

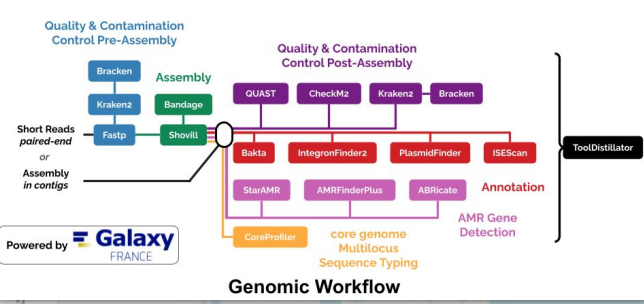
Plateforme d'analyse ABRomics

Julie Lao, Raphaël Tackx, Thomas Mignon, Fabien Mareuil

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



De Workflows omics ...



à des rapports d'analyses ...

GENOMIC RESULT REPORT

FILE RESULTS

SAMPLE SUMMARY

QUALITY METRICS

ANTIMICROBIAL RESISTANCE ANALYSES

ORGANISM IDENTIFICATION

Les fonctionnalités de la plateforme ABRomics ...

1 Gestion de projets

User list of Projects

CREATE PROJECT

- Add user to project
- Modify project
- Delete project

2 Upload de données & validation des métadonnées

Import Samples form

Metadata

Sample ID	Strain ID	R1 fastq filename	R2 fastq filename	Instrument model	Sample
ARDIG49	ARDIG49	ARDIG49_1.fastq.gz	ARDIG49_2.fastq.gz	NextSeq 500	human

Visualisation des résultats

4

3 Lancement d'analyses

Create analysis for Citrobacter freundii

List of available analysis:

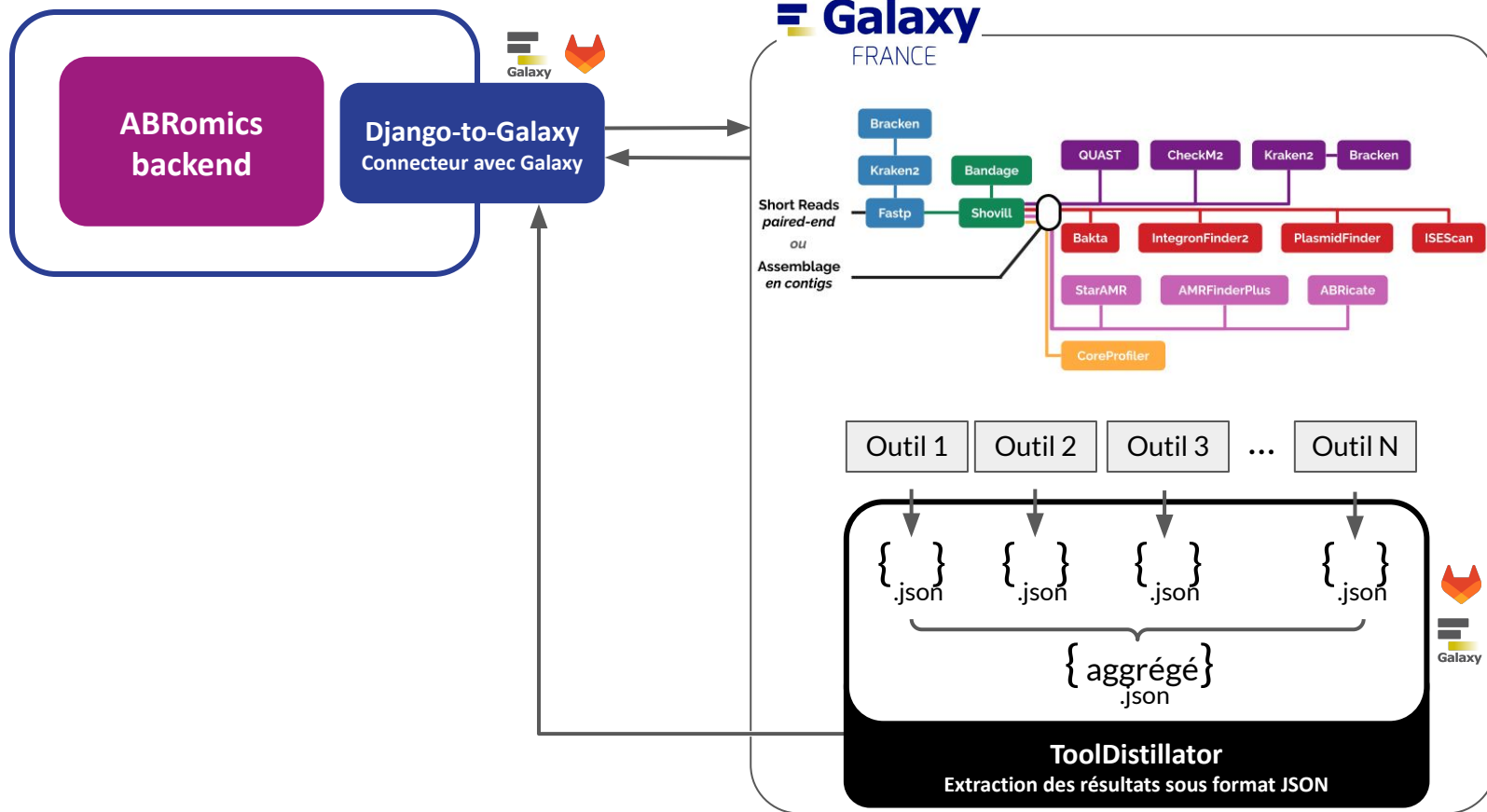
- (fasta) AMR detection
- (fastq paired) Trimming, QC, Assembly, Taxonomy, Annotation, AMR detection
- (fastq paired) Trimming, QC, Assembly, Taxonomy, AMR detection

5 Explorer les résultats de la communauté (Emplacement des échantillons de la base de données ABRomics disponible sur la carte)

ABRomics Database

DOWNLOAD XLSX



Sample Information	Taxonomy assignment	Resfinder	Genes	Plasmid	Analysis type	Permission	Actions
NAJ20050 United States of America (Genome test BEP) 2019	Not available	Estimate	Drug resistance	10 found	GENOMIC FASTA	Public	1
31221014 France (Genome test BEP) 2019	Bacteroides fragilis (H57.39) (resistance) (ST 584)	Estimate	Drug resistance	10 found	GENOMIC WGS paired (WGS FASTQ)	Public	1
France (Genome test BEP) 2022	Not available	Estimate	Drug resistance	73 found	GENOMIC FASTA	Public	1
ES1007 (Labson E. coli from (90.29) (resistance) ST 48)	Escherichia coli (90.29) (resistance) ST 48	Estimate	Drug resistance	39 found	GENOMIC WGS paired (WGS FASTQ)	Public	1
3427791 (Labson E. coli from (90.29) (resistance) ST 504)	Escherichia coli (90.29) (resistance) ST 504	Estimate	Drug resistance	55 found	GENOMIC WGS paired (WGS FASTQ)	Public	1
ES1007 (Labson E. coli from (90.29) (resistance) ST 48)	Escherichia coli (90.29) (resistance) ST 48	Estimate	Drug resistance	2 found	GENOMIC WGS paired (WGS FASTQ)	Public	1
ES1007 (Labson E. coli from (90.29) (resistance) ST 48)	Escherichia coli (90.29) (resistance) ST 48	Estimate	Drug resistance	2 found	GENOMIC WGS paired (WGS FASTQ)	Public	1
ES1007 (Labson E. coli from (90.29) (resistance) ST 48)	Escherichia coli (90.29) (resistance) ST 48	Estimate	Drug resistance	2 found	GENOMIC WGS paired (WGS FASTQ)	Public	1
ES1007 (Labson E. coli from (90.29) (resistance) ST 48)	Escherichia coli (90.29) (resistance) ST 48	Estimate	Drug resistance	2 found	GENOMIC WGS paired (WGS FASTQ)	Public	1



Alerts

Saved searches (database change count is updated daily)

Click on any row, or the search button, to query on the database (redirects to Database)

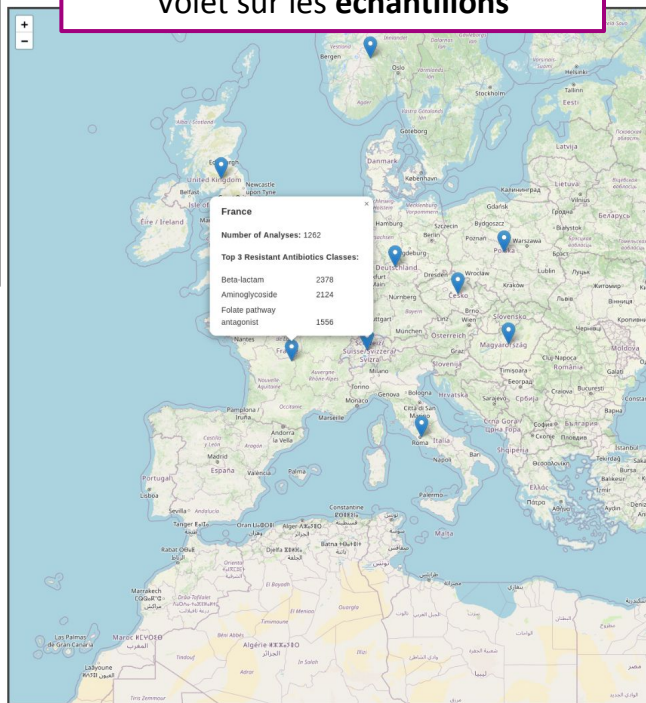
SEARCH NAME	CHANGE COUNT	LAST CHECKED	ACTIONS
<code>source_type == human and host_species == Homo sapiens and microorganism_scientific_name == Escherichia coli and sequence_type == 88</code>	15	mercredi 23 juillet 2025 à 09:42:11	 

“Je veux être informé des nouvelles analyses avec *Escherichia coli* ST 88 chez l’Homme”

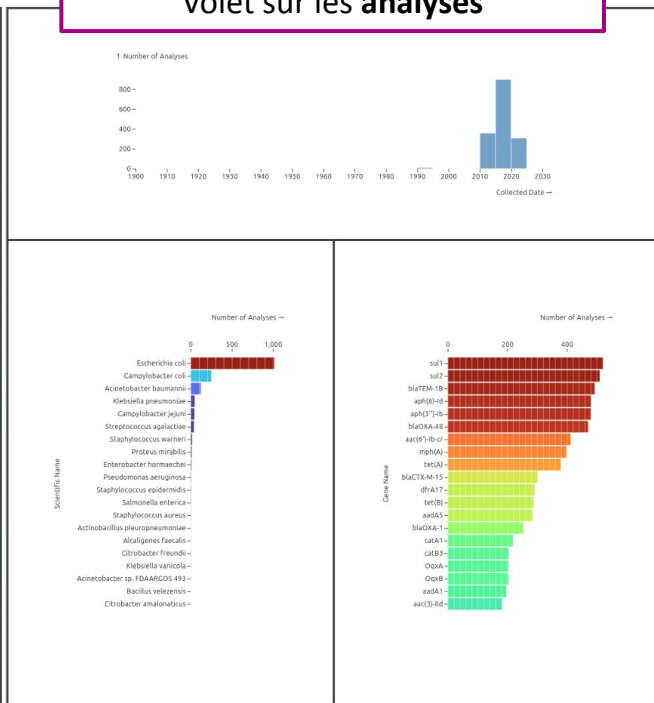
- *Escherichia coli* ST88 est un pathogène aviaire
- Si détecté chez l’Homme → Possible zoonose

Menu
Filtres

Volet sur les **échantillons**



Volet sur les **analyses**



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/>

QUALITY CONTROL SUMMARY - Pass
Quality control of Fastp v1.0.1, Bracken v3.1, CheckM2 v1.0.2 and Quast v5.3.0 results

- Quality Control of Reads**
FastP Q20 bases assessment passed. (values tested: Q20 bases: 323690320, Total bases: 330706656) Pass
- Quality Control of Contamination**
 - Bracken 1st hit reads assessment passed. (values tested: Top species detected: 97.268% reads) Pass
 - Bracken 1st hit species assessment passed. (values tested: Sample metadata and the species detected with the most reads) Pass
 - CheckM2 completeness assessment passed. (values tested: Completeness: 100.0) Pass
 - CheckM2 contamination number assessment passed. (values tested: Contamination: 0.43) Pass
- Quality Control of Assembly**
 - CheckM2 genome size flag assessment passed. (values tested: Assembled genome size: 5774547) Pass
 - QUAST average coverage depth assessment passed. (values tested: Average coverage depth: 57.0) Pass
 - QUAST and CheckM2 N50 assessment passed. (values tested: N50 for a 0bp base: 132865, N50 for a 200bp base: 136626) Pass
 - QUAST contigs number assessment passed. (values tested: #Contigs >= 0bp: 380, #Contigs >= 200bp: 148) Pass

Qualité des reads

basé sur le score Q20 des bases

Homogénéité taxonomique des reads

basé sur la concordance avec l'espèce renseignée par l'utilisateur et la proportion des reads correspondant

Complétude de l'assemblage

basé sur la proportion de gènes de ménage retrouvés

Contamination dans l'assemblage

basé sur le même critère

Taille du génome assemblé

basé sur la fourchette des tailles des génomes du NCBI

Profondeur de séquençage

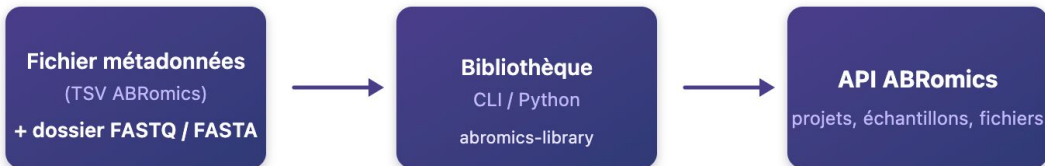
N50 de l'assemblage

Fragmentation du génome

basé sur le nombre de contigs

Qu'est-ce que c'est ?

- SDK Python léger (abromics-library sur PyPI: <https://pypi.org/project/abromics-library/> version 1.1.4) pour parler à l'API ABRomics : projets, samples, envoi de fichiers.
- Interface en ligne de commande (abromics) + bibliothèque Python (AbromicsClient) pour les mêmes opérations.
- Téléversement reprenable (protocole TUS) adapté aux gros jeux FASTQ/FASTA et aux connexions peu stables comme via le front.



A qui ça sert ?

Bioinformaticiens et ingénieurs qui veulent industrialiser le dépôt de données, et pour les gros déposants qui peuvent être confrontés à des limitations via le front.

A Retenir lors de l'usage de la librairie :

TSV Web = TSV SDK + une colonne `Project Name *`

	Place	Collected date*	Travel countries	Accession number	Sample comment	Project Name*
e	Paris	2024-01-15			FASTQ sample ARDIG49	ARDIG49Mass1
e	Paris	2024-01-15			FASTQ sample ARDIG49	ARDIG49Mass1
e	Paris	2024-01-15			FASTQ sample ARDIG49	ARDIG49Mass1
e	Paris	2024-01-15			FASTQ sample ARDIG49	ARDIG49Mass1

CLI

<code>project create</code>	Créer un projet FASTQ (template 1) ou FASTA (template 2)
<code>check-data</code>	Voir pour une souche : métadonnées / assemblage présents ou non
<code>complete-upload-workflow</code>	TSV + dossier → création des échantillons + envoi des fichiers

Ex: `abromics --api-key "abk_..." --base-url "https://analysis.abromics.fr" \`
`complete-upload-workflow \`
`--metadata-tsv "/chemin/samples.tsv" \`
`--data-dir "/chemin/fichiers_sequences"`

Création de l'API Key depuis son compte Abromics dans le front



Create API Key ×

Key Name

Select the permissions for this API key

Complete workflow upload ▾

Complete workflow upload

Datalink Abromics ENA Plugin

CANCEL CREATE KEY

Python

```
from abromics import AbromicsClient
client = AbromicsClient(
    api_key="abk_...",

base_url="https://analysis.abromics.f
r",
)
project = client.projects.create(
    name="Mon projet FASTQ",
    template=1,
    description="Import automatisé",
)
result =
client.batch.process_tsv_and_upload(
    project_id=project.id,
    tsv_file="samples_metadata.tsv",

files_directory="/chemin/vers/sequenc
es",
)
if result["success"]:
    print("Workflow terminé.")
```