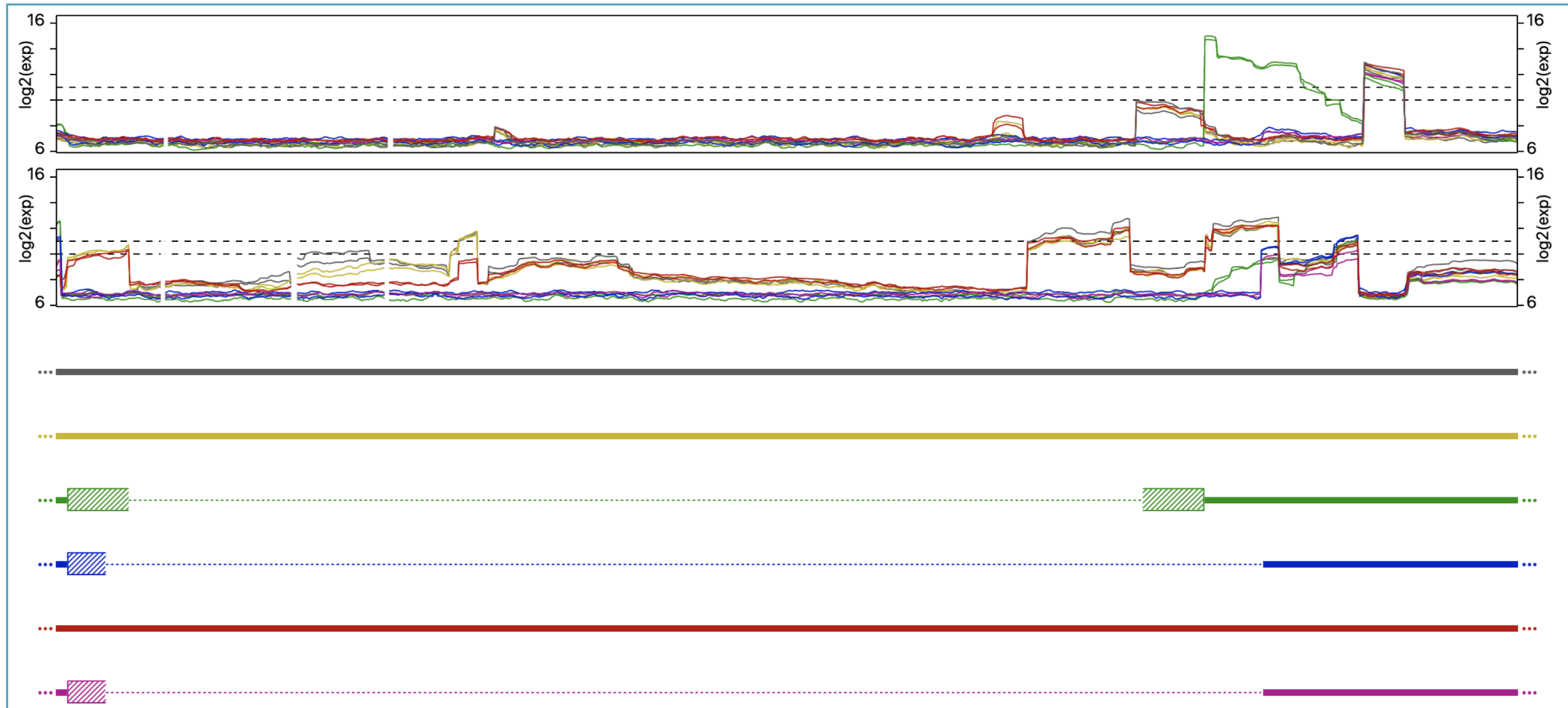
Portail agrégateur de visualisation inter-projet par génome de référence

Poursuite des déploiements associés à des projets de recherche



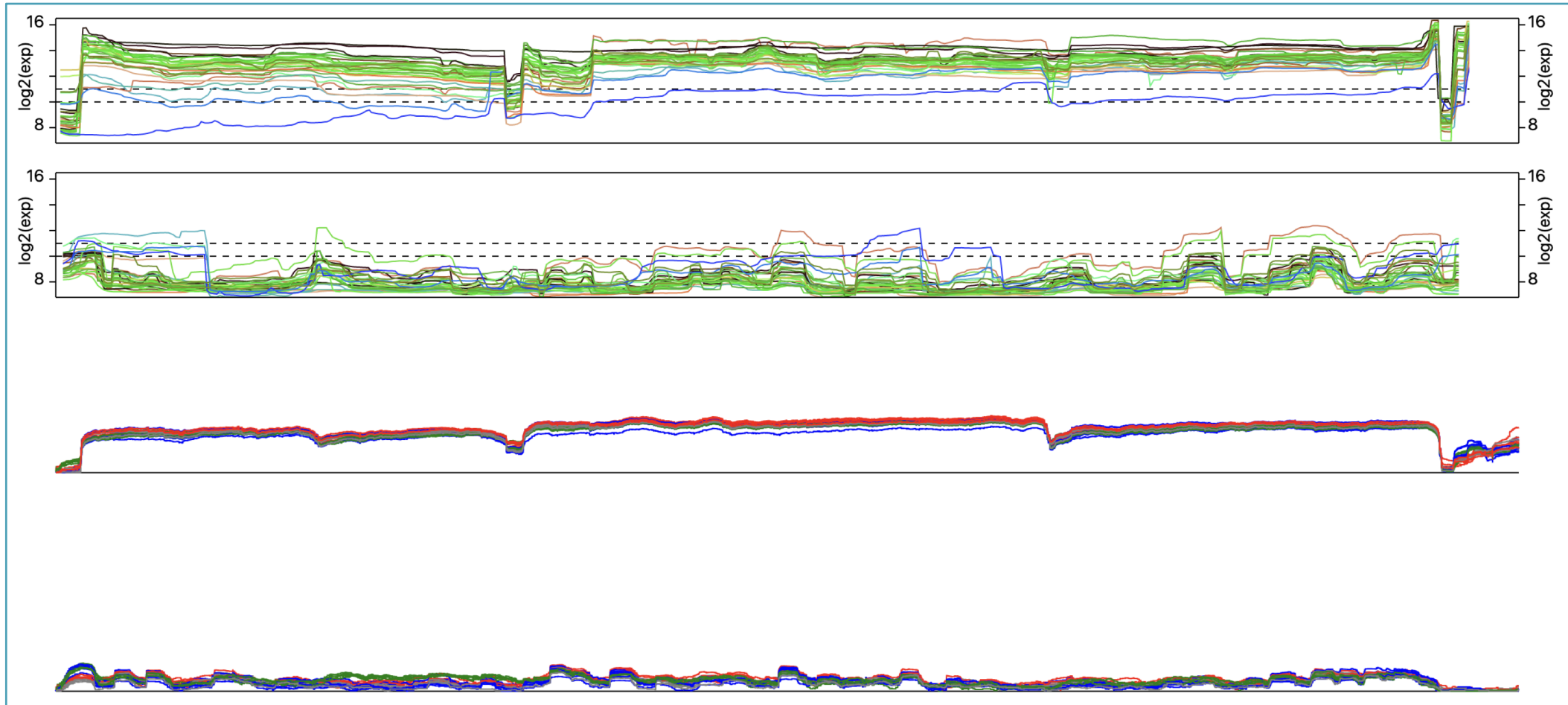
Minimal genome *Bacillus subtilis* (en cours)

Représentation de modifications génomiques (délétions, cassettes de sélection et cicatrices de construction).



RNA-Seq *Bacillus subtilis* (en cours)

Visualisation multi-échelle (vue large ou à la base près) en utilisant des solutions de segmentations de signal.



Evolutions & Perspectives

Visualisation d'autres types de données

- RNA-Seq et 5'-RNA-Seq
- utilisation de glyphes personnalisés en SVG

Personnalisation et gestion des « pistes de données » côté utilisateur

Chargement et visualisation de données à la volée



Portail agrégateur de visualisation inter-projet par génome de référence

Poursuite des déploiements associés à des projets de recherche

Projets

Projet de ré-annotation de *Staphylococcus aureus*

→ *tiling-array* (Dérozier et al. 2021)

Projet de ré-annotation de *Bacillus subtilis*

→ *tiling-array* (Dérozier et al. 2021)

Génome minimal de *Bacillus subtilis*

→ *tiling-array* et structure génomique (associé à une publication, 2021)

Modulation d'expression de Rho chez *Bacillus subtilis*

→ **RNA-seq** (associé à une publication, 2021)

ANR CoNoCo (*Bacillus subtilis*)

→ **RNA-seq** (associé à une publication)

Workflow d'annotation de pseudogènes

→ annotation (associé à une publication, 2021)



Perspectives du *workpackage* 3

Interfaces de visualisation le long du génome

- visualisation de profils RNA-Seq → étude des choix possibles de visualisation pour ce type de données qui présentent une granularité plus importante de données brutes par rapport aux données de *tiling array* déjà traitées.

Interfaces de visualisation de réseaux

- suite aux échanges lors du *workshop* Work4Graph, évaluation d'autres solutions de visualisation, notamment [MetExploreViz](#).

Interfaçage des outils de visualisation avec les données extraites du système informatique structuré proposé par le WP1

- l'objectif final du WP3 est d'interconnecter les outils de visualisation le long du génome ainsi que les outils de visualisation de réseaux avec l'entrepôt de données issu du WP1.

