



INSTITUT NATIONAL
DES SCIENCES
APPLIQUÉES
TOULOUSE



Utiliser des données méta-omiques pour piloter les procédés de méthanisation : le projet DeepOmics



Ariane Bize
ariane.bize@inrae.fr

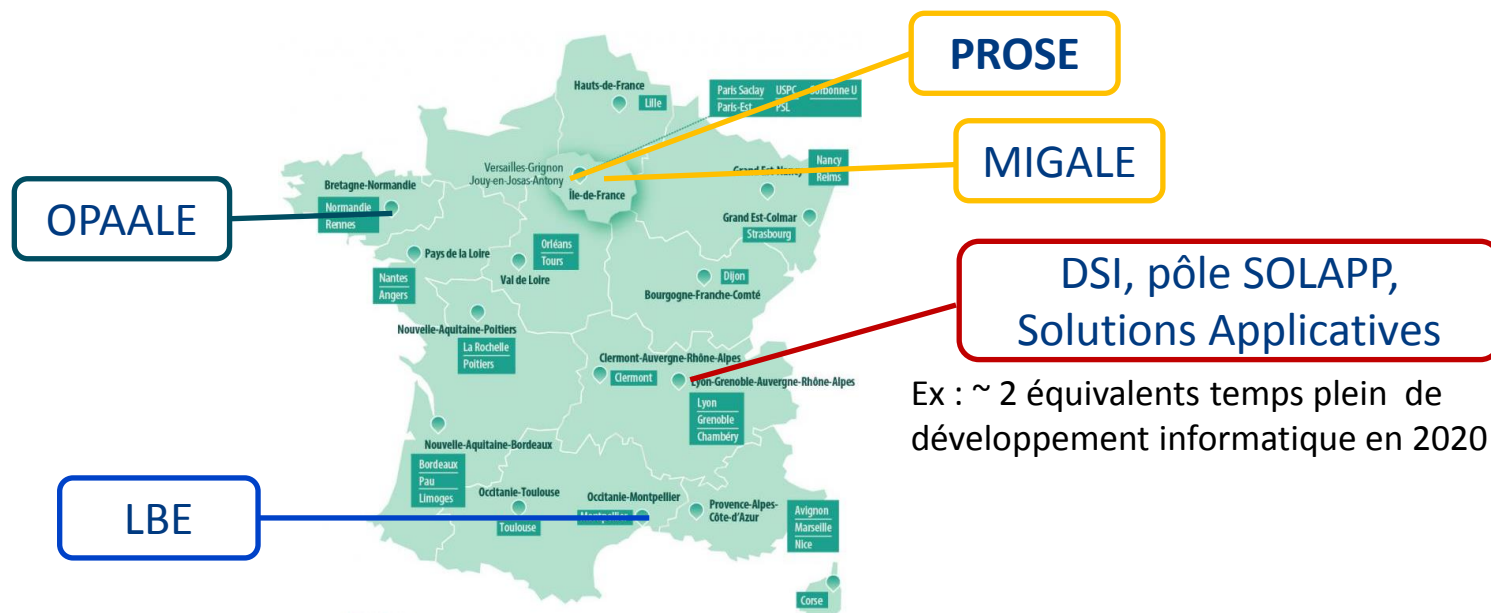


en partenariat avec



DeepOmics: un projet collaboratif INRAE

- **Projet INRAE** PROSE Antony, DSI Lyon, LBE Narbonne, OPAALE Rennes



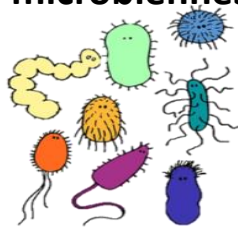
Ex : ~ 2 équivalents temps plein de développement informatique en 2020

- Perspectives : poursuivre le développement informatique et opérationnel de DeepOmics
- Etendre les collaborations à des partenaires extérieurs

Lier les performances des procédés au fonctionnement des communautés microbiennes

...

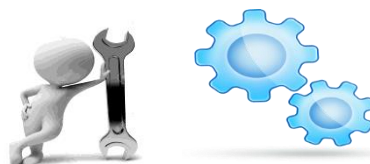
Structure, diversité et
fonctions métaboliques des
communautés
microbiennes



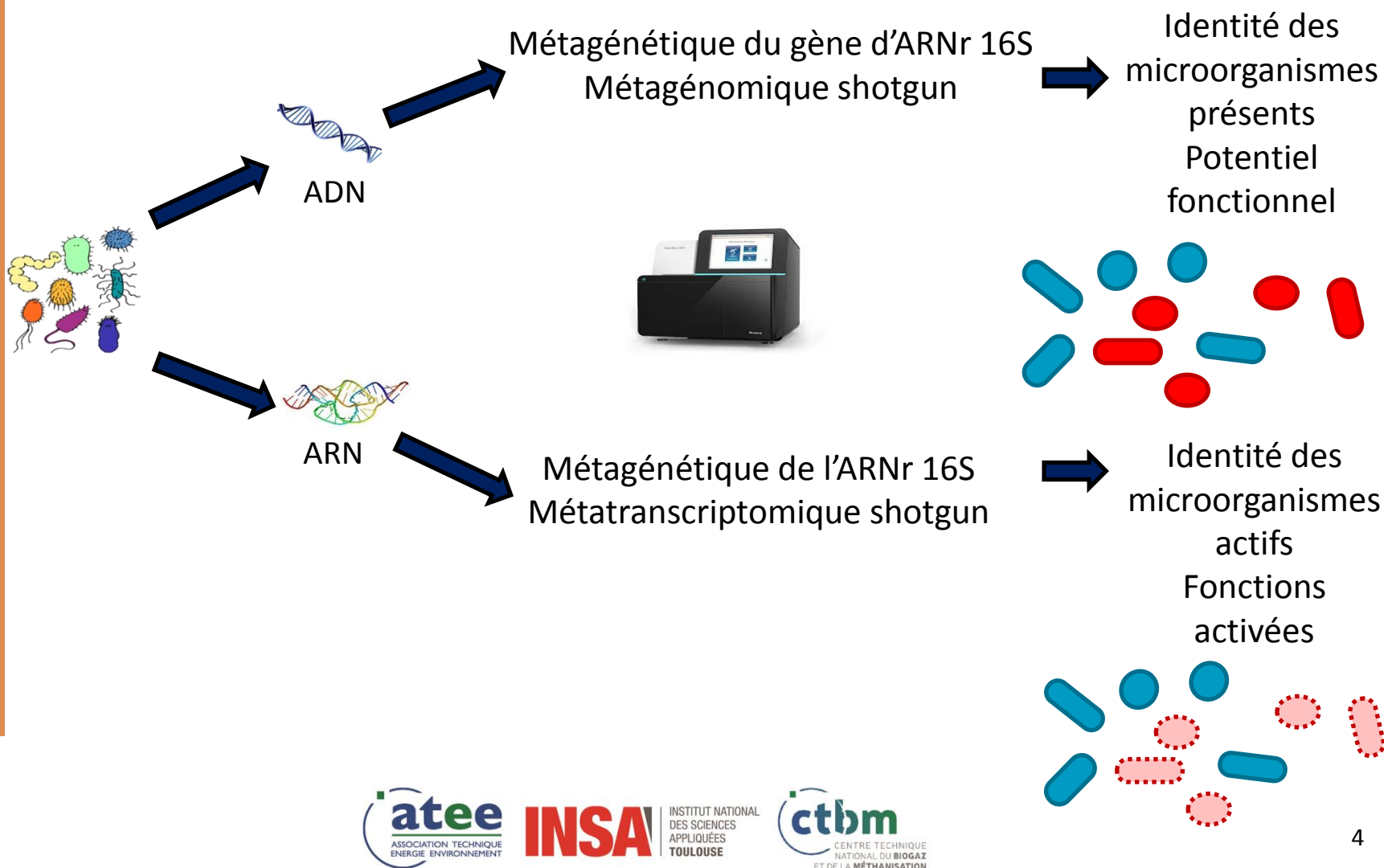
Performances et fonctions
du procédé



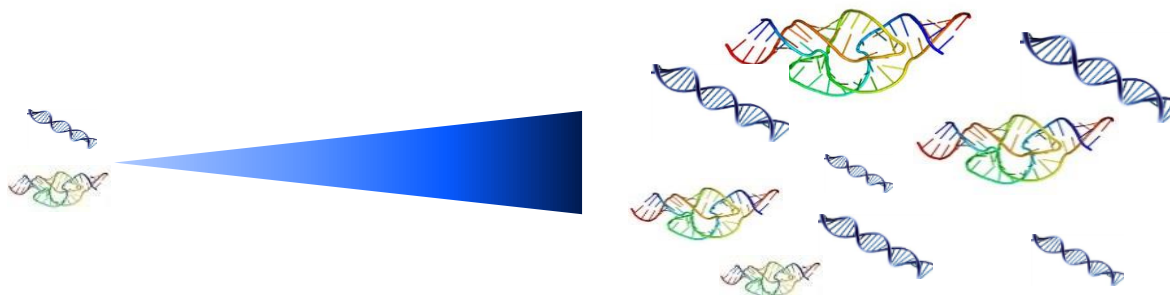
Paramètre opérationnels
Design du procédé



... à l'aide de méthodes « méta-omiques »



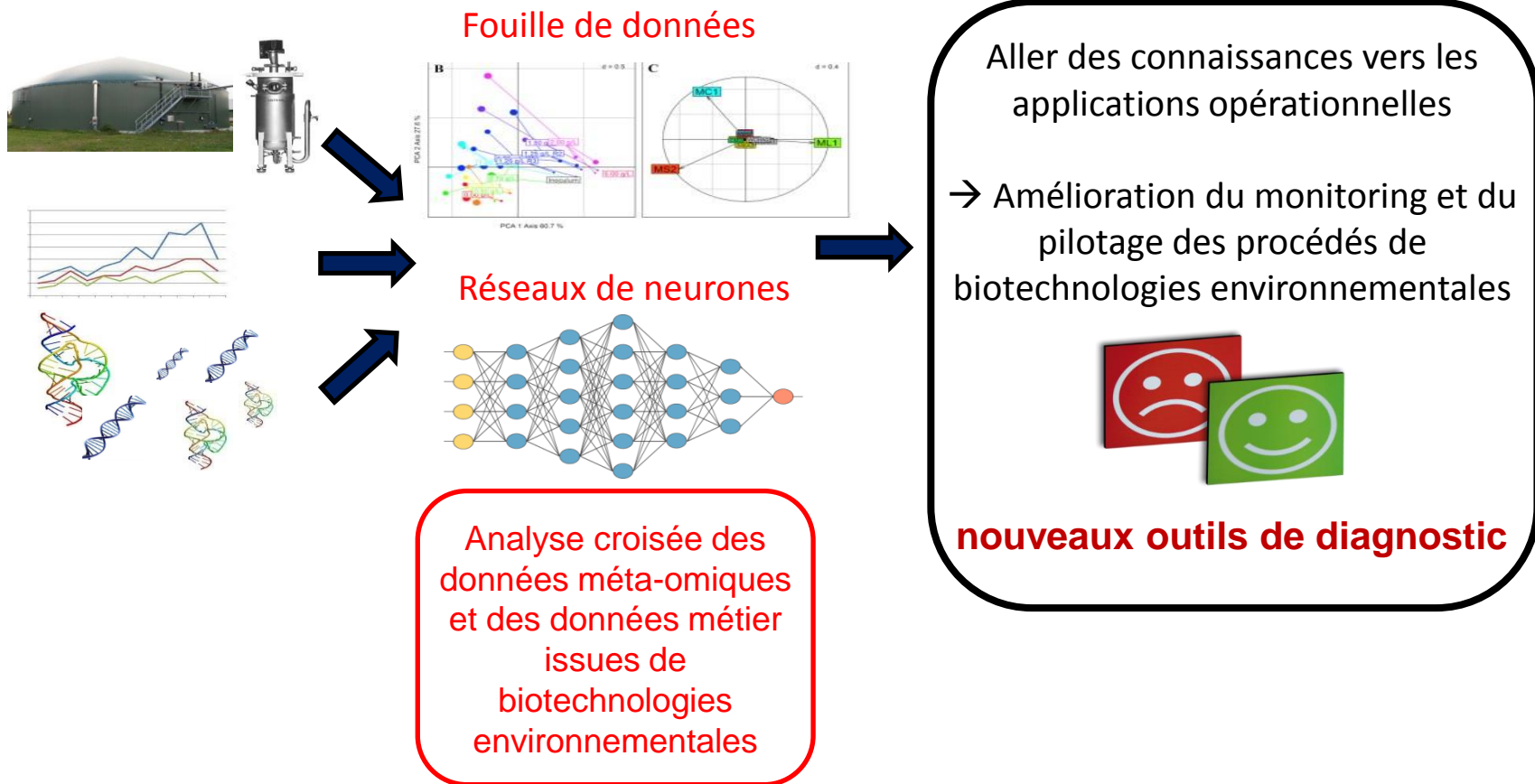
Les données méta-omiques s'accumulent rapidement ...



**mais les jeux de données sont souvent analysés
de manière isolée, à l'échelle de chaque projet**

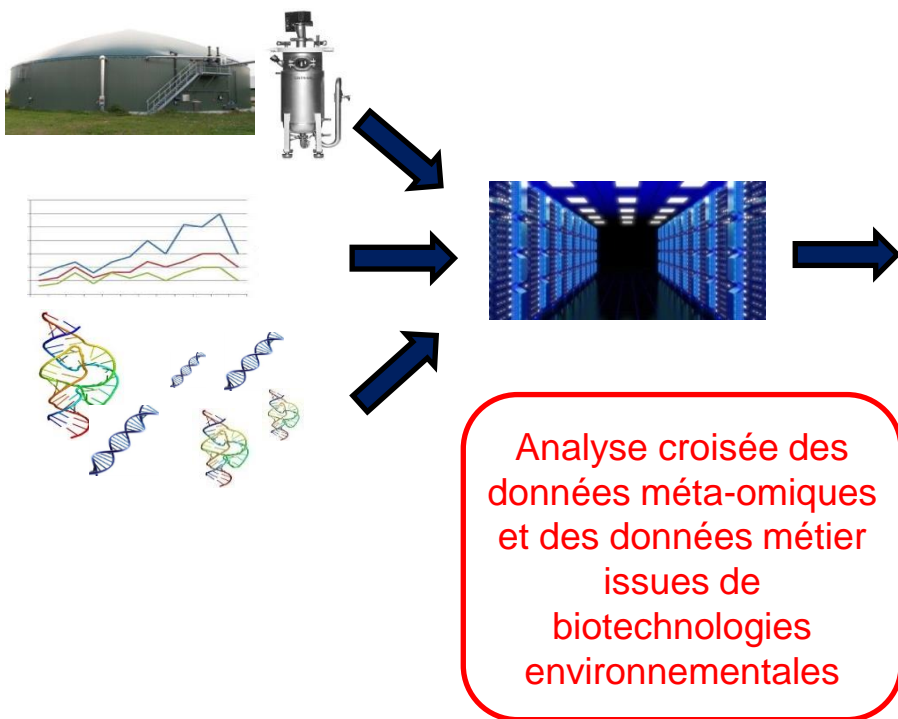


Pourrions-nous capitaliser sur ce type de données à plus large échelle ?



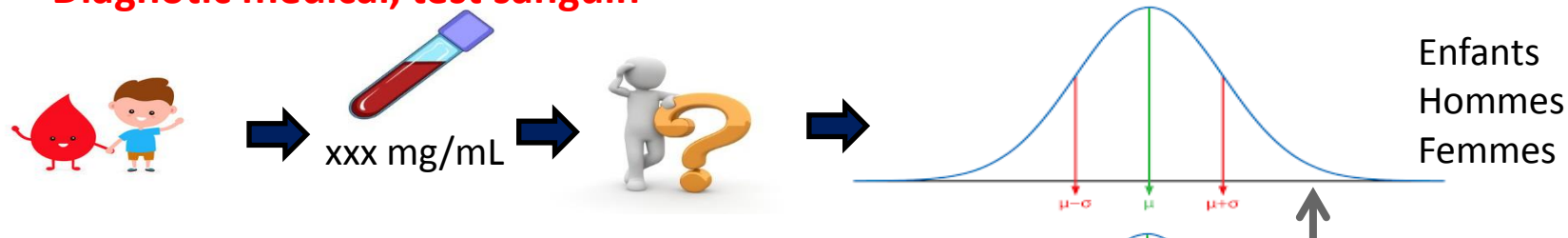
Le projet DeepOmics

- “Digital Environmental Engineering Platform for OMICS data”
- Développement d’un entrepôt dédié aux données méta-omiques de procédés de méthanisation et d’autres biotechnologies environnementales

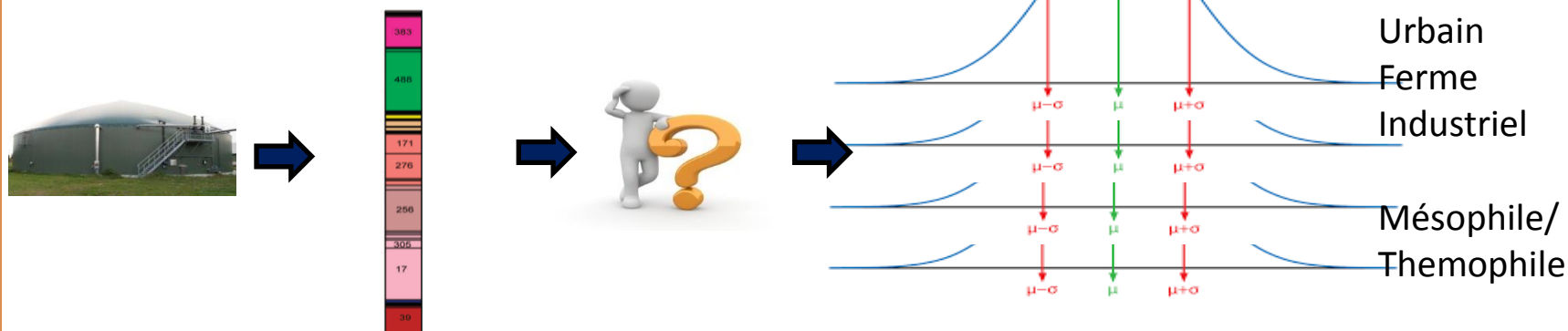


Le projet DeepOmics

Diagnostic médical, test sanguin

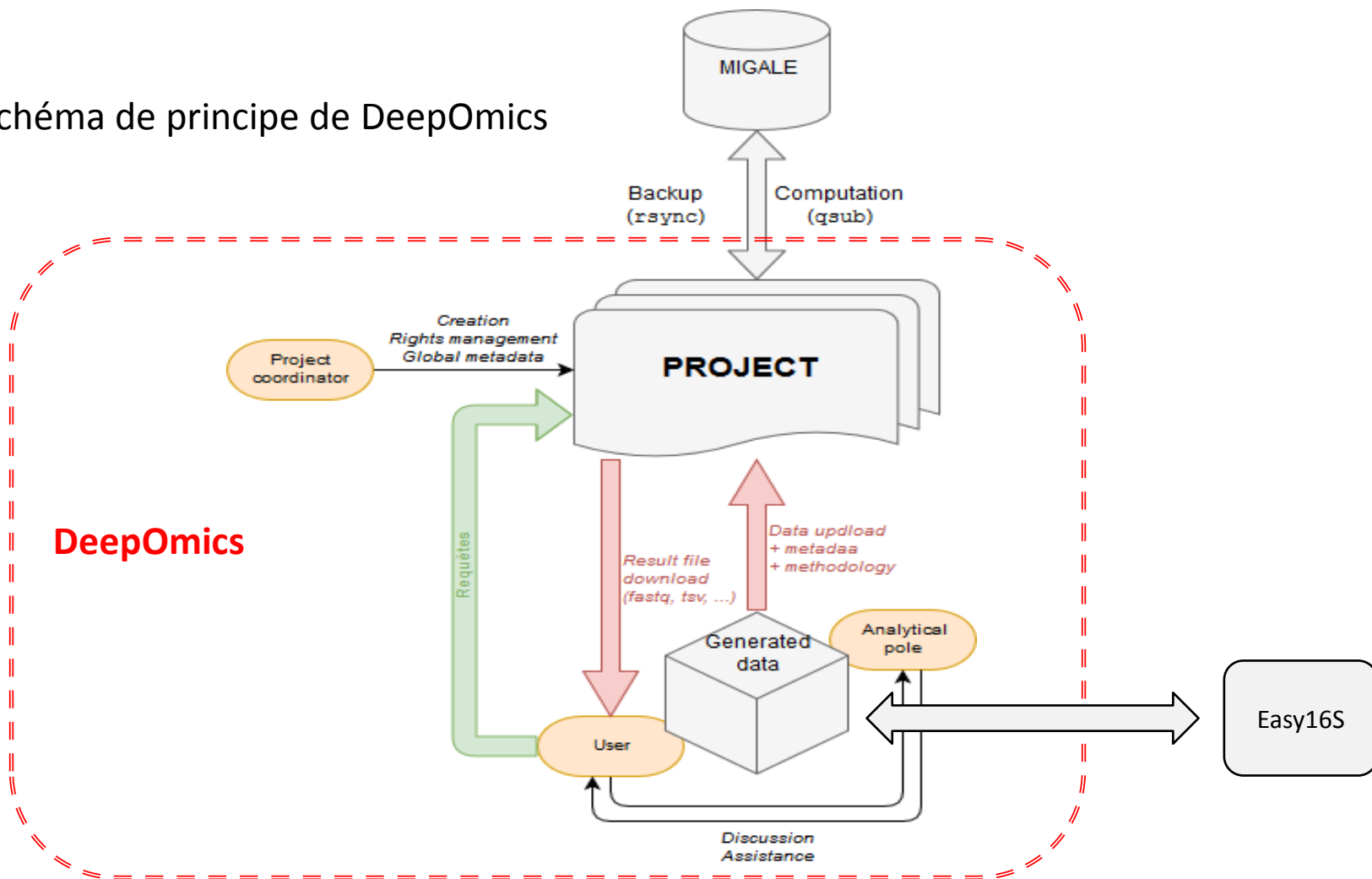


Diagnostic de la méthanisation, analyses « 16S »



Le projet DeepOmics

Schéma de principe de DeepOmics



Le projet DeepOmics

Objectifs

- Fournir un outil pour entreposer de manière organisée les données métiers et méta-omiques issues de procédés de méthanisation
- Favoriser les démarches de fouille de données
- Prendre en compte des données de laboratoire (pilotes) ou de terrain

<https://deepomics.test.irstea.fr/home>

Fonctionnalités clés de DeepOmics

- **Entreposage rationalisé** de données méta-omiques et de **métadonnées** issues d'installations industrielles et de réacteurs pilotes, à l'échelle de projets de recherche individuels

INRAE DeepOmics ariane.bize

Home | Projects | Meta-analysis

Project "Cellulose Nelly" [View details](#)

Lab-scale process data | Industrial and field process data | Meta-omics analysis

Laboratory

- Experimental series
- Project's input management

List of experimental series

Search... [+ New series](#)

	Experimental series	Contact person	Start date	Operators
>	LignocellulosicSubstrates Lignocellulosic substrates (complex substrates)	Ariane BIZE D'AVEZAC DE CASTERA	2011-01-15	Ariane BIZE D'AVEZAC DE CASTERA
>	CellulosicSubstrates Cellulosic substrates (simple substrates)	Marguerite BAZIN	2012-08-22	Martin BAILLY

Fonctionnalités clés de DeepOmics

- **Import et export de données** dans des formats standards (biom, fastq, fasta, csv, ... selon la nature des données)

Deepomics Log in

Cellulose Nelly Cotton Kleenex Tissue

Data Management

Add a monitored Measure: [1\) Generate csv file](#) [2\) Import data](#)

Number of Data :

Monitored Measure Type	Replicate Location															
	T1	T2	K1	K2	K3	K4	C1	C2	C3	C4	P1	P2	P3	P4		
Acetate concentration (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	
Butanol concentration (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	
Butyrate concentration (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	
Dissolved Inorganic Carbon (mg/L)			18 data	18 data	18 data	18 data	17 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	18 data	

ASSOCIATION TECHNIQUE ENERGIE ENVIRONNEMENT
 DES SCIENCES APPLIQUÉES TOULOUSE
 CENTRE TECHNIQUE NATIONAL DU BIOGAZ ET DE LA MÉTHANISATION

Fonctionnalités clé de DeepOmics

- **Visualisation** de graphiques en courbes définis par l'utilisateur, basés sur les métadonnées (mesures physico-chimiques), au sein d'un tableau de bord



Outil complémentaire : Easy16S

- **Analyses statistiques** et visualisation de données de métabarcoding 16S et métagénomiques grâce à l'utilisation de l'application conviviale **Easy16S** (<http://genome.jouy.inra.fr/shiny/easy16S/>)



Fonctionnalités clé de DeepOmics

- Réaliser des requêtes au sein de données et de méta-données, de réacteurs de laboratoires et d'installations industrielles
- Utiliser la partie publique de l'entrepôt et ses propres données privées
- **Exemple d'analyse**

DeepOmics

- *Requête dans DeepOmics pour trouver toutes les données de séquençage 16S et toutes les analyses physico-chimiques d'échantillons issus de séries d'expériences pour lesquelles du phénol ou de l'ammonium ont été introduits dans certains des réacteurs*
- *Première exploration statistique avec Easy16S*
- *Export des résultats dans des formats standards*
- *Analyses statistiques spécifiques avec les outils préférés de l'utilisateur*

Exemple d'application possible

2 séries d'expériences

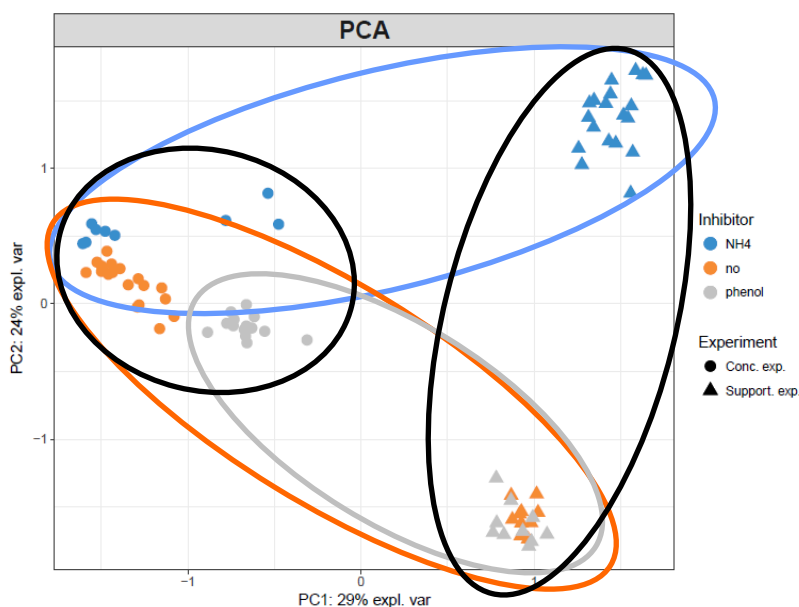
144 échantillons analysés par métabarcoding 16S

Inhibé par
le phénol

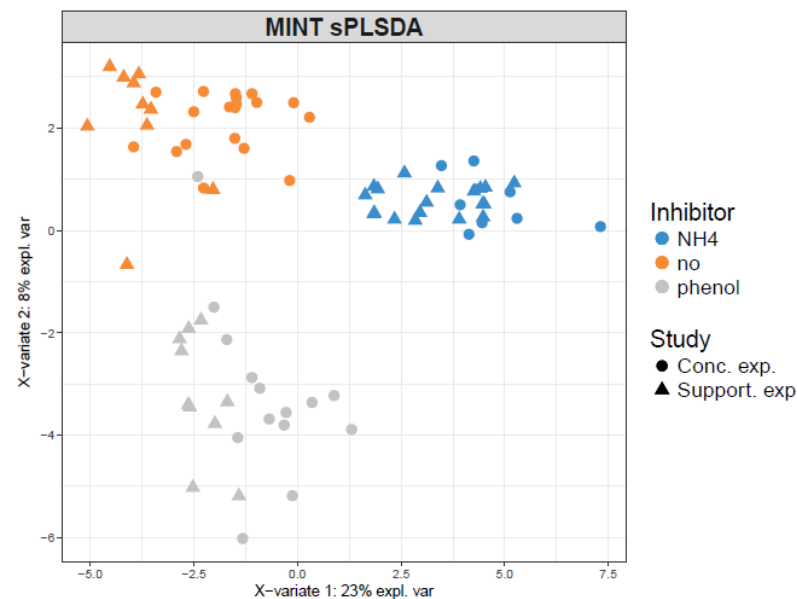
Non
inhibé

Inhibé par
L'ammoniaque

Ordination des échantillons par ACP



Multivariate INTEgrative method, MINT



Important effet "étude"
Pas de possibilité de généraliser ?

Correction de l'effet "étude"
Les résultats peuvent
être généralisés



Conclusions

- **DeepOmics**
 - Un entrepôt de données dédié au **stockage et à l'analyse croisée** de données méta-omiques et de procédés, sur la méthanisation
 - Un environnement pour préparer des approches de méta-analyse et de fouille de données
 - Un outil pour aider les professionnels à produire des données FAIR*
- **Notre approche est très complémentaire de la base de données de connaissance "MiDAS", du groupe danois de Per Nielsen**

McIlroy, Simon Jon, et al. "MiDAS: the field guide to the microbes of activated sludge." *Database* (2015).

* Findable, Accessible, Interoperable, and Reusable



Merci pour votre attention !



Remerciements aux différentes personnes impliquées



Théodore Bouchez
Cédric Midoux
Yannick Fayolle

Elie Le Quémener
Virginie Rossard
Eric Latrille



**DSI
Irstea
puis
INRAE**

Nicolas Raidelet
Guillaume Perréal
Aurélie Gramusset
Manon Predhumeau
Christophe Biasini
Harold Boissenin
Walter Roa Serrano

Patrick Dabert
Valentin Loux

