



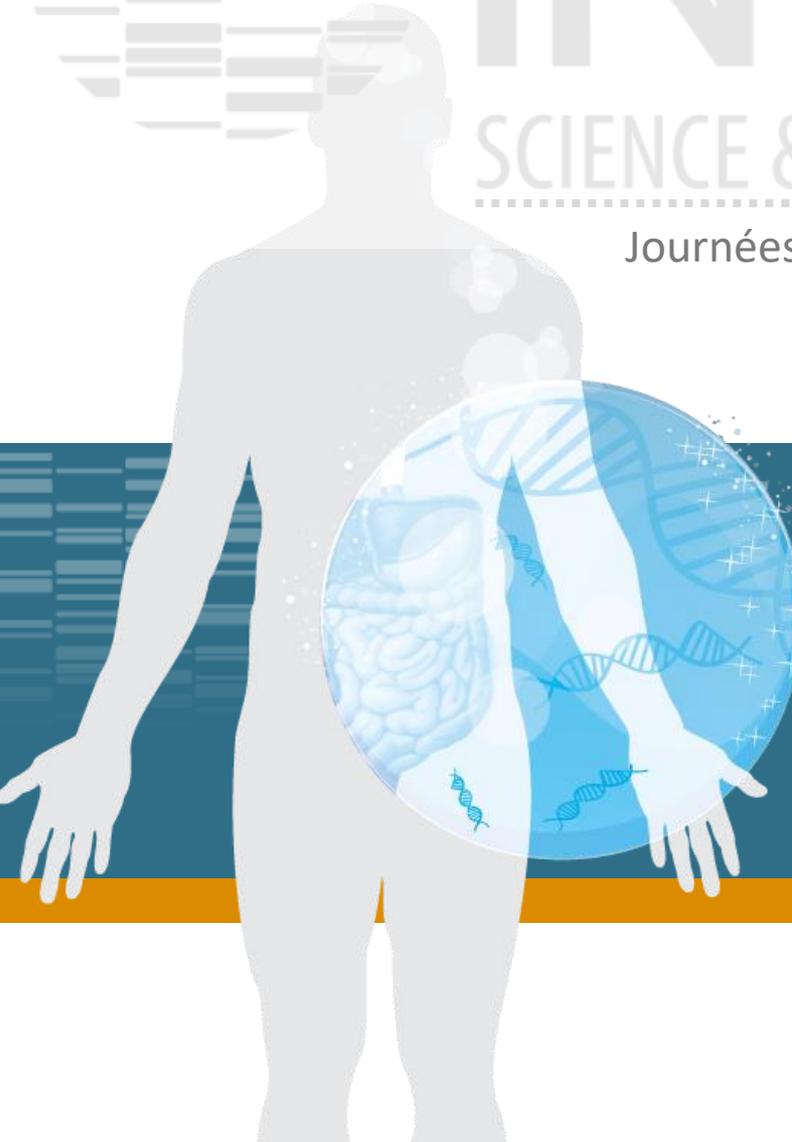
INRA

SCIENCE & IMPACT



metagenopolis
mgps.eu

Journées PEPI IBIS

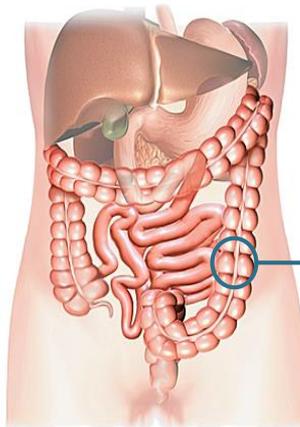


Les outils d'inférence de réseaux
dans le contexte de l'analyse du
microbiote intestinal humain

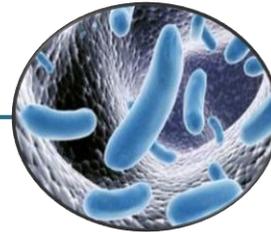
Louisa Hadj Abed – Unité MetaGenoPolis de l'INRA de Jouy-en-Josas (78)



Le rôle du microbiote intestinal dans la santé et la maladie



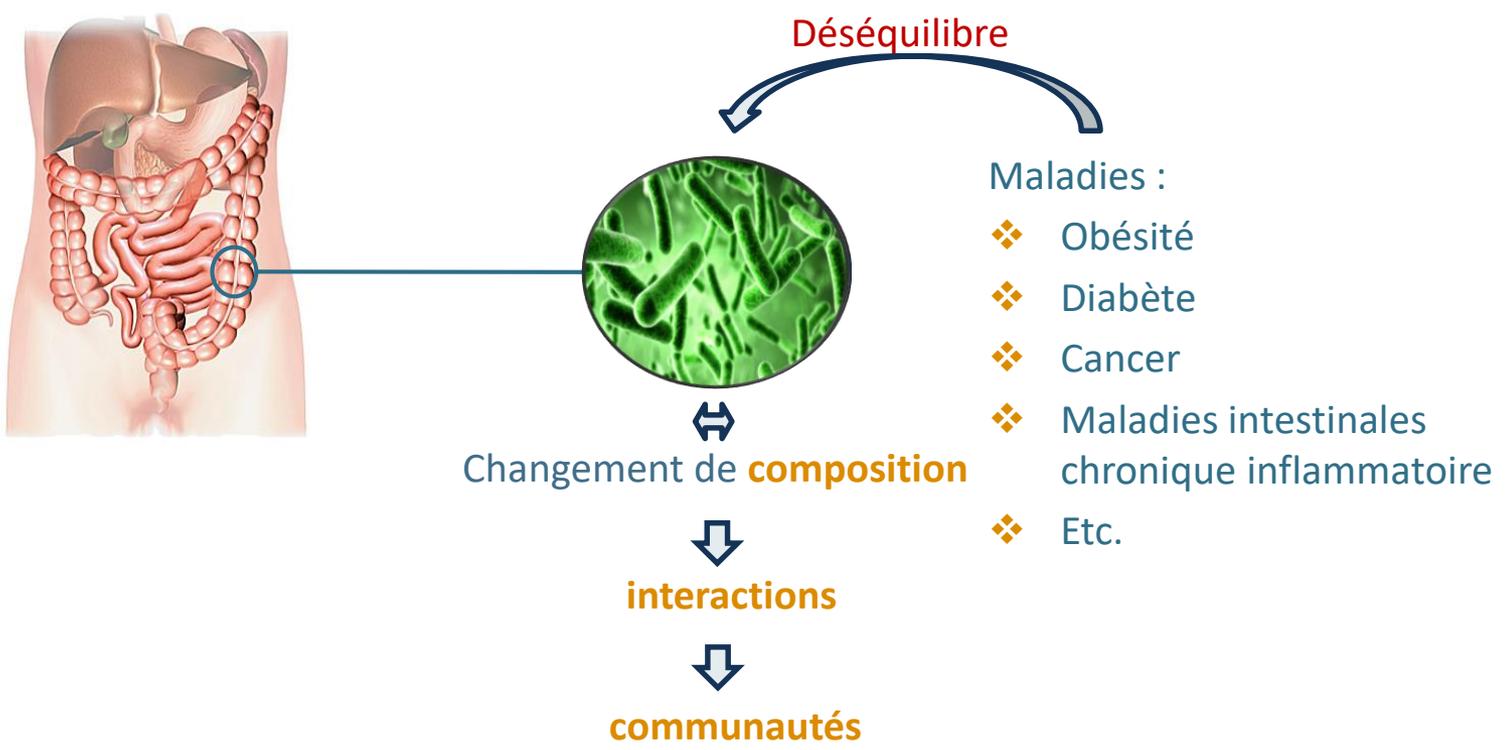
Microbiote



Fonctions :

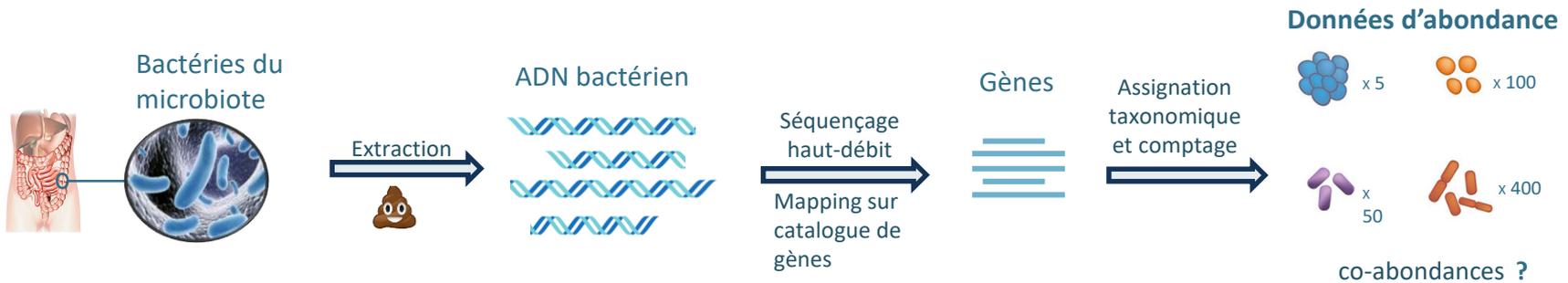
- ❖ Digestives
 - ❖ Métaboliques
 - ❖ Immunitaires
 - ❖ Neurologiques
- ↔ Santé

Le rôle du microbiote intestinal dans la santé et la maladie

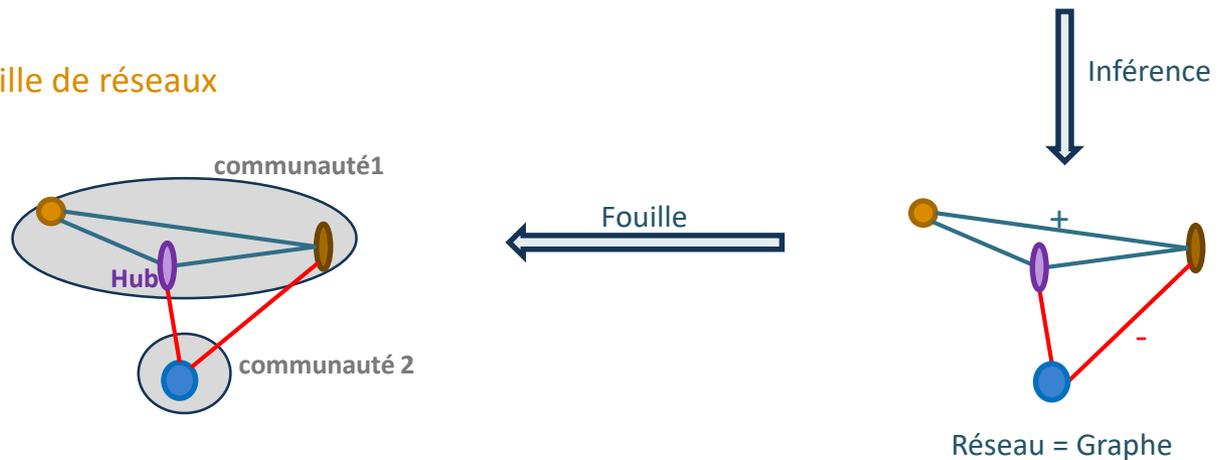


La métagénomique quantitative et les réseaux bactériens

Etape 1 : La métagénomique quantitative



Etape 2 : l'inférence et la fouille de réseaux





Quels sont les **outils** d'inférence de **réseaux** qui existent aujourd'hui et quels sont leurs **caractéristiques** ?



Panorama des outils d'inférence de réseaux



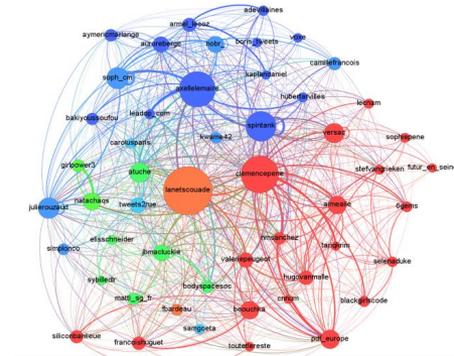
➔ Accélération du nombre d'outils développés depuis 4 ans

➔ Sujet d'étude en expansion

Les étapes de l'inférence de réseaux

Table de comptage

	Échant. 1	Échant. 2	Échant. j
Bactérie 1	500	25 000	...
Bactérie 2	20	0	...
Bactérie i



Filtre sur l'occurrence et l'abondance

Transformation des données

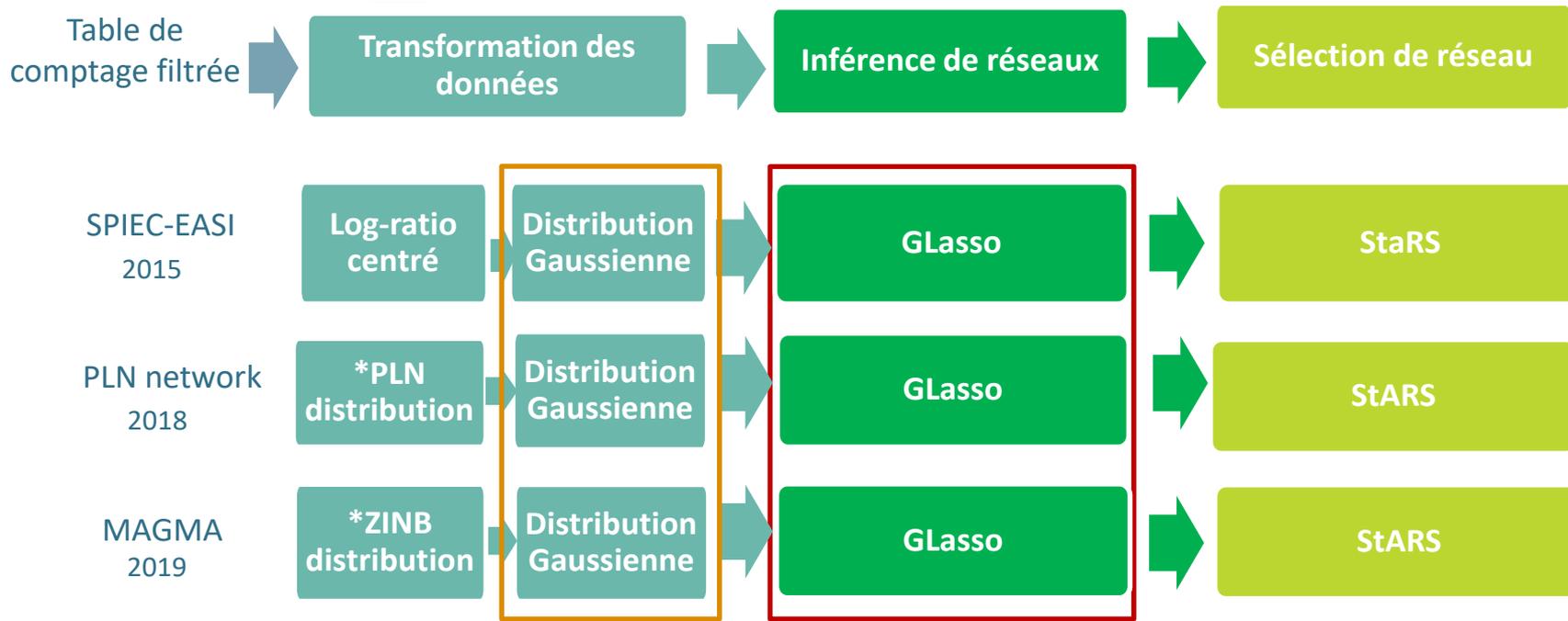
Inférence de réseaux

Sélection

Fouille de réseau & visualisation



Exemples d'étapes effectuées par trois outils d'inférence



*PLN : Poisson log-Normal distribution

*ZINB : Zero-Inflated Negative Binomial distribution

➔ Pourquoi vouloir obtenir une distribution Gaussienne ?

GLasso: *Graphical Lasso*

Inférence de graphes

GLasso = Modèle Graphique Gaussien + pénalité Lasso

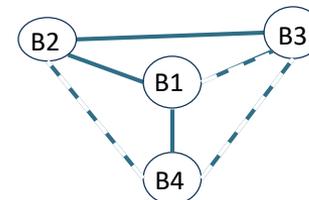
1. Modèle Graphique Gaussien : Calcul de covariances conditionnelles entre les variables

- Que des valeurs de covariance différentes de zéro = liens
- Lesquelles sont des vrais zéro ?

Matrice de covariance

	B1	B2	B3	B4
B1				
B2				
B3	0.9			
B4	0.8	0.01		

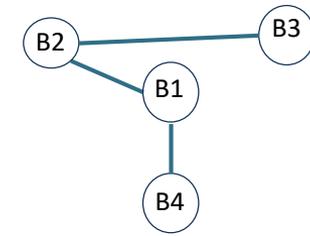
Graphe de liens directs et indirects



Matrice de covariance

	B1	B2	B3	B4
B1				
B2				
B3				
B4	0			

Graphe de liens directs



2. Pénalité Lasso : Ajout d'une pénalité de sparsité λ

- Valeurs égales à zéro \Leftrightarrow pas de lien
- Différentes de zéro \Leftrightarrow liens

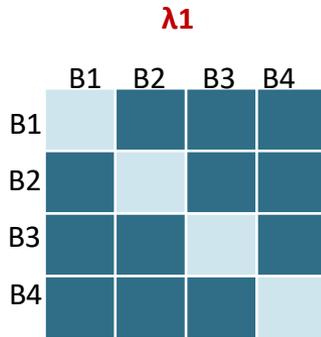
Glasso: Graphical Lasso

Covariance $\neq 0$ \Leftrightarrow interaction
 Covariance = 0 \Leftrightarrow pas d'interaction

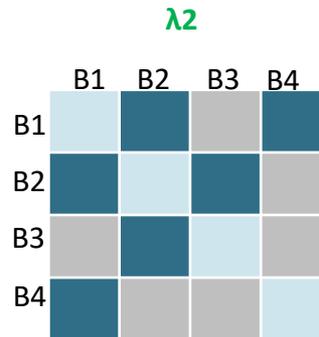
$\uparrow \lambda$: graphe sparse

$\downarrow \lambda$: graphe dense

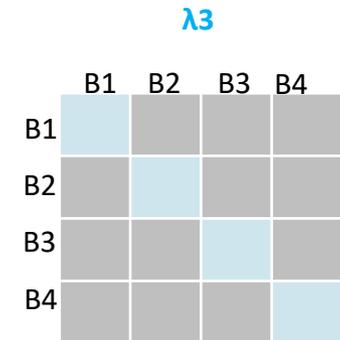
Matrices des Covariances :



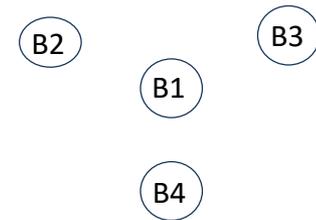
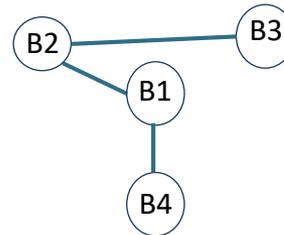
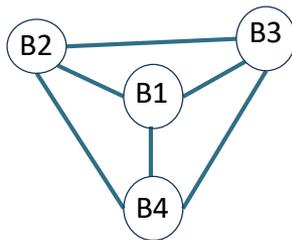
<



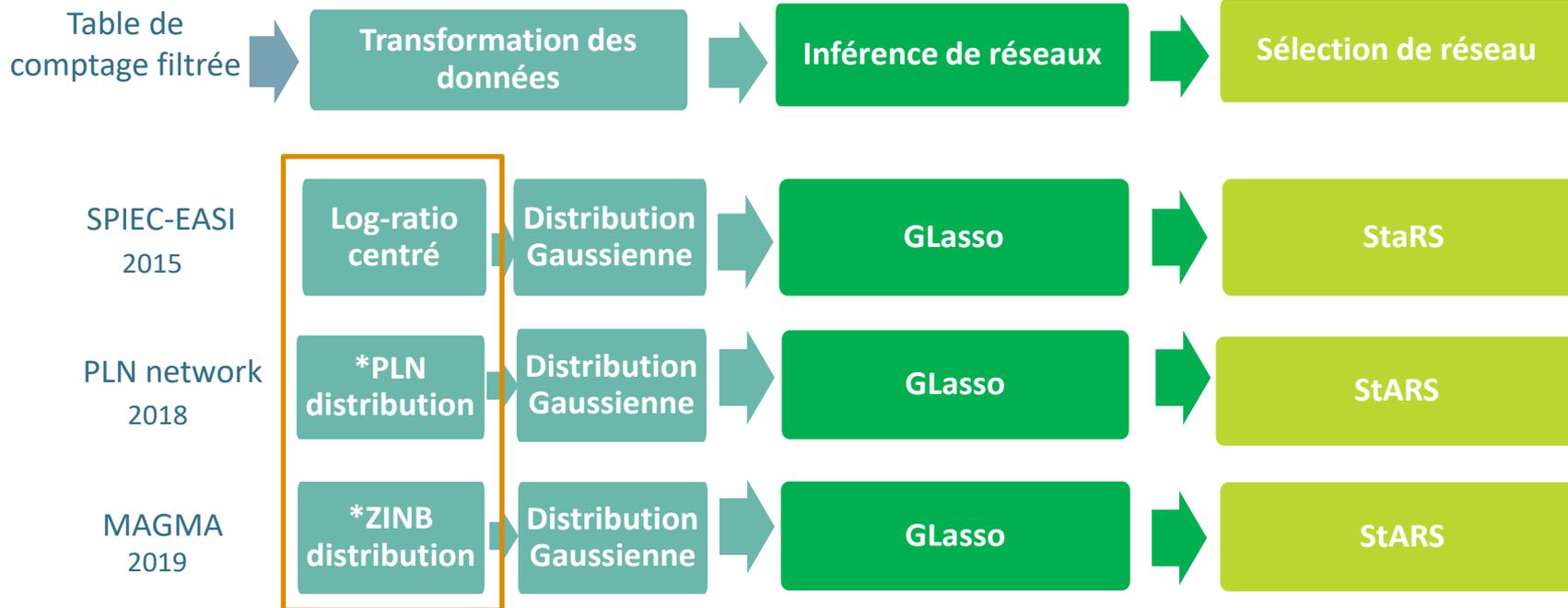
<



Graphes :



Exemples d'étapes effectuées par trois outils d'inférence



*PLN: Poisson log-Normal distribution

*ZINB: Zero-Inflated Negative Binomial distribution

➔ Quelles sont les caractéristiques des données de métagénomiques ?

Caractéristiques des données métagénomiques



- ❖ **Grande dimensionnalité** : nombre de bactéries \gg nombre d'échantillons
 - ❖ **Sur-dispersées** : Amplitude de valeurs d'abondance très large
 - ❖ **Sparse** : Beaucoup de zéros dans le tableau d'abondance
 - ❖ **Compositionnelles** : Données d'abondance relatives \rightarrow Somme contrainte à 1
- \rightarrow Pas de lois de distribution de probabilité propre à ce type de données
- \rightarrow Celles qui s'en rapprochent le plus :
- ❖ distribution Poisson Log-Normale (PLN)
 - ❖ distribution Zero-inflated Binomiale Negative (ZINB)

Les transformations

SPIEC-EASI
2015

Log-ratio
centré

→ Prend le **log du ratio de chaque abondance** et le divise par la moyenne géométrique
1986, Aitchison et al.

→ Palie au problème de **compositionnalité**

PLN network
2018

*PLN
distribution

Distribution
Gaussienne

→ **Approximation** des paramètres de la distribution PLN → Normale

→ Prend en compte la **compositionnalité**, la **sur-dispersion** et la **sparsité**
→ Possibilité de prendre en compte des **covariables** (sain/malade)

MAGMA
2019

*ZINB
distribution

Distribution
Gaussienne

→ **Approximation** des paramètres de la distribution PLN → Normale

→ Prend en compte la **compositionnalité**, la **sur-dispersion** et la **sparsité**
→ Possibilité de prendre en compte des **covariables** (sain/malade)

Exemples d'étapes effectuées par trois outils d'inférence



SPIEC-EASI
2015



PLN network
2018



MAGMA
2019



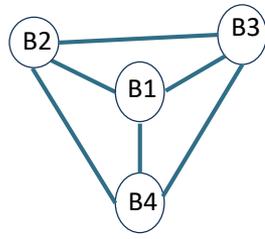
*PLN : Poisson log-Normal distribution

*ZINB : Zero-Inflated Negative Binomial distribution

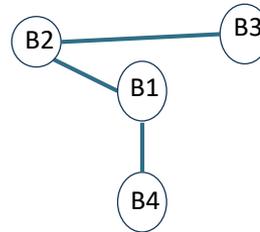
Sélection d'un graphe

❖ Après le Glasso :

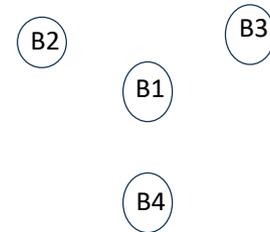
Différents graphes inférés :



λ_1



λ_2

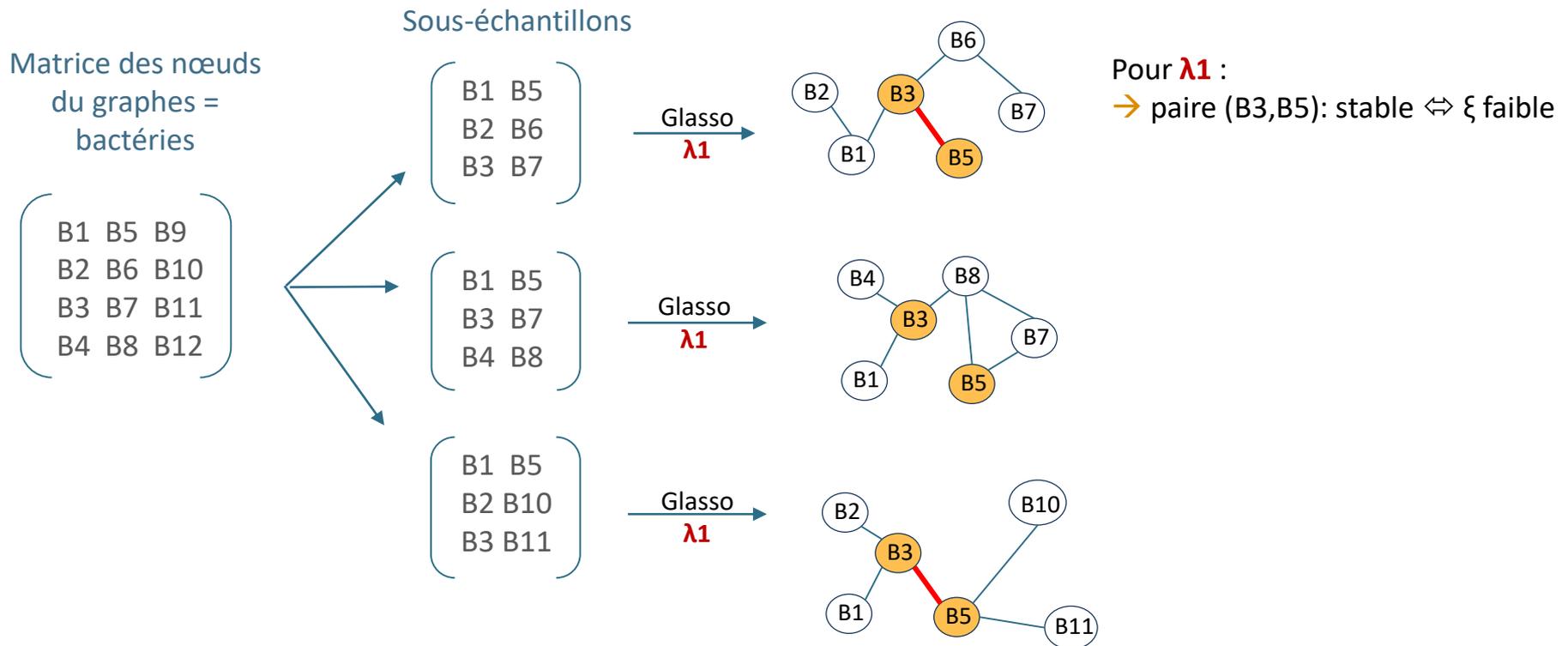


λ_3

➔ Quel lambda choisir ?

StARS (*Stability Approach to Regularization Selection*)

- But de StARS: Sélectionner les liens **robustes** (= retrouvés dans un maximum de sous-échantillons)
- Calcul d'une **moyenne d'instabilité** ξ



StARS (*Stability Approach to Regularization Selection*)

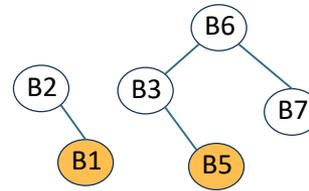
Matrice des nœuds
du graphes =
bactéries



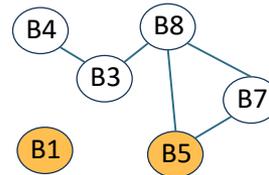
Sous-échantillons



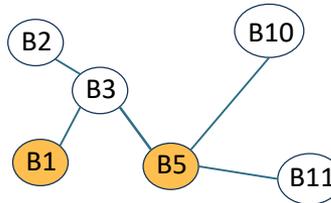
→ Glasso
 λ_1



→ Glasso
 λ_1



→ Glasso
 λ_1



Pour λ_1 :

→ paire (B3,B5): stable $\Leftrightarrow \xi$ faible

→ paire (B1,B5): stable $\Leftrightarrow \xi$ faible

StARS (*Stability Approach to Regularization Selection*)

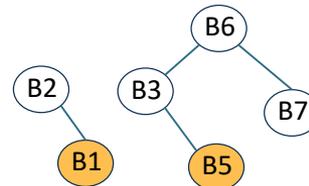
Matrice des nœuds
du graphes =
bactéries

B1	B5	B9
B2	B6	B10
B3	B7	B11
B4	B8	B12

Sous-échantillons

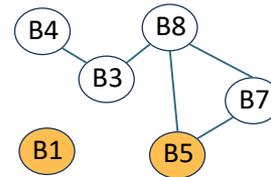
B1	B5
B2	B6
B3	B7

→
Glasso
 λ_1



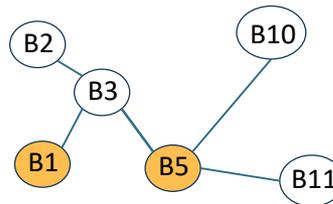
B1	B5
B3	B7
B4	B8

→
Glasso
 λ_1



B1	B5
B2	B10
B3	B11

→
Glasso
 λ_1



Pour λ_1 :

- paire (B3,B5): stable $\Leftrightarrow \xi$ faible
- paire (B1,B5): stable $\Leftrightarrow \xi$ faible
- paire (Bx,By) ...

→ Moyenne des ξ_1 pour λ_1
== Moyenne d'instabilité

StARS (*Stability Approach to Regularization Selection*)

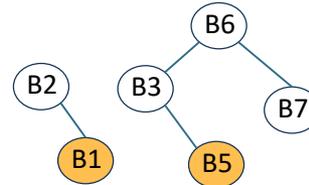
Matrice des nœuds
du graphes =
bactéries

B1	B5	B9
B2	B6	B10
B3	B7	B11
B4	B8	B12

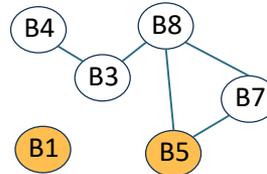
Sous-échantillons

B1	B5
B2	B6
B3	B7

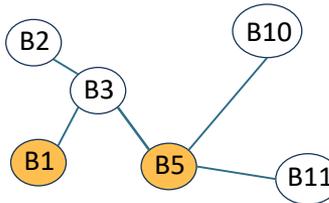
→
Glasso
 λ_2



→
Glasso
 λ_2



→
Glasso
 λ_2



Pour λ_2 :

- paire (B3,B5): stable $\Leftrightarrow \xi$ faible
- paire (B1,B5): stable $\Leftrightarrow \xi$ faible
- paire (Bx,By)

→ Moyenne des ξ_2 pour λ_2
== **Moyenne d'instabilité**

StARS (Stability Approach to Regularization Selection)

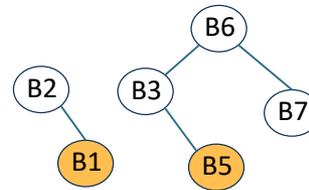
Matrice des nœuds
du graphes =
bactéries

$$\begin{pmatrix} B1 & B5 & B9 \\ B2 & B6 & B10 \\ B3 & B7 & B11 \\ B4 & B8 & B12 \end{pmatrix}$$

Sous-échantillons

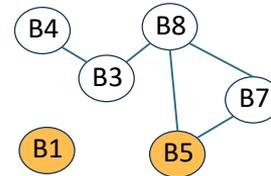
$$\begin{pmatrix} B1 & B5 \\ B2 & B6 \\ B3 & B7 \end{pmatrix}$$

Glasso
 λ_3



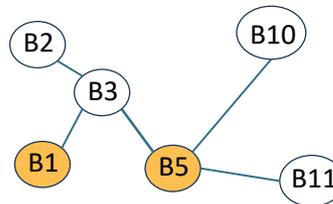
$$\begin{pmatrix} B1 & B5 \\ B3 & B7 \\ B4 & B8 \end{pmatrix}$$

Glasso
 λ_3



$$\begin{pmatrix} B1 & B5 \\ B2 & B10 \\ B3 & B11 \end{pmatrix}$$

Glasso
 λ_3

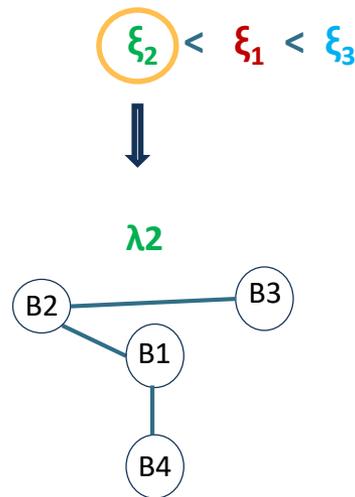


Pour λ_3 :

- paire (B3,B5): stable $\Leftrightarrow \xi$ faible
- paire (B1,B5): stable $\Leftrightarrow \xi$ faible
- paire (Bx,By)

→ Moyenne des ξ_3 pour λ_3
== Moyenne d'instabilité

❖ Quel lambda choisir ?

➔ Celui dont la moyenne d'instabilité est **minimale**

Fouille de réseaux et visualisation

Transformation
des données



Inférence
de
réseaux

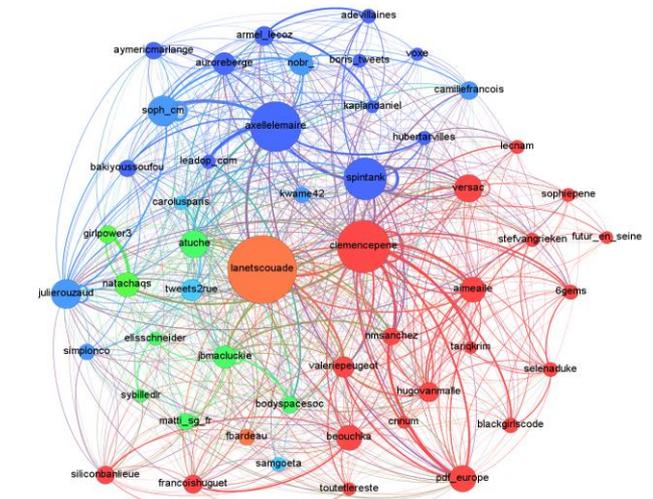


Sélection



Fouille de
réseaux &
visualisation

Exemple de réseau final :



Conclusion

❖ But de l'inférence et de la fouille de réseaux → caractériser le microbiote

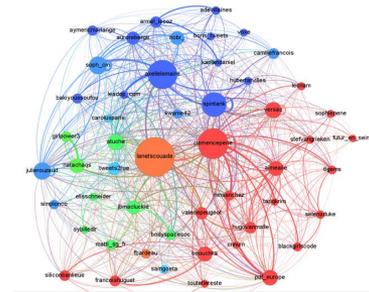


Co-abondances



communautés
& hub

❖ Réseau → Vue d'ensemble



❖ Beaucoup d'outils développés ces dernières années



❖ Glasso → la référence pour l'inférence de réseaux

❖ D'autres outils n'utilisent pas Glasso

- ❖ Comparer

- ❖ Tester

➔ Lequel semble le plus adapté :

- métagénomique quantitative
- microbiote intestinal

❖ Inférence ➔ Fouille ➔ Faire le lien avec les données biologiques

Remerciements



Direction

Alexandre Cavezza
Florence Haimet

Sambo

Joël Doré
Hugo Roume
Sarah Jegou

MetaQuant

S. Dusko Ehrlich
Hugo Roume
Nathalie Galleron
Benoit Quinquis

Prépa

Anthony Dobez

InfoBioStat

S. Dusko Ehrlich
Nicolas Pons
Laurie Alla
Mathieu Almeida
Ariane Bassignani
Magali Berland
Samar Berreira Ibraim
Kevin Da Silva

MetaFun

Hervé Blottière
Véronique Léjard
Alexandre Evans
Parfait Evouna Mengue
Océane Kacimi

Léonard Dubois
Sébastien Fromentin
Louisa Hadj Abed
Franck Gauthier
Oscar Gitton-Quent
Susie Guilly
Marie Jeammet
Emmanuelle Le Chatelier

Business Development

Karine Valeille

Prévention / Qualité

Benoit Quinquis
Franck Gauthier
Nathalie Galleron

Soufiane Maski
Nicolas Maziers
Victoria Meslier
Florian Plaza-Oñate
Florence Thirion
Fatoumata-Adama Traoré
Kevin Weiszer
Sana Zaghouani

Gestion RH et Finances

Caroline Lion
Rebecca Valide



Merci de votre attention