

Plateforme d'analyse ABRomics

Un retour d'expérience sur le développement d'une plateforme web nationale d'analyse bioinformatique

Julie Lao

Institut Français de Bioinformatique (CNRS) Hub de Bioinformatique (Institut Pasteur)











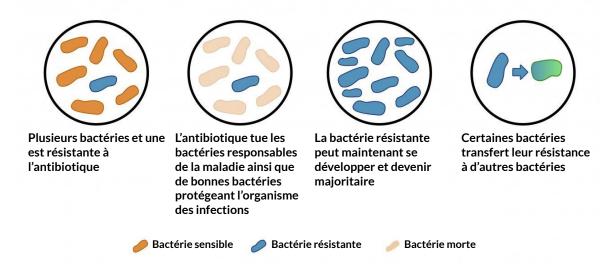






Les bactéries, virus, champignons et parasites ne réagissent plus aux antimicrobiens.

Comment la résistance aux antibiotiques apparaît?



Source: AdobeStock



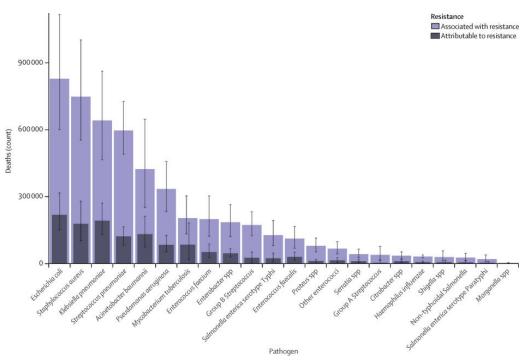


Figure 2 - Antimicrobial Resistance Collaborators et al, The Lancet, 2022.



Objectifs d'ABRomics

- Amélioration des connaissances sur la circulation des bactéries résistantes aux antibiotiques (gènes de résistance aux antibiotiques et éléments génétiques mobiles)
- Simplification des analyses à travers des pipelines standardisés et contrôlés :
 - Utilisable par les non bio-informaticiens et les experts en analyses des résistances antimicrobiennes
 (AMR) et de l'antibiorésistance (ABR)
- Intégration/interopérabilité des outils et des données (multi-omiques) pour limiter les analyses d' échantillons (méta) génomiques bactériens cloisonnées/individuelles
- Assurer la pérennité du stockage et de partage de données ; soumission dans les « entrepôts » internationaux (ENA)





⇒ AAP « structurants » du PPR antibiorésistance
 4 projets dont ABRomics « A multi-omics data platform »
 2 M€ pour 4+1 ans début 01/11/2021

⇒ Conception d'une plateforme web partagée pour améliorer la <u>surveillance</u> et la <u>recherche</u> en AMR dans un <u>objectif « One Health »</u>

L'organisation dans le projet







Co-coordination:

C. Médigue

(Institut Français de Bioinformatique)

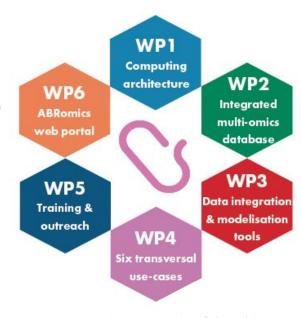
&

P. Glaser

(Institut Pasteur)

Consortium multidisciplinaire national de 43 partenaires

- Microbiologistes
 - Centres nationaux de référence (CNR)
 - Hospitalo-universitaires
 - Vétérinaires (ANSES)
 - Chercheurs
- Bioinformaticien·nes et mathématicien·nes
 - Plateforme d'ingénierie
 - Équipes de R&D

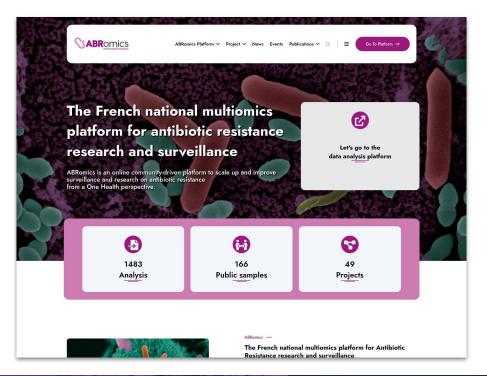


WP1, WP2 et WP3 : objectifs en relation avec l'AAP du PPR

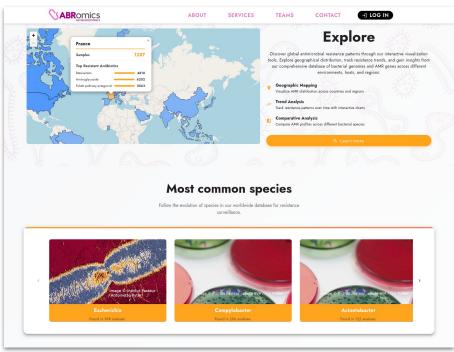
- Utilisateurs des trois secteurs : humain, animal et environnement
- Laboratoires de microbiologie (Clinique, vétérinaire, agroalimentaire, environnement ...)
 - Analyser les données et les stocker
- Les laboratoires de santé publique pour aider à la surveillance
 - Intégrer les données génomiques globales à la surveillance et assurer un suivi spatio-temporel (sur plusieurs années et de manière rétrospective)
- Laboratoires de recherche
 - Stocker, partager, soumettre les données, notamment pour les plans de gestion des données
 - Utiliser des méthodes développées par le consortium

Le portail ABRomics & La plateforme d'analyse ABRomics

Portail ABRomics: Information sur le projet et le consortium https://www.abromics.fr/



ABRomics analysis: La plateforme web d'analyse https://www.analysis.abromics.fr/





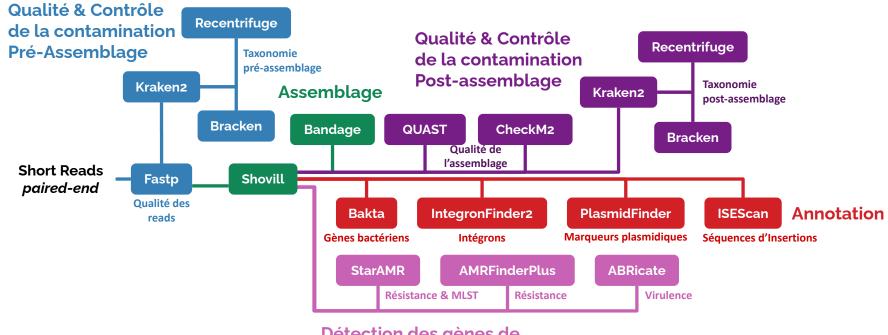








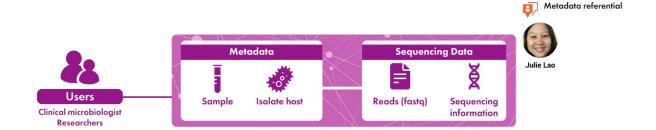




Détection des gènes de résistances aux antimicrobiens



Version alpha — Développement du 1^{er} référentiel de métadonnées





Métadonnées pour les données de séquençage short read paired-end

Sample ID
Strain ID
Sample type
Sample source
Host species
Country
Region
Place
Collected date
Travel countries
Accession(s) number
Sample comment

16 champs décrivant l'échantillon biologique, l'isolat et les données de séquençage pour 1 échantillon

 11 champs obligatoires et 5 optionnels

Utilisation de vocabulaires contrôlés

- pour le stockage dans notre base de données
- formatés pour la publication des données et résultats sur l'ENA (ENA prokaryotic pathogen minimal sample checklist)

Microorganism scientific name

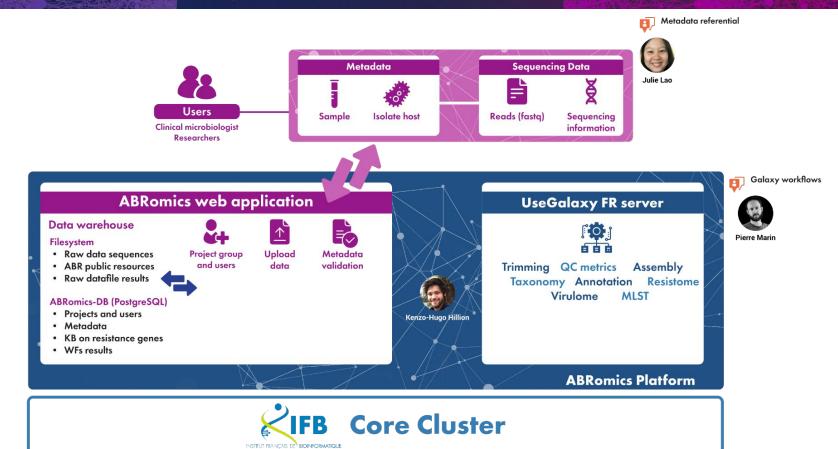
Instrument model R1 fastq filename

R2 fastq filename

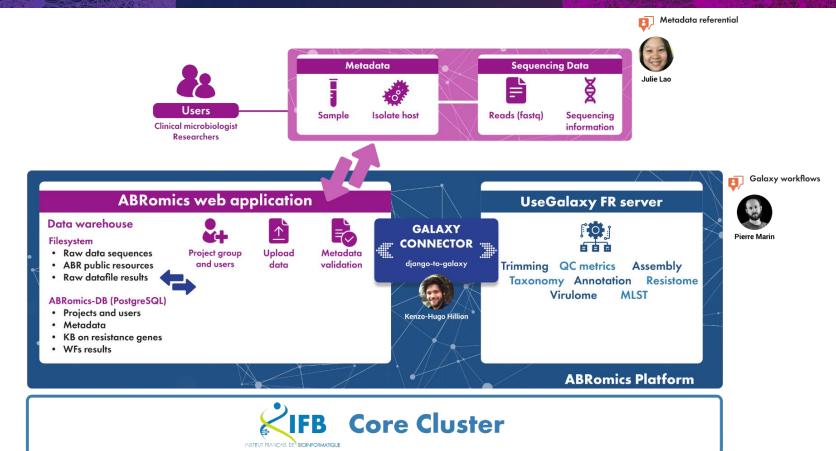
https://www.abromics.fr/home/abromics-platform/metadata-referential/



Version alpha — Mise en place du backend



Version alpha — Mise en place du backend



Django-to-Galaxy: Django extension that eases communication with Galaxy instance to execute workflows.

https://gitlab.com/ifb-elixirfr/abromics/django-to-galaxy

Django to Galaxy

The extension might be similar to this project.

Django extension that eases communication with Galaxy instance to execute workflows.

This should includes:

- · List available workflows on a Galaxy instance
- · Add selected workflow from Galaxy instance to the app for execution from admin panel
- · Run workflow from a Galaxy instance

For more information, please check the Wiki





https://pypi.org/project/django-to-galaxy/





Romain Dallet

Kenzo-Hugo Hillion



Fabien Mareuil

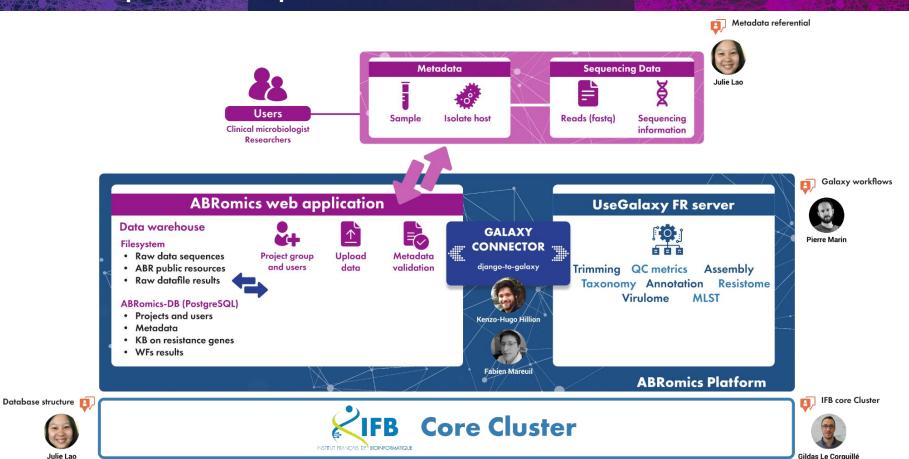


Amanda Dieuaide

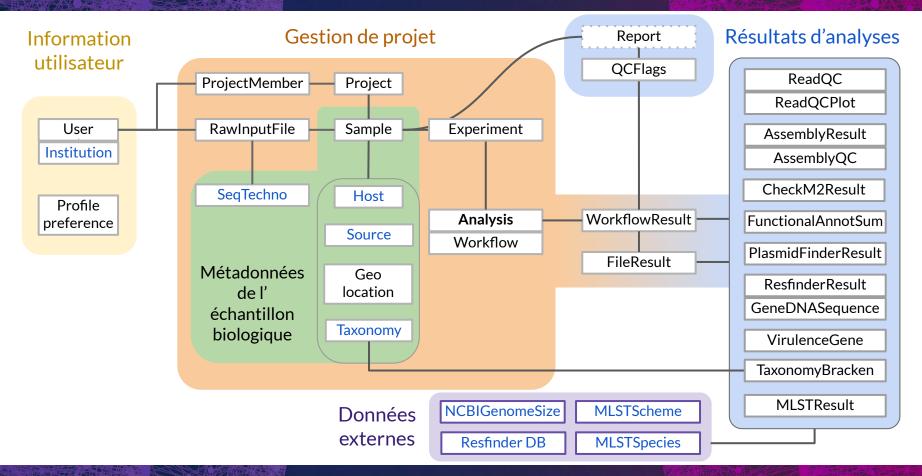
Raphaël Tackx

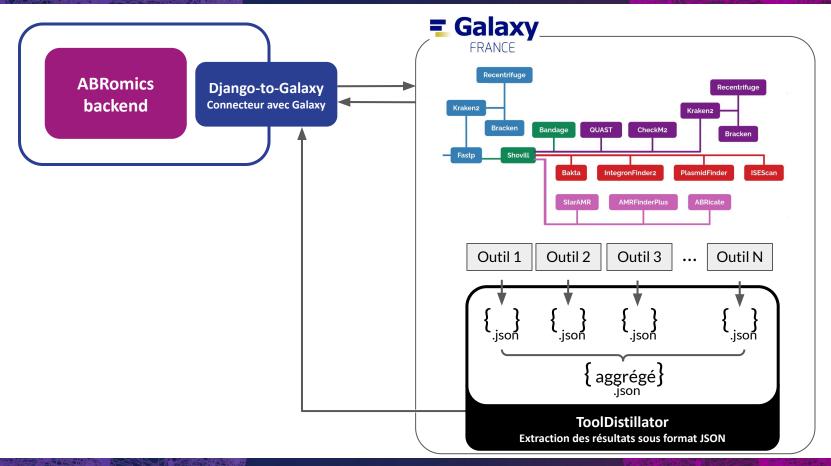


Version alpha — Conception des tables de la base de données



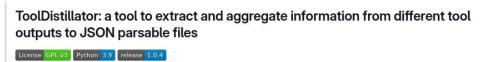
Base de données





ABRomics consortium (2023). **ToolDistillator: a tool to extract and aggregate information** from different tool outputs to JSON parsable files.

https://gitlab.com/ifb-elixirfr/abromics/tooldistillator



ToolDistillator is a tool to extract information from output files of specific tools, expose it as JSON files, and aggregate over several tools.

It can produce both a single file to each tool or a summarized file from a set of reports.

It was initially developed to be used on Galaxy and some options are only available on Galaxy (e.g. extract the historic ID from a galaxy analysis).

Tool was inspirated from the hAMRonization project (author: @dfornika, @fmaguire, @raphenya, @jodyphelan, @pvanheus)

https://bio.tools/tooldistillator













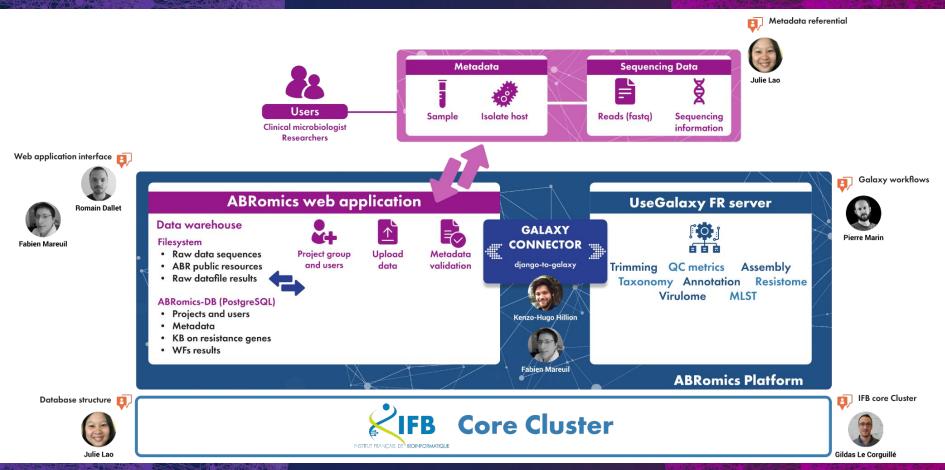




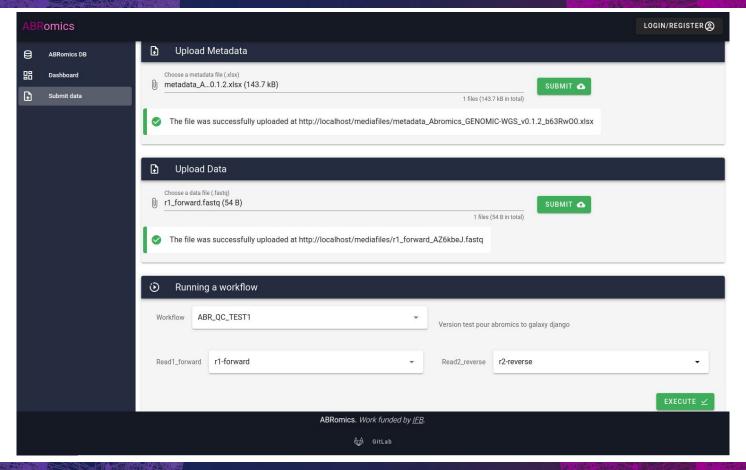




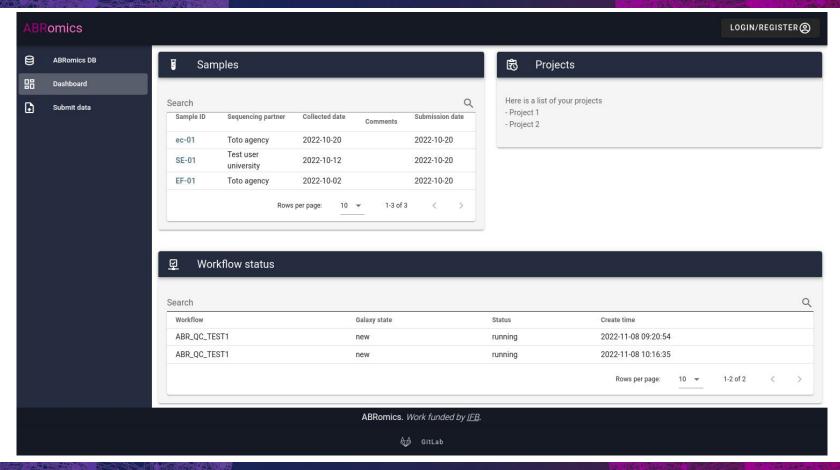
Version alpha — 1^{ère} interface web



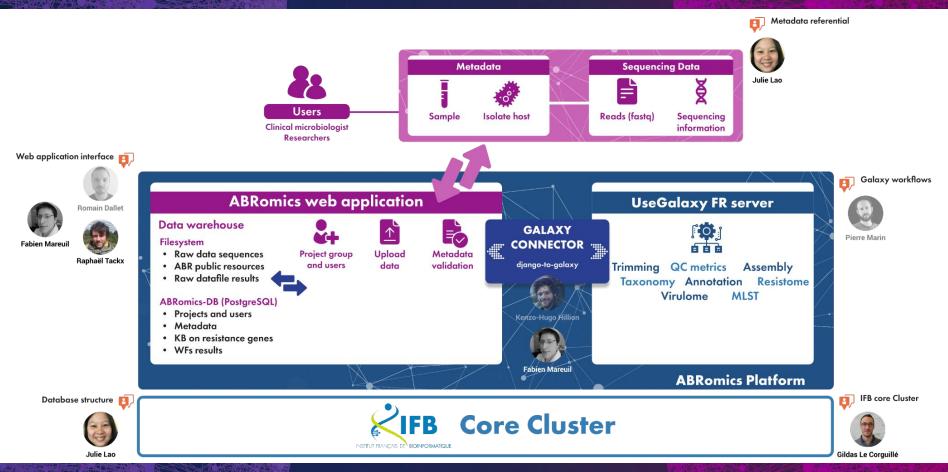
Version alpha — 1^{ère} interface web



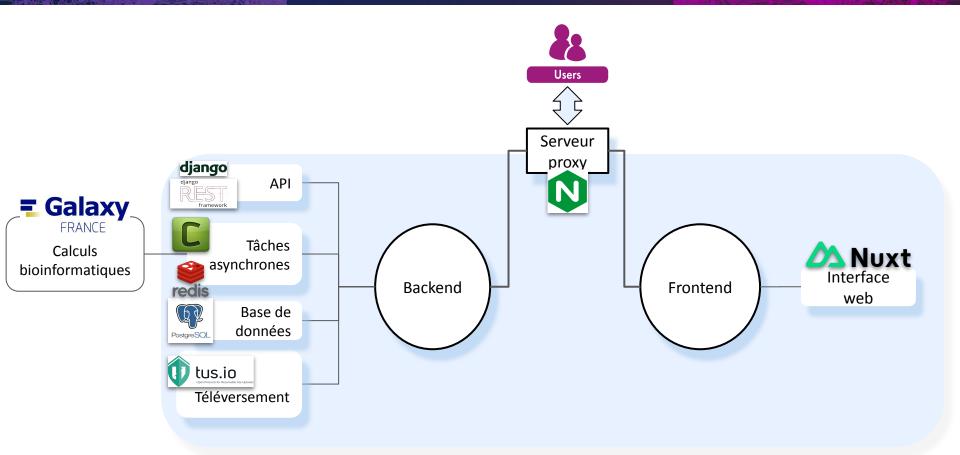
Version alpha — 1^{ère} interface web



Version alpha — Consolidation du backend



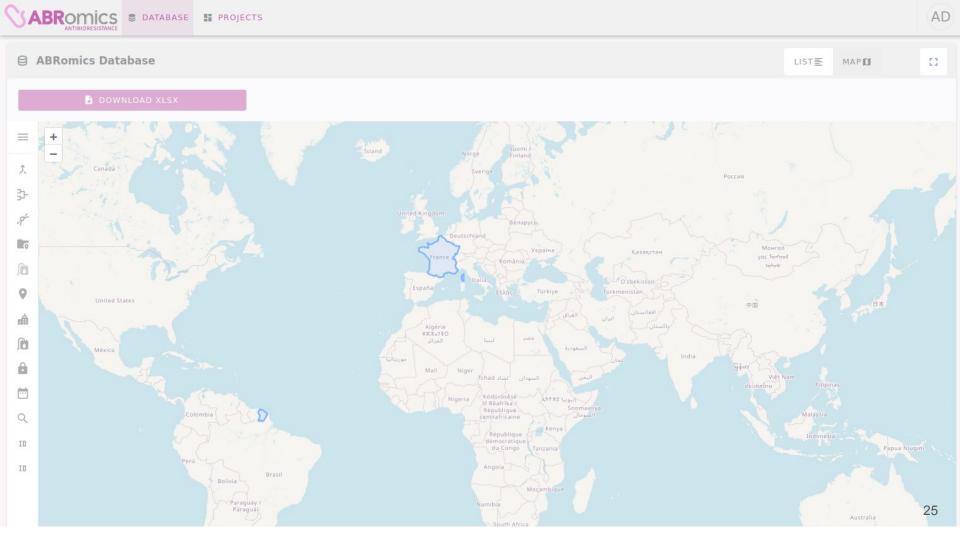
Architecture

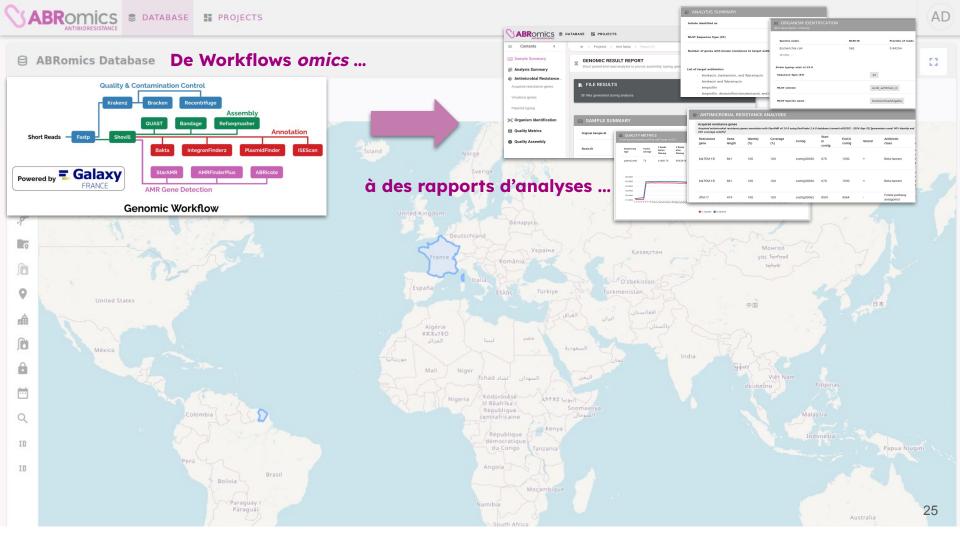


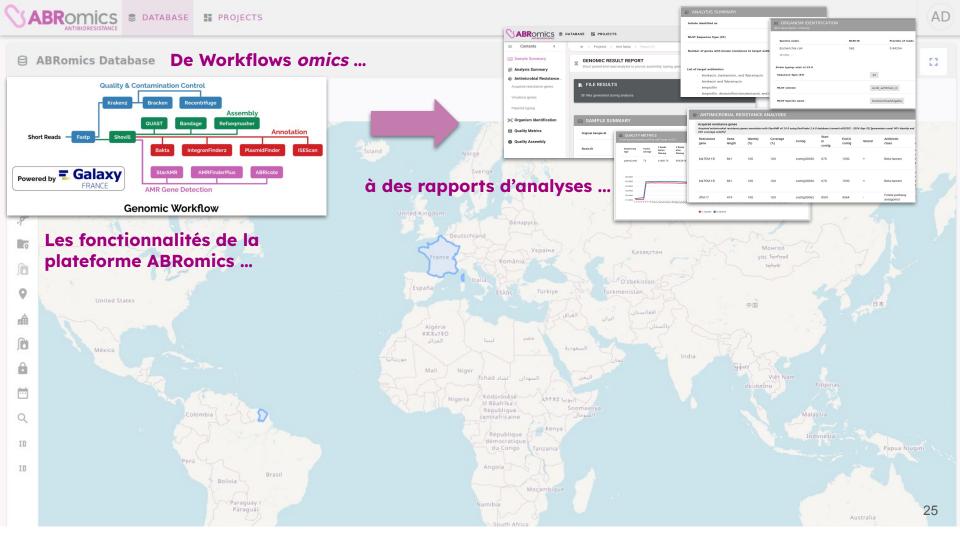
Julie Lao

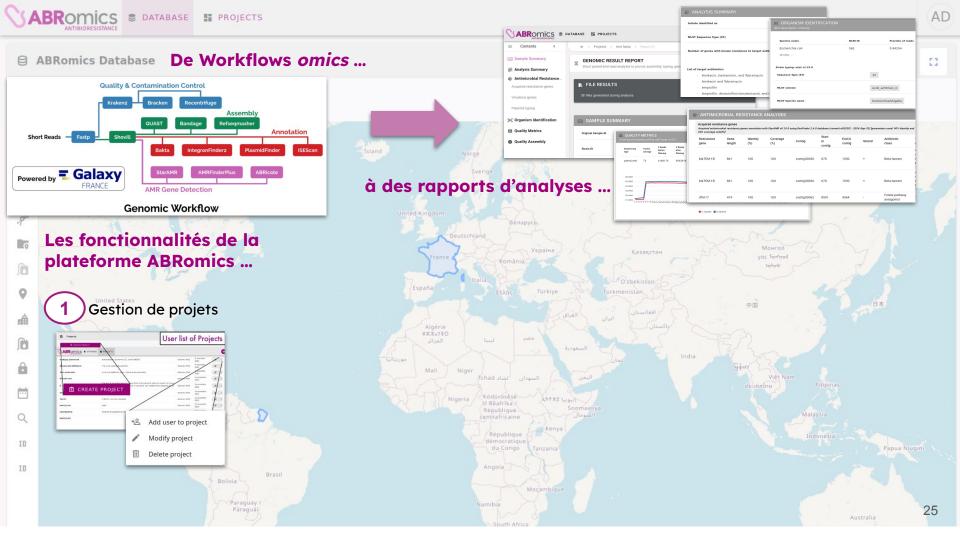


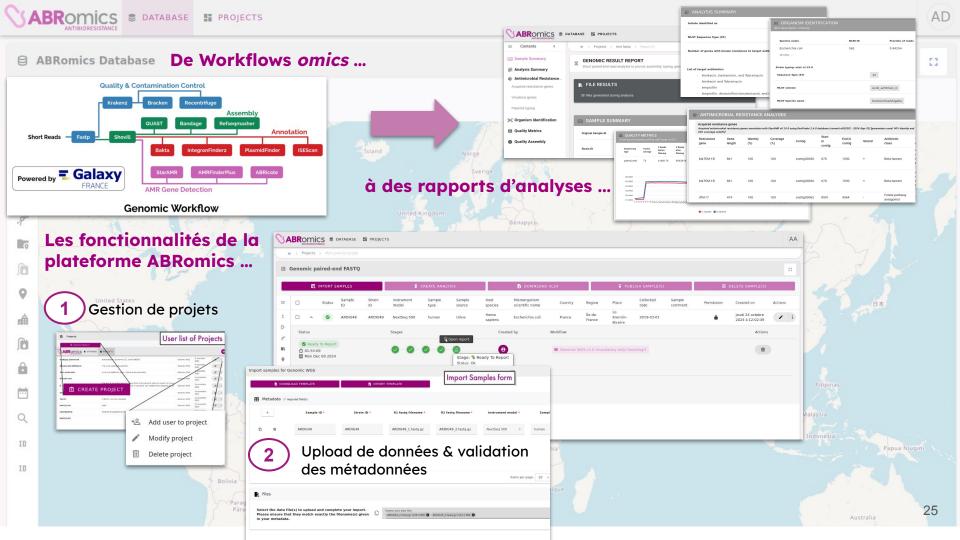
Gildas Le Corquillé

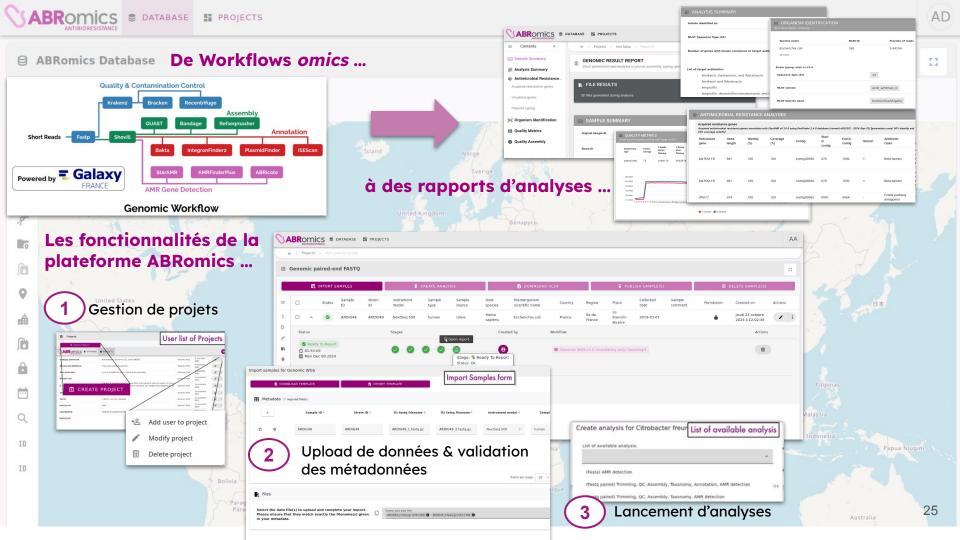


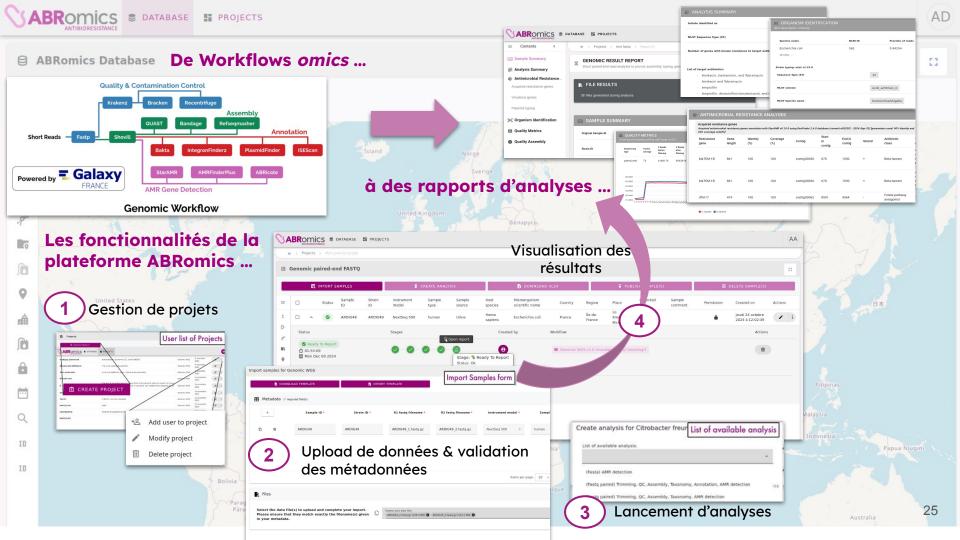


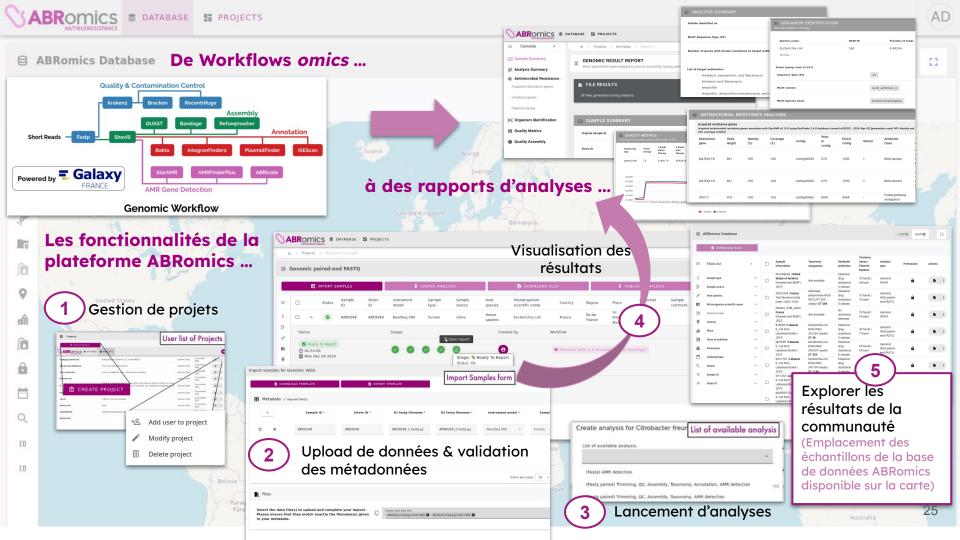




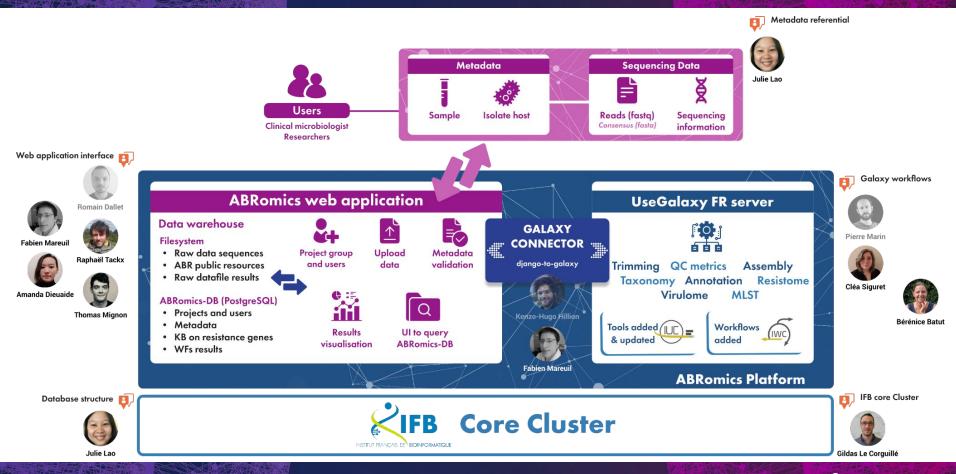




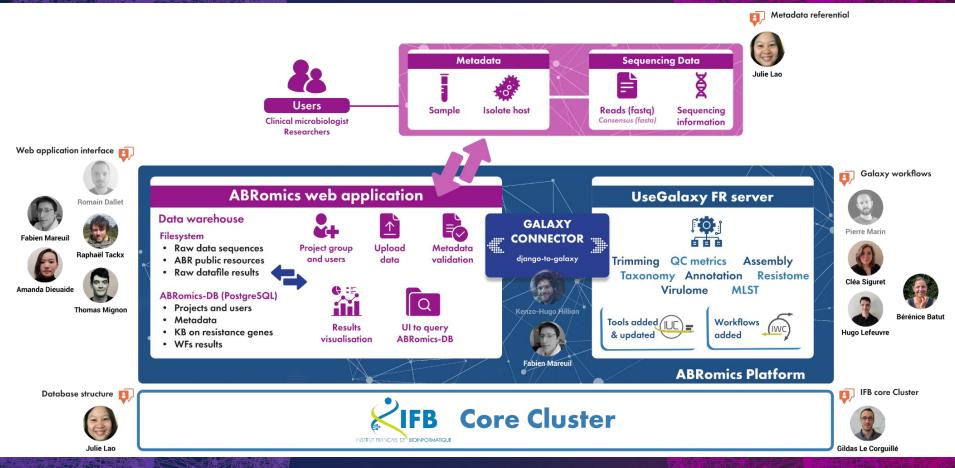




Version beta — Consolidation et amélioration des workflows



Version beta — Consolidation et amélioration des workflows



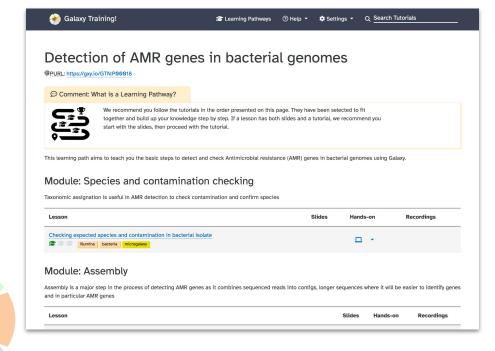
Sous-workflows short reads

 Disponible sur des registres de workflows: WorkflowHub et Dockstore via Galaxy IWC





 Documenté comme Learning Pathways sur le Galaxy Training Network https://gxy.io/GTN:P00018



FAIR — Partage et réutilisation

Tutoriel: Annoter, préparer, tester et publier ses workflows Galaxy dans les registres de workflows



Atelier(s): Promotion du partage et de la réutilisation des données de recherche autour des workflows Galaxy



https://hal.inrae.fr/hal-05095680v1

Réutilisation dans la formation Galaxy Avancée en 2026

https://gxv.io/GTN:T00533

Recherche d'éléments sur un contig contre 3 bases de données

- ResFinder (gènes de résistance)
- PointFinder (mutations ponctuelles confèrent une résistance)
- PlasmidFinder (marqueurs plasmidiques)

Limitations FAIR

- Versions des bases de données sont dépendantes de la version de l'outil
- Comment mettre à jour les bases de données dans usegalaxy?
 - Pull request, mettre à jour l'outil, mettre à jour le wrapper Galaxy
 - Ou montée de version manuelle par l'administrateur Galaxy

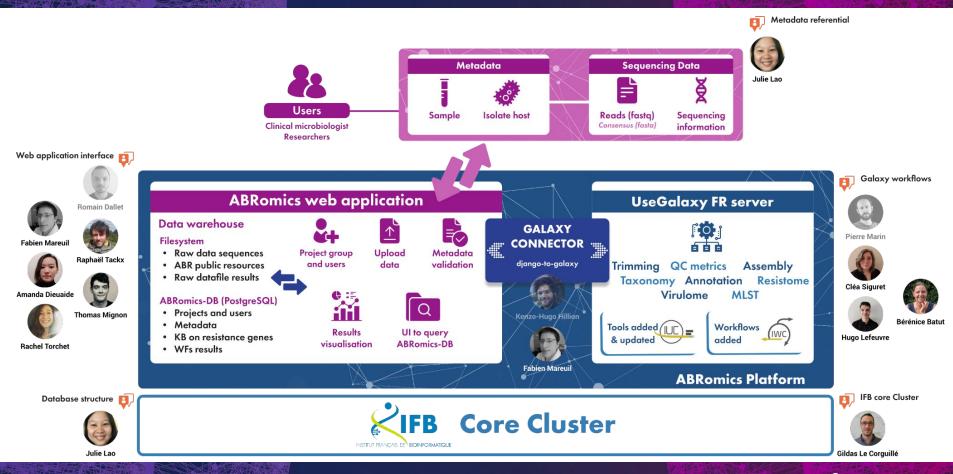
Amélioration réalisée

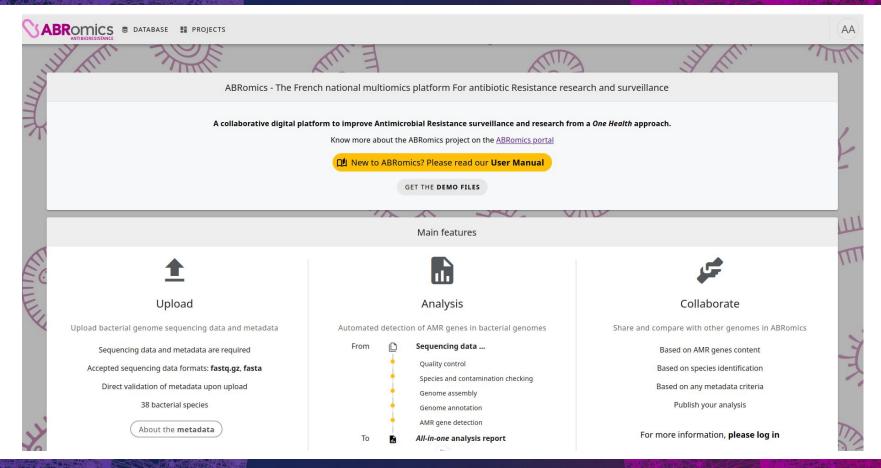
Data Manager : **Bonne pratiques** pour la gestion des données Galaxy



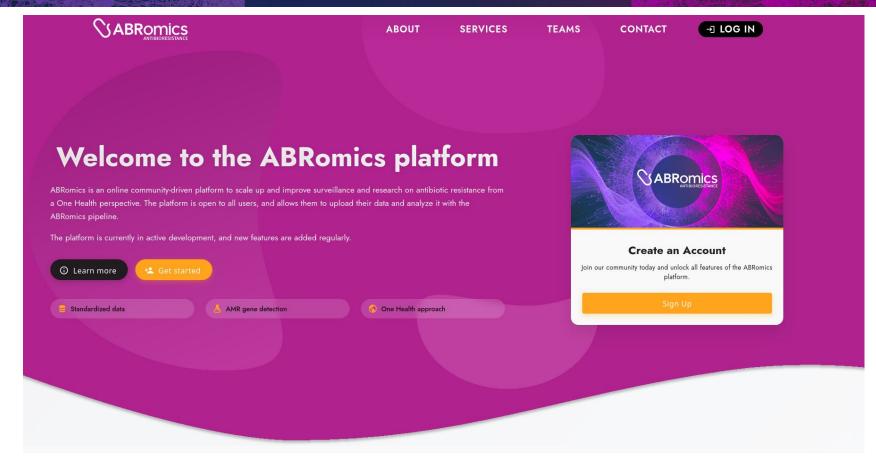


Version 1 de la plateforme ABRomics











Pratique mais on se rend dépendant d'une ressource qu'on maîtrise pas

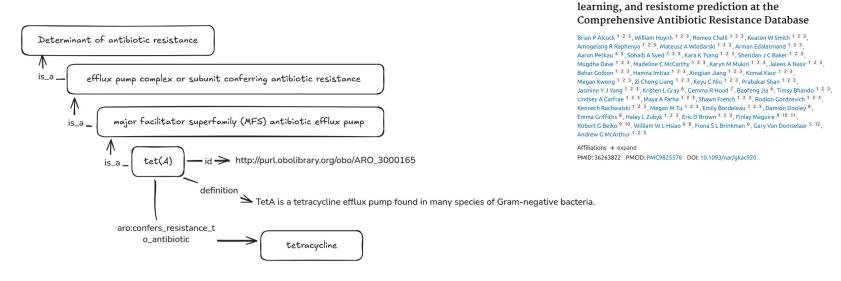
Données	Source des données	Récupération automatique?
Institution	ROR	API
Technologies de séquençages	?	?
Hôtes	?	?
Source	?	?
Taxonomie	NCBI Taxonomy	API
Taille de génome	NCBI Genome Size	API
Gènes de Resfinder	Resfinder DB	Repository Bitbucket
Schémas MLST	pubMLST, BigsDB,	API, FTP,

Pratique mais on se rend dépendant d'une ressource qu'on maîtrise pas

Données	Source des données	Récupération automatique?
Institution	ROR	API
Technologies de séquençages	?	?
Hôtes	Problème de sémantique Solution envisagée : Utilisation d'ontologies	
Source		
Taxonomie	NCBI Taxonomy	API
Taille de génome	NCBI Genome Size	API
Gènes de Resfinder	Resfinder DB	Repository Bitbucket
Schémas MLST	pubMLST, BigsDB,	API, FTP,

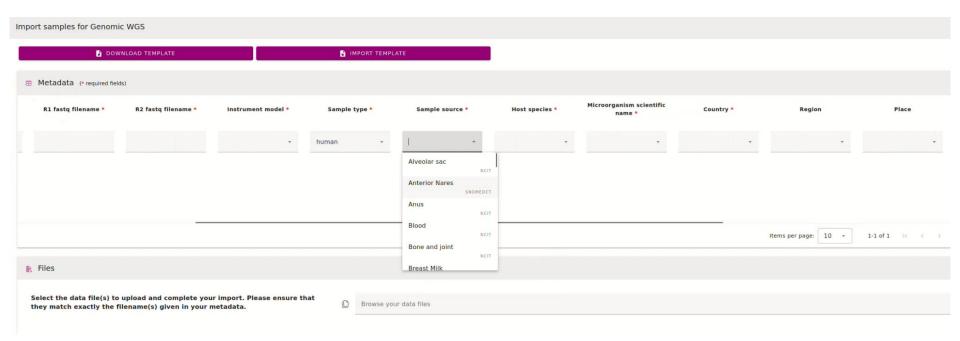
"A formal specification of a shared conceptualization" - Borst 1997

Antibiotic Resistance Ontology (ARO)



CARD 2023: expanded curation, support for machine

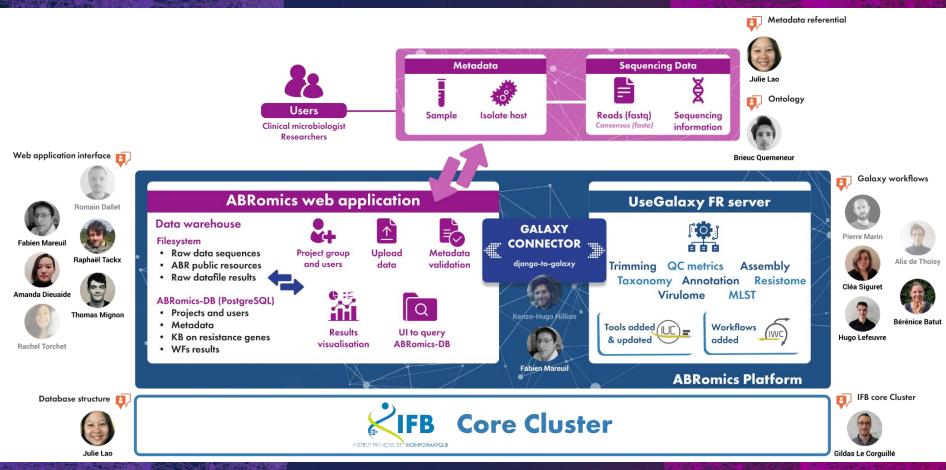
Ontologies



Prochaines étapes (déjà en cours) :

- cgMLST
- Librairie ABRomics pour les CNR
- La métagénomique

Dernier aperçu général



Merci pour votre attention!

