INRAE

DeepOmics user guide

Digital Environmental Engineering Platform for Omics data

Ariane Bize
08/12/2023

v 2022-02-05

Table of contents

1. Introduction: what is DeepOmics?

DeepOmics is an information system (IS) dedicated to meta-omics data from environmental biotechnology processes, such as wastewater treatment or anaerobic digestion. It enables the management of data from samples that originate either from full-scale processes, or from laboratory or pilot reactors.

It intends to support the production of FAIR data, thereby promoting data valorization, exchange and reuse. Through its wide use, it will enable data mining and facilitate biostatistical meta-analysis. It could foster innovation by accelerating the development of a microbial management for environmental processes.

In the present version, DeepOmics enables the storage of **amplicon sequencing data** (typically **16S rDNA metabarcoding** data but not limited to them) as well as very rich data describing process design, operating parameters and physico-chemical monitoring measurements. The data stored in DeepOmics can be exported in standard formats (csv, biom, fastq, …). It accepts both **single-end or pair-end data**.

For lab-scale and pilot processes, DeepOmics presently covers reactors with up to 3 compartments. Batch processes are more easily described in DeepOmics, but semi-continuous and continuous processes can also be entered with some limitations.

For full-scale processes, the current version of DeepOmics mainly covers wet and dry digestion, as well as activated sludge. The other types of processes can still be entered with a more limited and standard description.

A documentation website is available: https://deepomics-info.hub.inrae.fr/.

In the near future, we intend to enrich DeepOmics by developing new features. New types of meta-omics data (e.g. shotgun metagenomics, metatranscriptomics) and process types (bioelectrochemical systems) should be covered. Moreover, additional invaluable functionalities should be included such as a userfriendly search interface and the integration with complementary tools (easier sequence submission in the European Nucleotide Archive (ENA), coupling to Easy16S, a userfriendly tool for the interactive statistical analysis of count data from microbial communities, https://shiny.migale.inrae.fr/app/easy16S, doi attribution, etc).

2. License

3. Funding and acknowledgements

DeepOmics was originally developed by the Information Systems Division of INRAE, under the coordination of INRAE-PROSE unit (https://www6.jouy.inrae.fr/prose_eng/), in collaboration with INRAE-LBE (https://www6.montpellier.inrae.fr/narbonne_eng/Laboratory-of-Environmental-Biotechnology/Welcome), INRAE-OPAALE (https://www6.rennes.inrae.fr/opaale_eng/) and INRAE-MaIAGE, MIGALE platform (https://migale.inrae.fr/).

DeepOmics recently benefited from the financial support of the division **Microbiology and the food chain of INRAE** (2020-2021), and of the **3BCAR network** (2022-2025).

We are greatful to Prof. Jo De Vrieze (CREAS, KU Leuven, Belgium) and Dr Claudia Etchebehere (Microbial ecology laboratory, Clemente Estable Biological Research Institute, Montevideo, Uruguay) for helpful discussions.

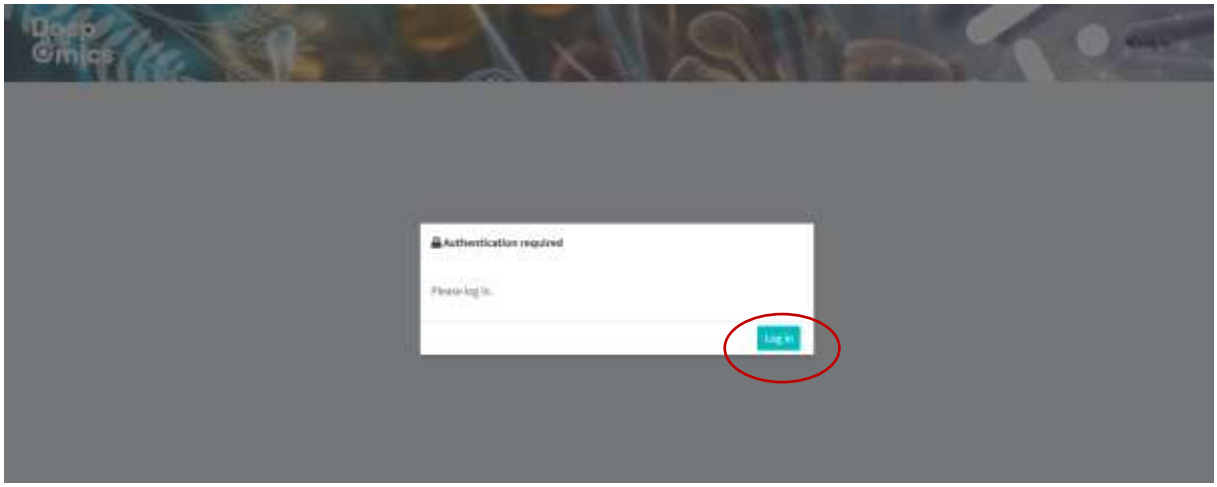
4. How to acess DeepOmics server?

DeepOmics server is located at the following address:
https://deepomics.inrae.fr/

If you depend from Renater federation (national research and education network in France), you can connect to DeepOmics with your Renater login and password. You will then directly access to DeepOmics public projects in read-only mode.

If you are interested in creating your own projects and entering data, please contact us!

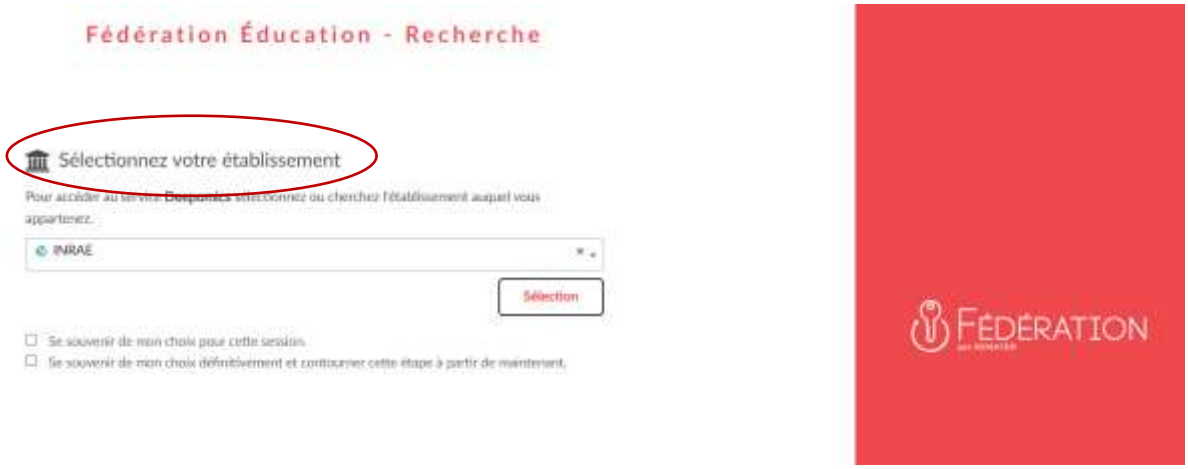
If you do not depend from Renater federation (e.g. private entities, abroad academic entities), please contact us!



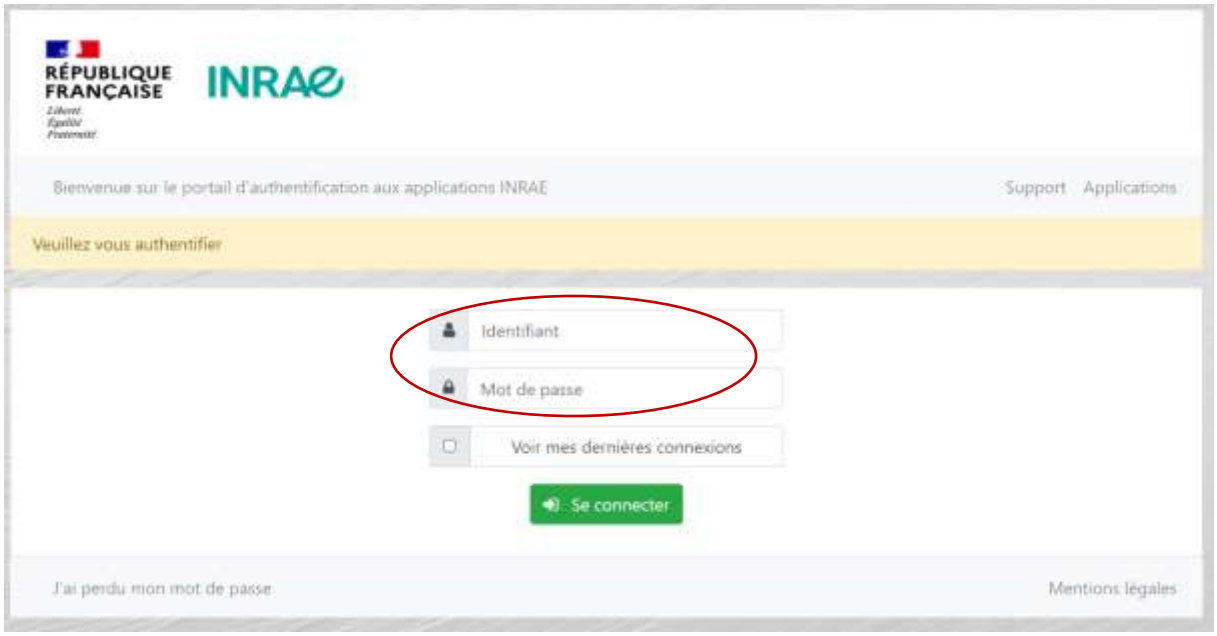
Overview of the sign-in interface

Click on *Log in*.

You are directed to the authentication interface of Renater federation. Select your organization, and then enter your usual login and password.



Overview of the sign-in page in Renater federation (step 1).



Overview of the sign-in page in Renater federation (step 2)

You are now connected to DeepOmics IS through your personal account.

Overview of a user welcome page

You can start using the interface. When connected, you have access to the public projects of DeepOmics, including 2 demo datasets:

- *00_demo_AD_inhib_Poirier*: a demo dataset for lab-scale experiments; it contains data from Dr Simon Poirier's PhD work (2013-2016) on anaerobic digestion inhibition by phenol or ammonia (2 experimental series, with 48 biosamples for each). This project was supervised by Dr Olivier Chapleur (INRAE-PROSE).
- *00_demo_AD_plants*: a demo dataset for full-scale processes; it contains data related to the feeding and sludge sampled from 6 various AD plants, (one time point per plant, total of 20 biosamples). The data has been anonymized.

5. Which bioinformatics pipeline to use?

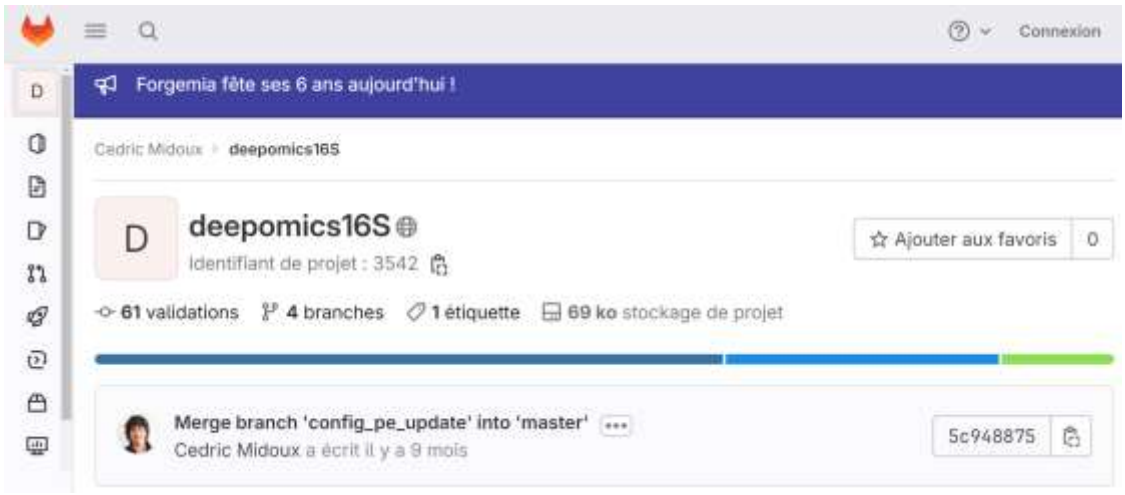
To analyze amplicon sequencing data, processing them with bioinformatics tools is required before data interpretation.

In DeepOmics IS, we provide the link https://forgemia.inra.fr/cedric.midoux/deepomics16S) to the pipeline we recommend for 16S rRNA gene metabarocoding analysis, for 4 main reasons:

- this pipeline is based on state-of-the-art tools, including DADA2[1] and FROGs[2];
- the output files generated by the pipeline are directly compatible with import into DeepOmics;
- if all the users employ the same bioinformatics pipeline, it will promote data homogeneity and intercomparability;
- the pipeline is suitable for an information system, enabling the addition of datasets by batches or individually, without the need to run the pipeline again on the whole database.

As stated in the introduction, **both single-end and pair-end data** are supported presently.

Link to the recommended bioinformatic pipeline (gitlab page)



Overview of the Gitlab page dedicated to deepomics16S bioinformatic pipeline

[1]Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13,** 581–583 (2016). https://doi.org/10.1038/nmeth.3869

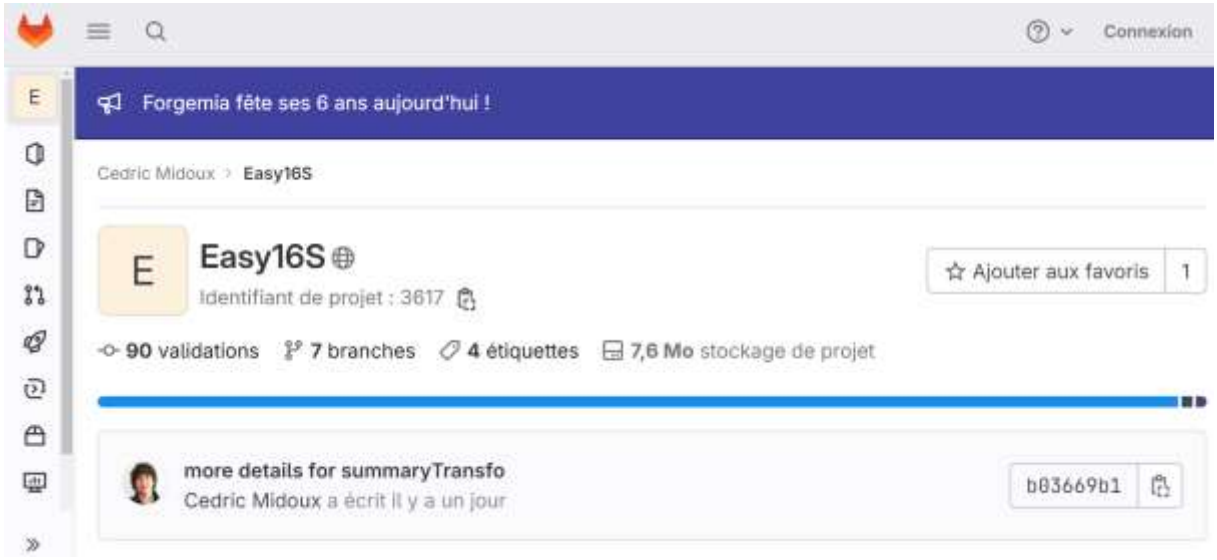
[2] Escudié, P, Auer, L, Bernard, M, Mariadassou, M, Cauquil, L, Vidal, K, Maman, S, Hernandez-Raquet, G, Combes, S, Pascal, G. FROGS: Find, Rapidly, OTUs with Galaxy Solution, *Bioinformatics*, Volume 34, Issue 8, 15 April 2018, Pages 1287–1294, https://doi.org/10.1093/bioinformatics/btx791

6. Easy16S, a complementary tool for statistical analysis of microbial ecology omics data

In DeepOmics, we also provide the link to a complementary tool named Easy16S, which can be used to lead classical multivariate and other statistical analysis of biom files generated by DeepOmics. Easy16S code is available on Gitlab (https://forgemia.inra.fr/cedric.midoux/easy16s).

Link to Easy16S, a userfriendly tool for microbial ecology data analysis

Easy16S (https://shiny.migale.inrae.fr/app/easy16S) is a user-friendly and free access shiny web application that enables the dynamics visualization of count data in microbial ecology (biom or other entry formats) and the mapping of covariates.



Overview of the Gitlab page dedicated to Easy16S

The plots generated in the interface can be downloaded as images. Moreover, the code used to produce each plot can be displayed, copied and pasted in an external text file, to keep trace of the analysis.

Overview of Easy16S interface (using one of Easy16S demo dataset)

Overview of the code display functionality in Easy16S

7. File formats for the import of biom files and metrics

If you are using the pipe-line recommended for DeepOmics, the biom file will be in the correct format and the metrics file in the correct format will automatically be generated.

The biom file must be provided in format biom-1.0.

https://biom-format.org/documentation/format_versions/biom-1.0.html

A metrics file must also be provided. It must be a tab-separated file with the following headers:

“Sequencing sample code”: the code for the biosample/fastq (**same as in DeepOmics and as in the biom file**).

“Input reads”: number of raw reads for each sample

“Post-process reads”: number of reads after the preprocessing, for each sample

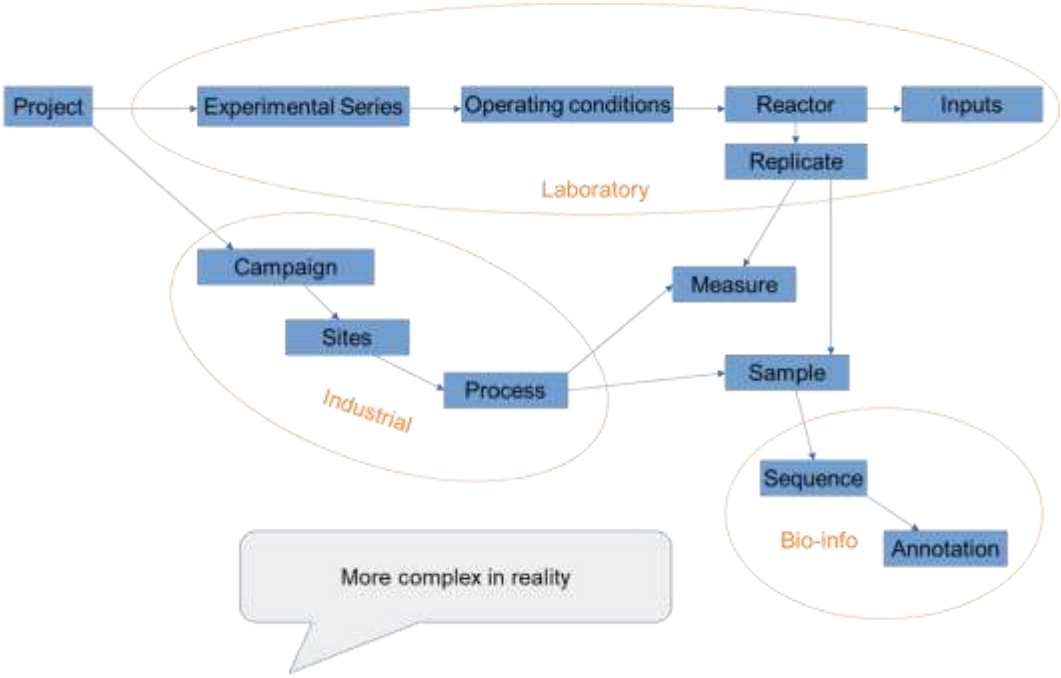
“Nb ASV”: number of amplicon sequencing variants in each sample.

In this metrics file, each line corresponds to one biosample.

An example of such file is provided below:

Sequencing sample code	Input reads	Post-process reads	Nb ASV
DIG_3_f_MAX	43006	33621	56
DIG_3a_s	63190	31427	192
DIG_3a_s_MAX	71649	37428	237
DIG_3b_s	81791	42241	207
DIG_4_f	108888	80711	236
DIG_4_s_MAX	68089	38158	243

8. DeepOmics key concepts



Simplified model scheme of DeepOmics concepts

8.1 Lab-scale process data

Project's input management

The inputs are the compounds and elements which will be used in the reactors (feeding, gas, inocula, buffers, matrix, pure microbial strain, …). They are described according to a controlled vocabulary and they are defined at the scale of the project.

Experimental series

They represent a consistent batch of experiments led into reactors or pilots, with a project. They are structured into operating conditions and replicates.

Operating conditions

Each reactor can be composed of 1 to 10 distinct compartments.
A given operating condition can be defined at the level of the reactor (if it is identical for all compartments) or at the level of each compartment.
In each operating condition, you will be able to define replicates.

Replicates

Reactors which were subjected to the exact same treatment; they are grouped by operating conditions.

8.2 Industrial and field process data

Sampling campaigns, Sampling site and Biotechnological Process

You need to start by creating a **'Sampling site'**, which corresponds to the industrial site from which the samples originate (e.g: a wastewater treatment plant).
Secondly, you need to create a **'Biotechnological process'**, with the precise description of the process and reactor from which the samples originate. Indeed, a single industrial site can gather several processes (e.g. activated sludge, anaerobic digester), hence the relevant ones.
Finally, you need to create a **'Sampling campaign'** and you will then be able to enter your data.

Biosamples

Biosamples represent physical samples for which you plan to perform meta-omics analysis. It is advised to create them in DeepOmics before the acquisition of the corresponding meta-omics data.
You can add multiple biosamples at a time by clicking on **'+ Import biosamples'** (batch mode, through the filling and upload of a template).
Alternatively, it is possible to create biosamples manually, one at a time, by clicking on **'+ New biosample'** (interactive mode).

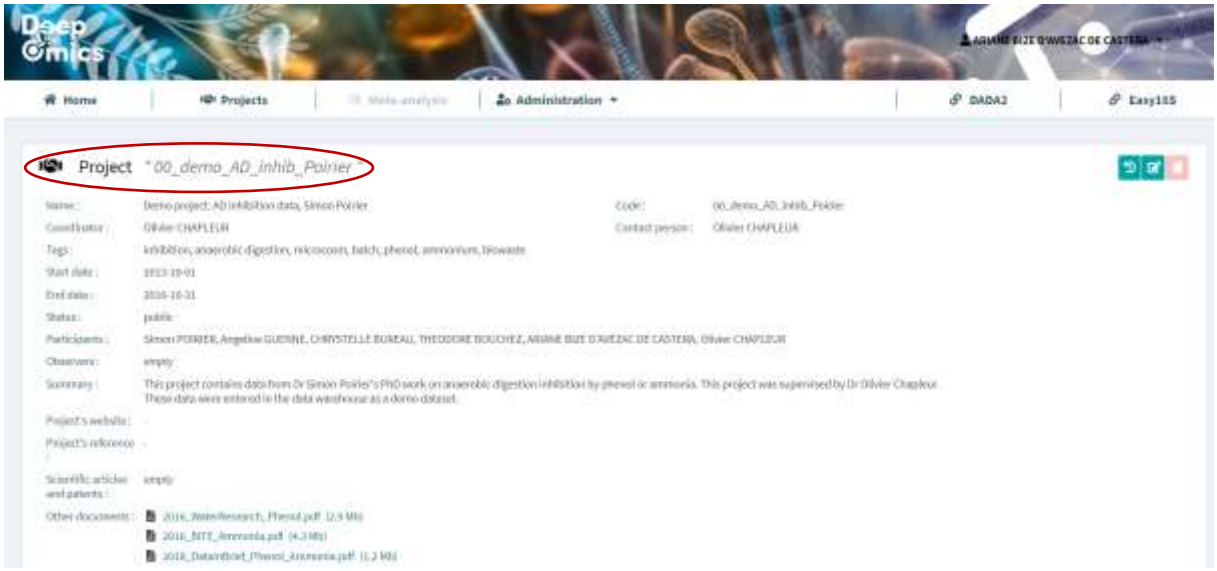
9. DeepOmics structure and key functionalities

9.1 Projects

The DeepOmics data warehouse is organized into *Projects*. Each connected user can create projects and indicate the project *Coordinator* (himself or a third person). The project *Coordinator* can in particular define:

- the project *Status*: *private* (all the project data are private) or *public* (all the project data are public); the project *Status* can be modified at all time;
- the list of *Participants*: all participants of the project have the same rights, namely reading and writing data;
- the list of *Observers*: observers have access to the project in read only mode.

Overview of a user welcome page, with the project table

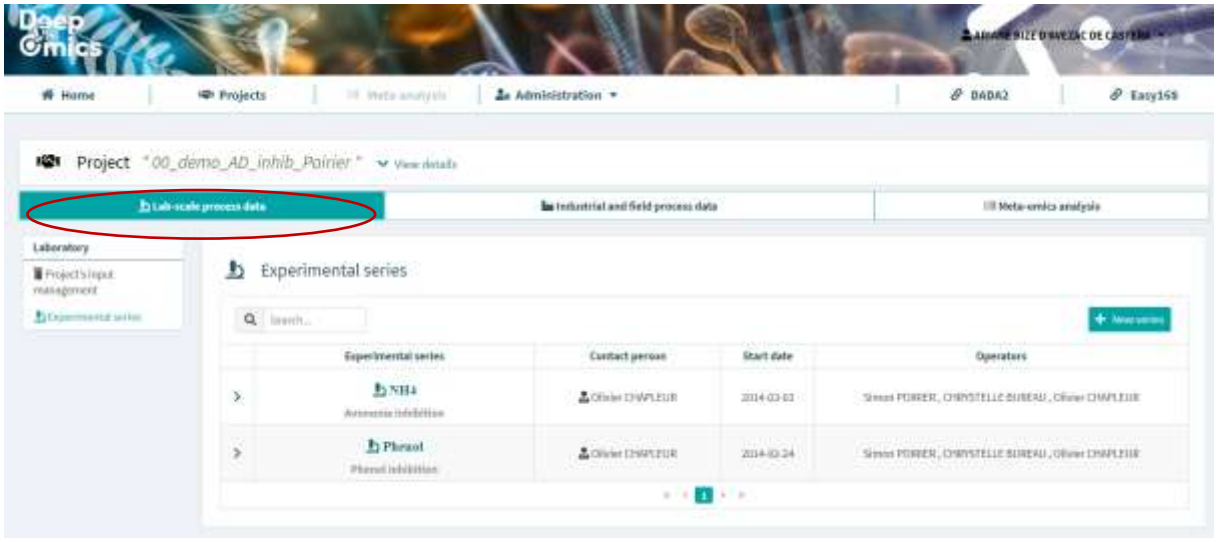


Overview of a project description

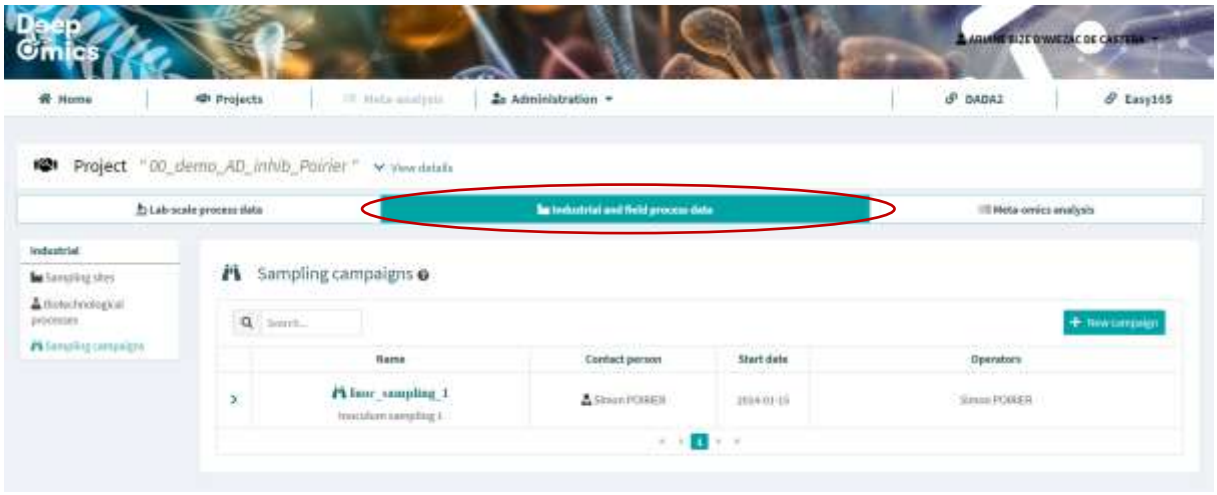
Within a project, 3 main modules are available, accessibles through tabs, from left to right:

1. a tab dedicated to *Lab-scale process data* (or small reactors or pilots)
2. a tab dedicated to *Industrial and field process data* (from full-scale processes)
3. a tab dedicated to *Meta-omics analysis* and data

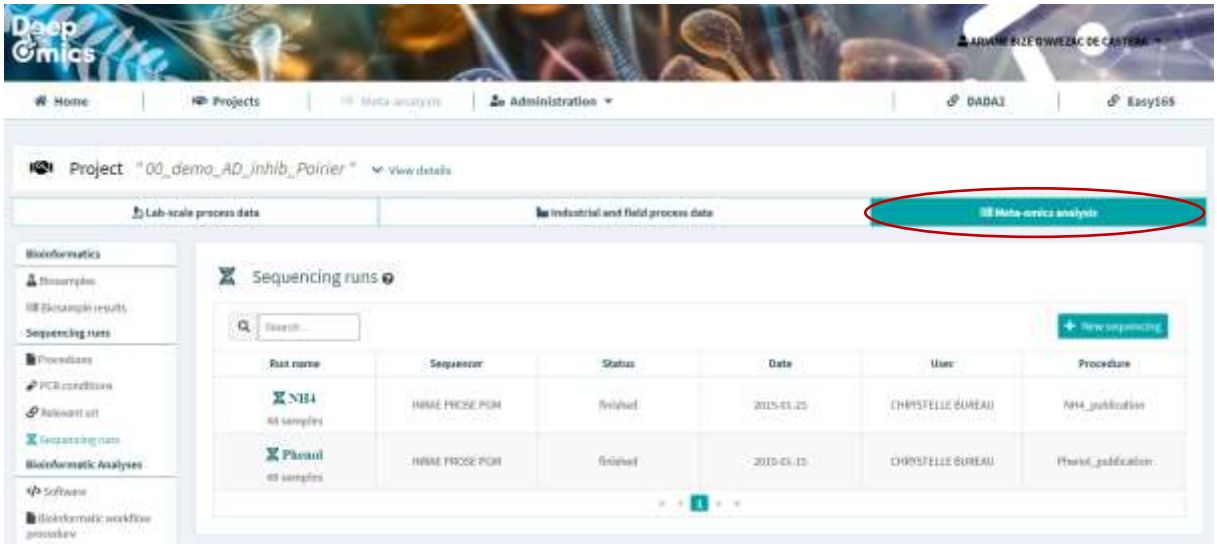
Within each tab, a dynamics menu bar is available on the left, whose item list is contextualized according to the current screen.



Overview of the tab "Lab-scale process data"



Overview of the tab "Industrial and field data"



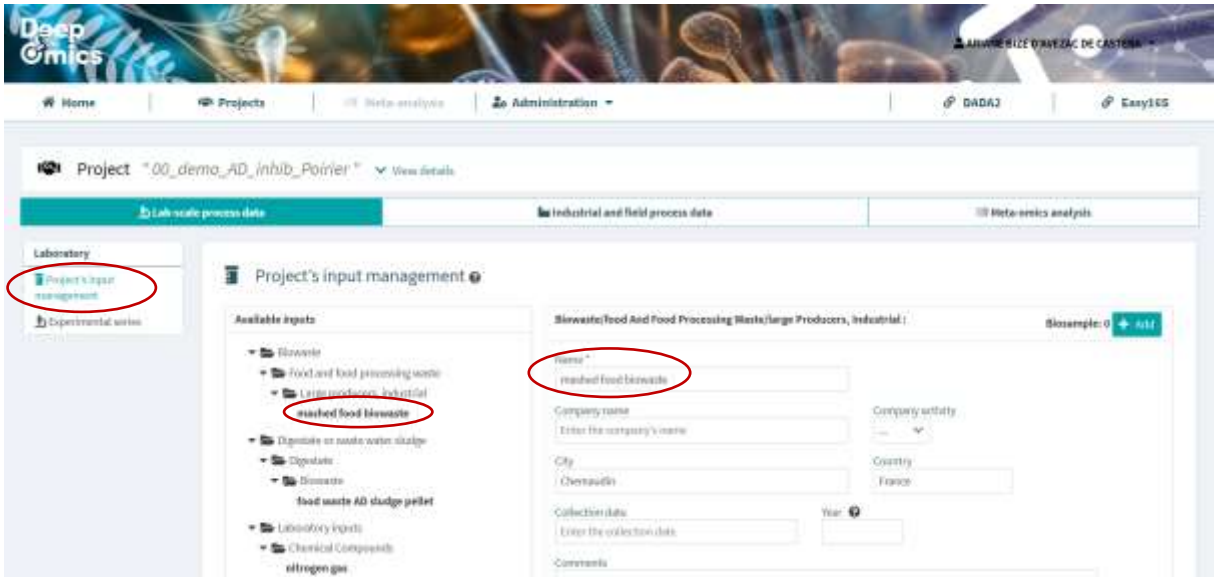
Overview of the tab "Meta-omics analysis"

9.2 Lab-scale process data

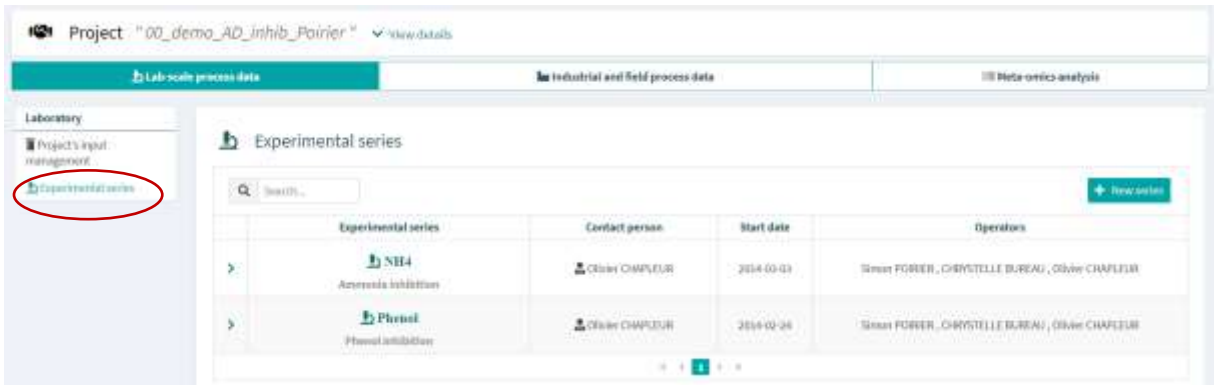
This tab is adapted to process reactors for which highly detailed information is available (operating conditions, nature and amount of the influents, etc), and/or if the experimental design includes replicates.

It contains functionalities related to:

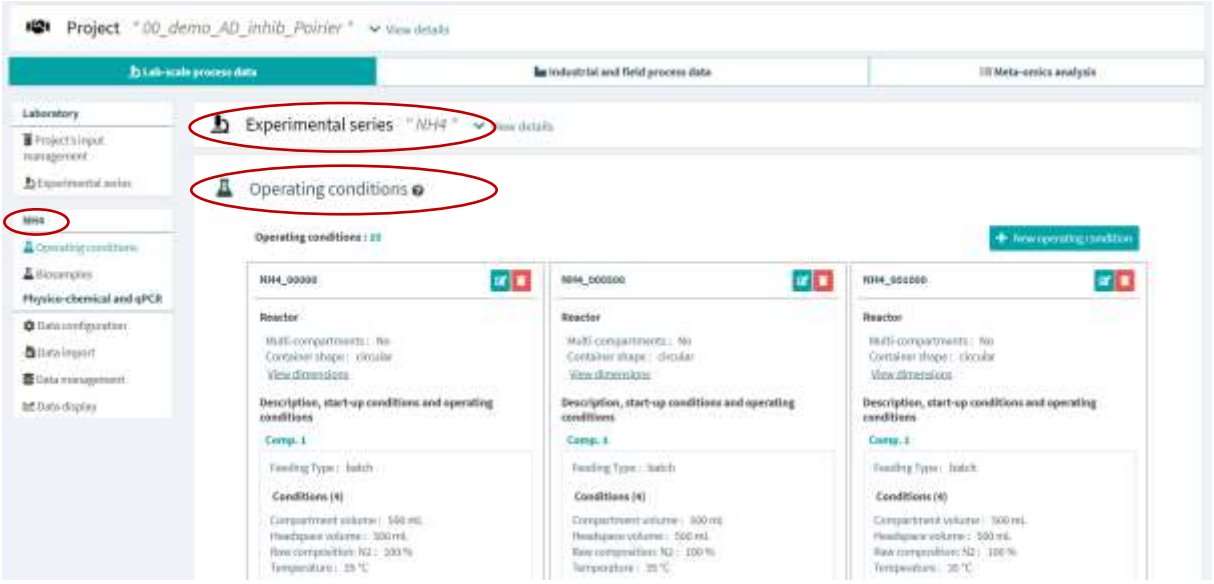
- *Project's input management*: for lab-scale process data, the pool of inputs added to the reactors must be defined in the first place; inputs are categorized according to a controlled vocabulary and they include all types of residual bioresources (waste, sludge, biomass, effluent), as well as gas, chemical compounds, matrices, pure microbial strains, etc. The inputs are defined at the project level.
- *Experimental series*: the experimental series are structured into *Operating conditions* and, within each operating condition, into *Replicates*.



Overview of an input in the interface "Project's input management"

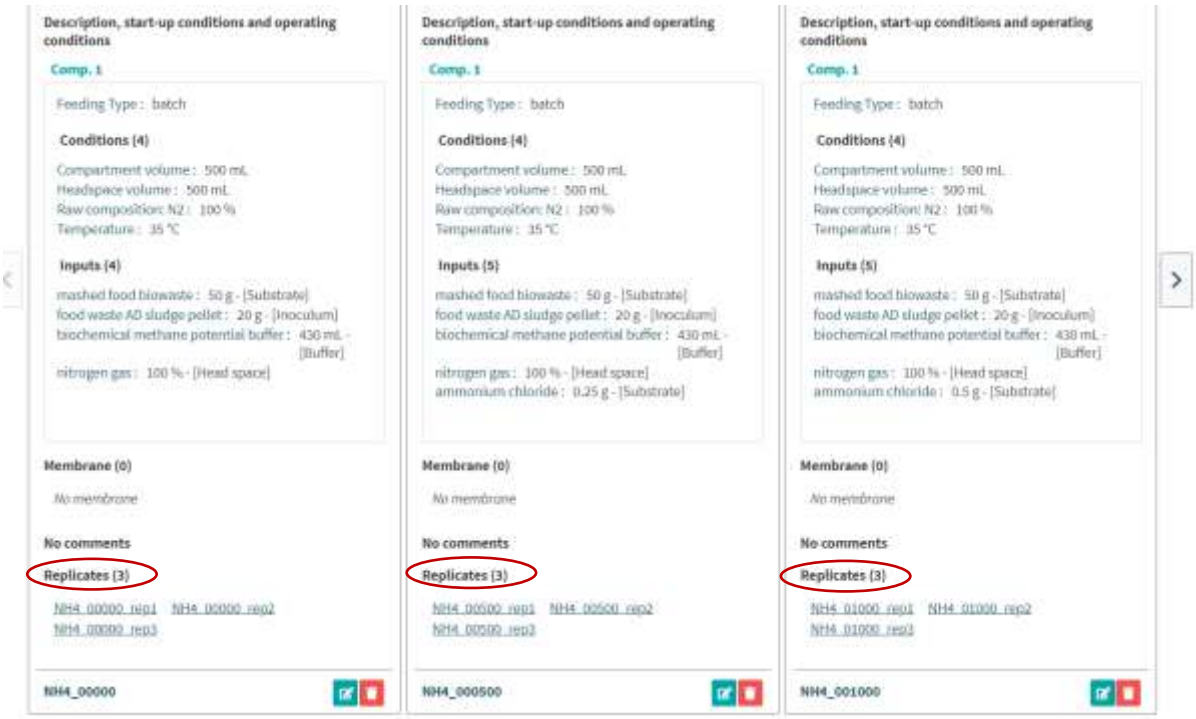


Overview of the interface "Experimental series"



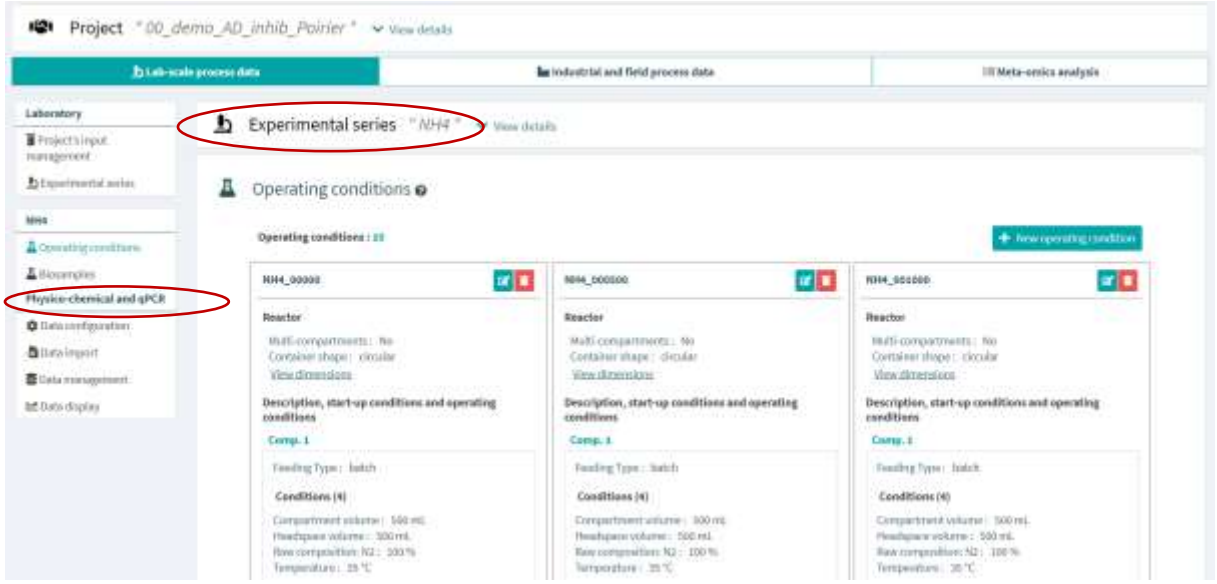
Overview of an experimental series selected in the list of "Experimental series" (Operating conditions)

The *Replicates* are visible at the bottom of the interface for *Operating Conditions*.



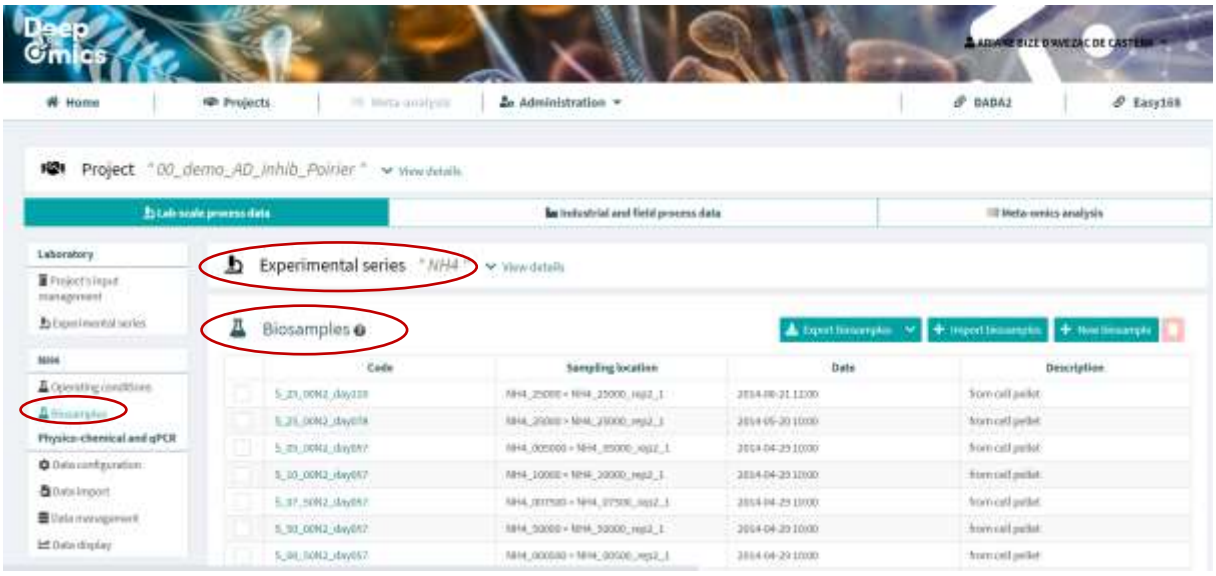
Overview of an experimental series selected in the list of "Experimental series" (lower part: Replicates)

Within the *Experimental series* interface, you can notice, in the left menu bar, a group of functionalities dedicated to *Physico-chemical and qPCR* data (*Data configuration*, *Data import*, *Data management*, *Data display*).



Overview of an experimental series selected in the list of "Experimental series" (Physico-chemical and qPCR)

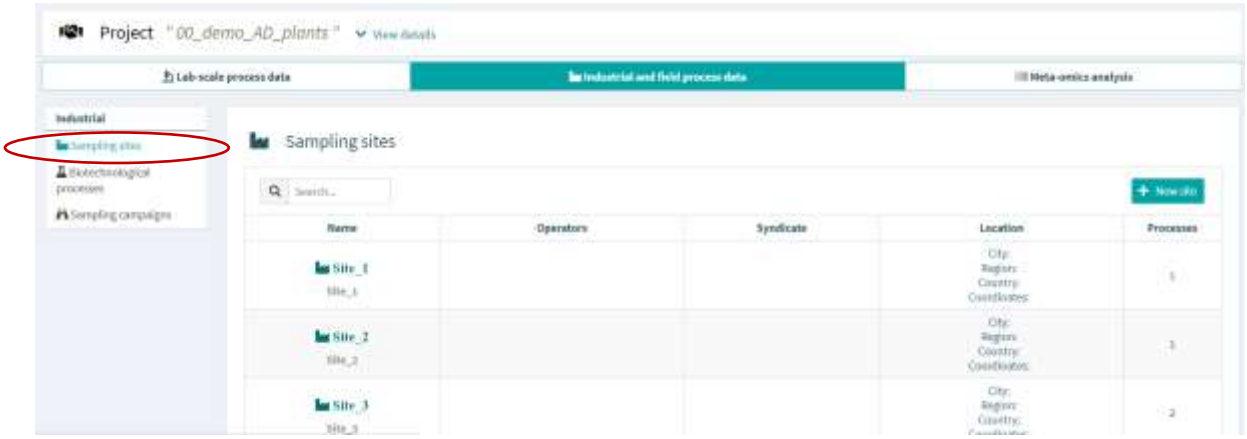
In the left menu bar, the *Biosamples* functionality displays all the biosamples of the considered *Experimental series*. *Biosamples*, created by the project users, are the virtual equivalent of a real experimental sample and they are defined in order to subsequently upload the associated amplicon sequencing data.



Overview of an experimental series selected in the list of "Experimental series" (Biosamples)

9.3 Industrial and field process data

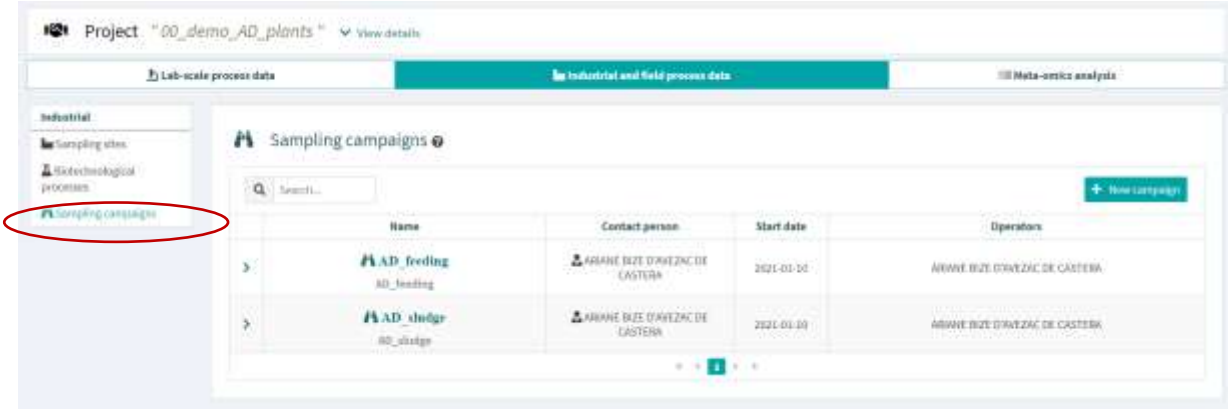
This tab is adapted to data from full-scale processes (no replicates, usually less detailed information on the process, parameters and influents). It is structured into *Biotechnological processes*, which are associated to *Sampling sites*. The process type is defined according to a controlled vocabulary. Moreover, a form specific for each type of process is available to describe the process parameters (currently available: wet and dry anaerobic digestion, activated sludge; the other processes have a basic form until further developments).



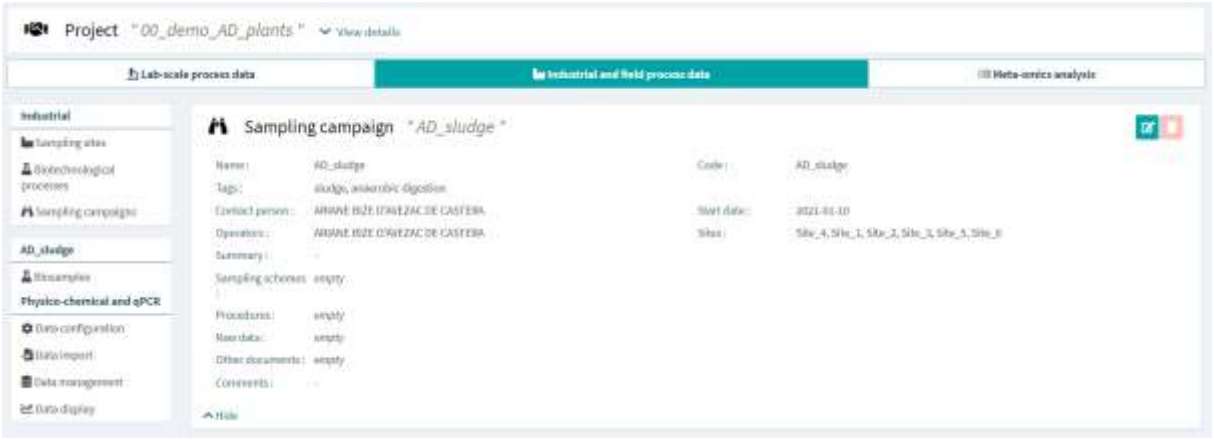
Overview of the interface "Sampling sites"

Overview of the interface "Biotechnological processes"

Sampling campaigns can then be defined and each of them is associated to one or several *Sampling Sites*.



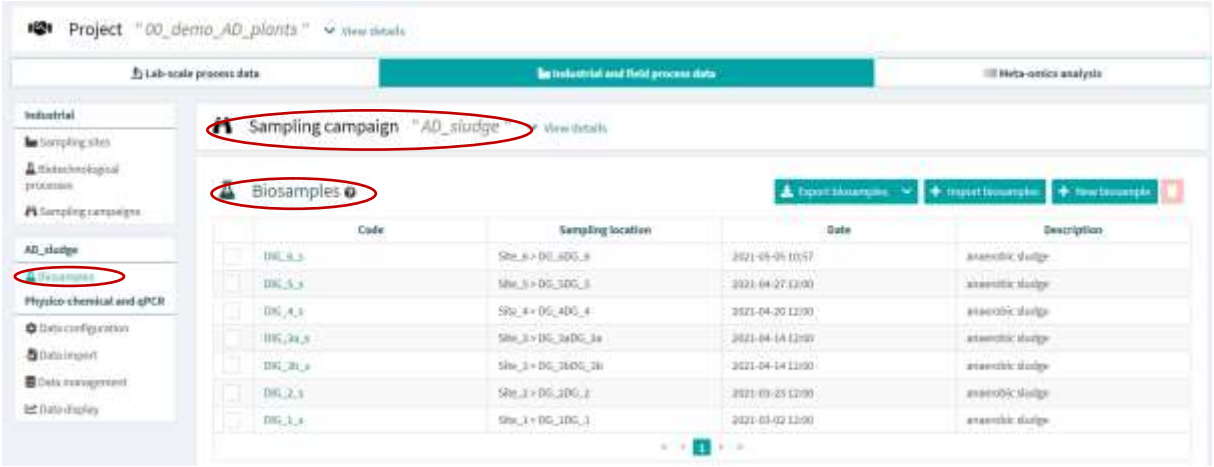
Overview of the interface "Sampling campaigns"



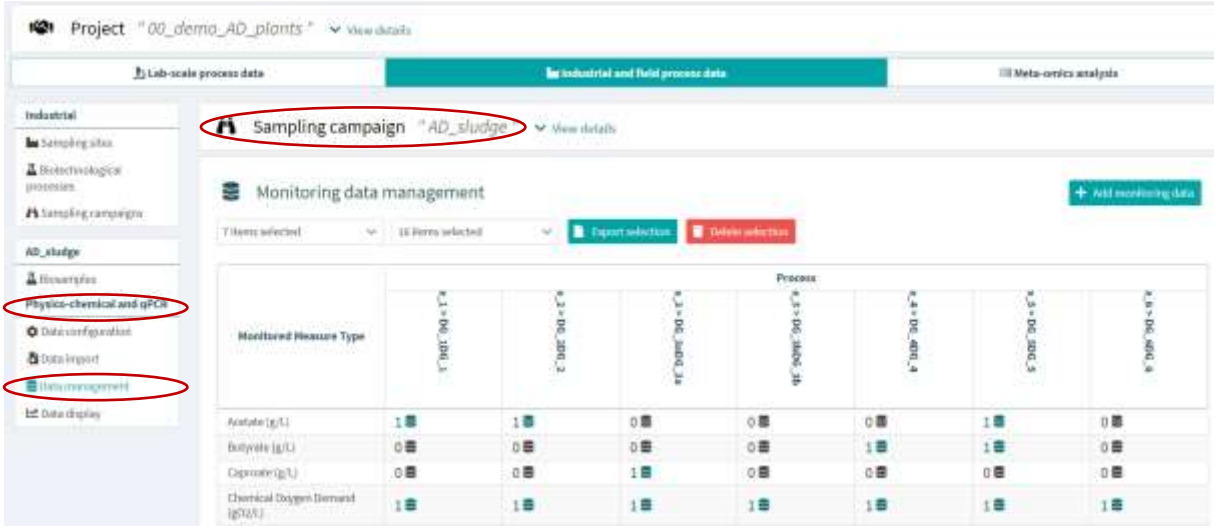
Overview of a selected "Sampling campaign" associated to several "Sampling sites"

Similar to the *Experimental series* interface, you can access to the *Biosamples* related to a selected *Sampling campaign* in the left menu bar.

And you can also notice, in the left menu bar, the group of functionalities dedicated to *Physico-chemical and qPCR* data (*Data configuration*, *Data import*, *Data management*, *Data display*).



Overview of the "Biosamples" interface, within a "Sampling campaign"

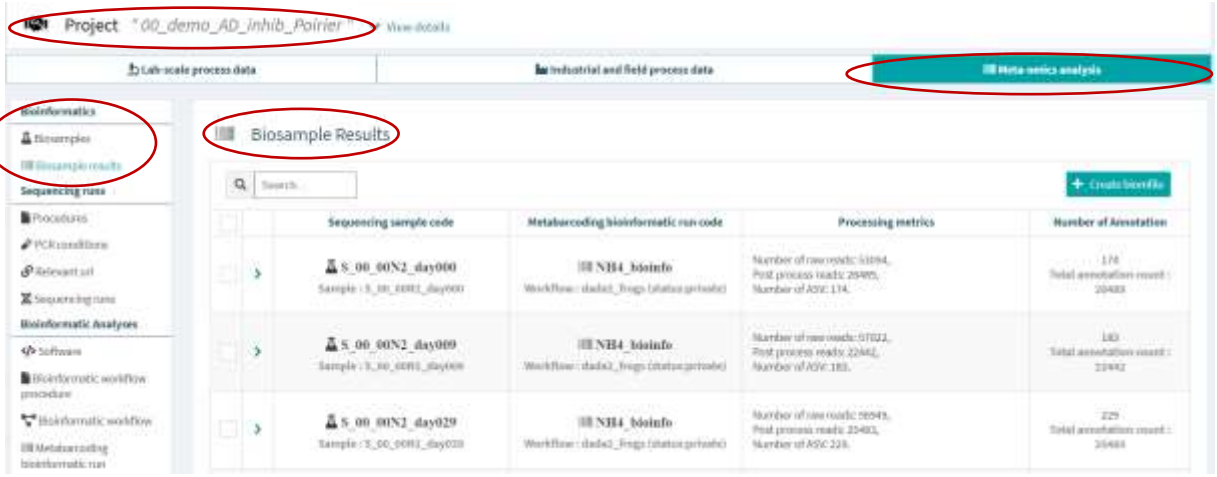


Overview of a sampling campaign the interface "Sampling campaigns" (Physico-chemical and qPCR)

9.4 Meta-omics analysis

This tab is dedicated to amplicon sequencing data. The left menu bar contains 3 main groups of functionalities.

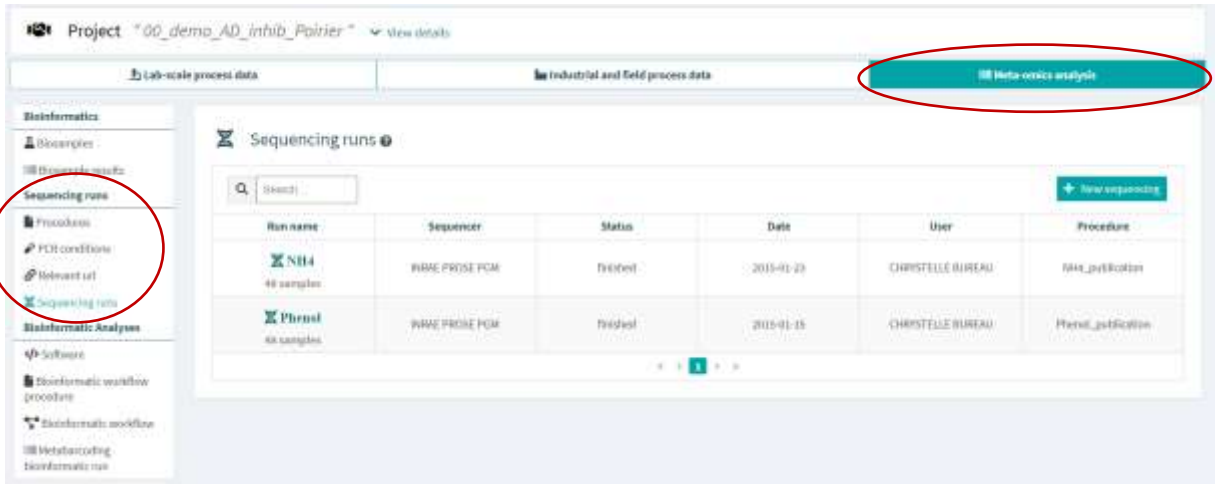
The first group, *Bioinformatics,* enables the visualization of all the *Biosamples* from the *Project* (all *Experimental series* and all *Sampling campaigns* together), and their associated amplicon sequencing results.



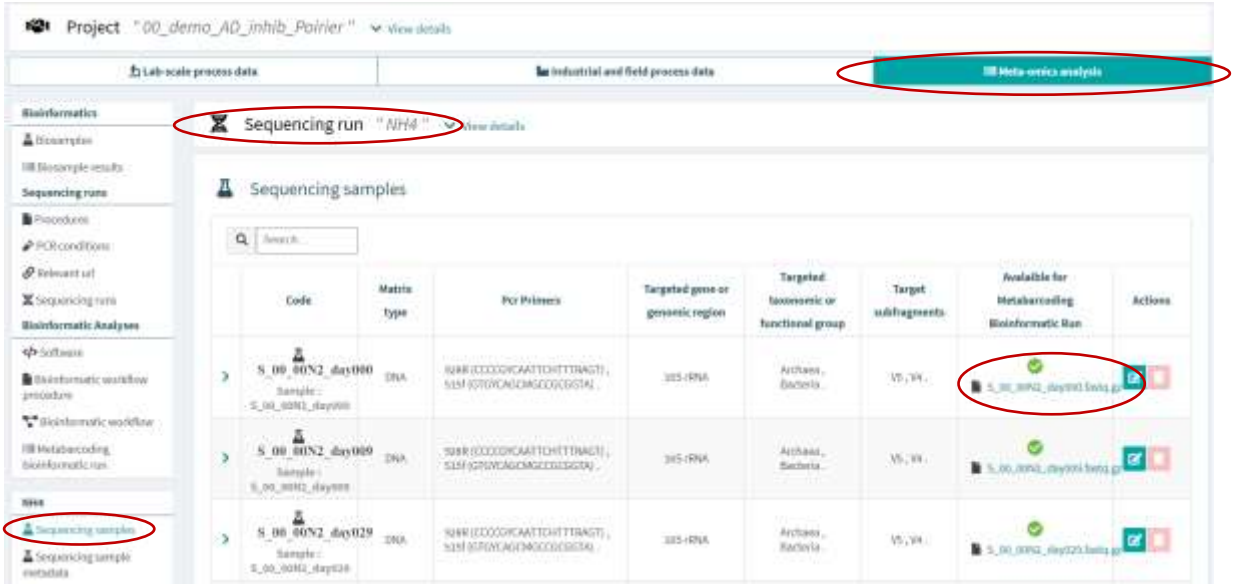
Overview of "Biosample results", for the project "00_demo_AD_inhib_Poirier"

The second group is dedicated to the *Sequencing runs*. The *Sequencing runs* can include *Biosamples* from all the considered *Project*. In the results of a *Sequencing runs*, the raw *fastq* files are in particular available for download, in a compressed format (*.gz*). Several functionalities are dedicated to the

metadata required to describe the *Sequencing runs* with relevant parameters (*Procedures*, *PCR conditions*, *Relevant url*).

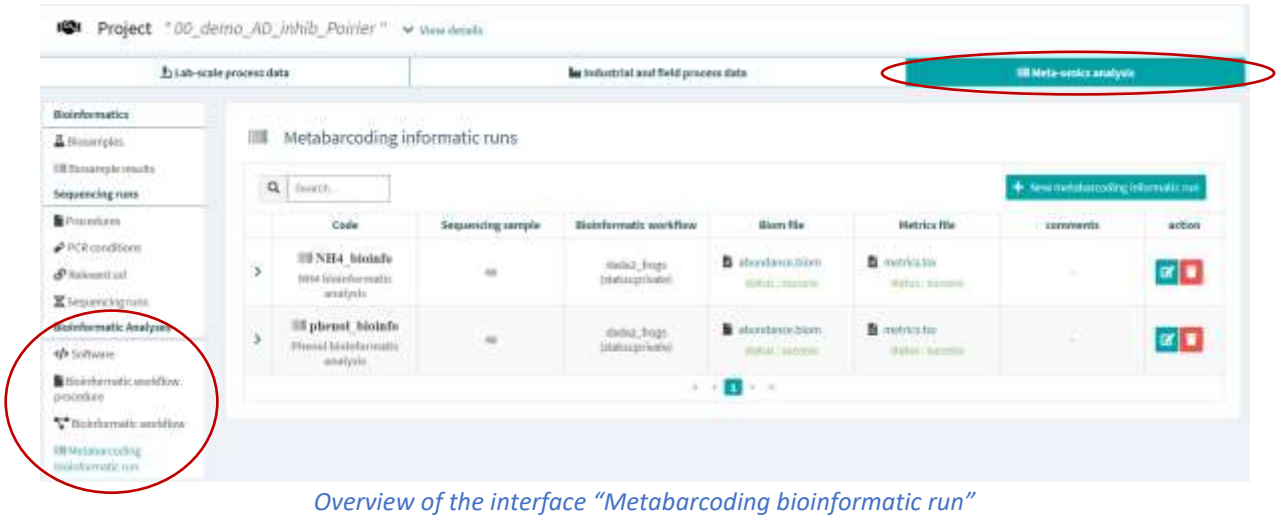


Overview of the interface "Sequencing runs"



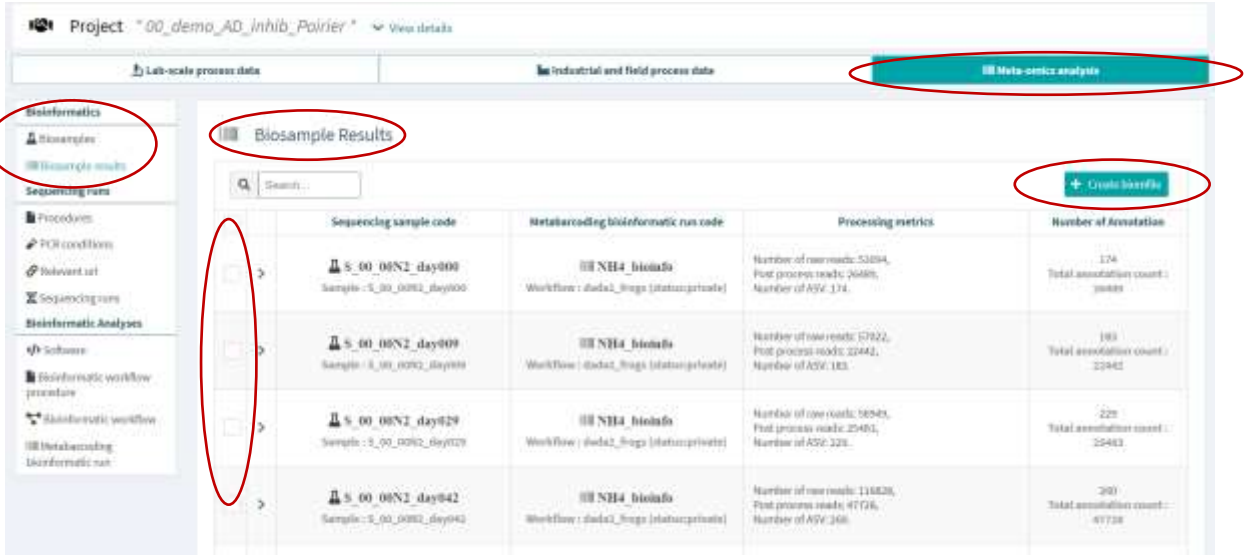
Overview of a sequencing run selected in the interface "Sequencing runs"

The third group, finally, is related to the *Bioinformatic Analyses* of the raw reads. Again, several functionalities ensure the adequate description of the tools and pipe-lines (*Software*, *Bioinformatic workflow procedure*, *Bioinformatic workflow*) used for the analysis. For each sequencing run, at least a biom file and a specific metrics file are loaded, and they can be subsequently downloaded from the interface.



Overview of the interface "Metabarcoding bioinformatic run"

The bioinformatic results from individual samples can be combined into a new *biom* file in the *Biosample results* interface, by selecting the desired *Biosamples*.



Overview of "Biosample results", including the funciton "Create biomfile"

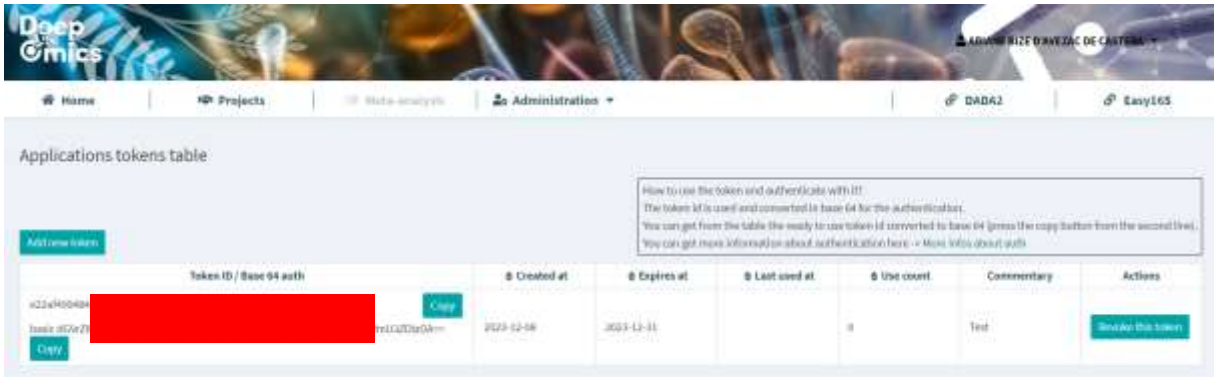
10. Query in DeepOmics

DeepOmics data can be queried and modified through an API (Application Programming Interface).

https://deepomics-api.prose.inrae.fr/

Application tokens can be obtained directly on DeepOmics server, by clicking on the *App tokens* link below the *Logout* button.

Overview of the access to "Application tokens" page



Overview of the "Application tokens" page

No user-friendly query interface is available at the moment, but it is planned to develop one.

11. Useful browsing tips

11.1 Physico-chemical and qPCR

Data configuration

How to add a new monitored parameter?

Select one parameter in the left panel tree and fill the form.

How to edit or delete a monitored parameter?

Select one Monitored Measure in the right panel tree.
If data has already been entered, it is impossible to:
Delete Monitored Measure or compartment with data and **Edit unit**

Data import

Generate a CSV or XLSX template to import your data

1. In the left panel, select at least one **monitored parameter**
2. In the right panel, select at least one **location** (for lab-scale processes, the possible locations are the compartments of the reactor replicates; for industrial and field processes, the possible locations are the processes).
3. In the middle panel, select your **template type**
4. Click on **"Generate template"** below
5. Once the template is filled, **upload** it in the middle panel, at the bottom of the page

Data management

Select at least **one Monitored Data** and **one Replicate/Process**

Data display

This interface enables the creation of **Graph collections**, which will automatically updated according to newly entered data in the considered **Experimental series** or **Sampling campaign**.

11.2 Meta-omics analysis

Sequencing runs

Each sequencing run gathers biosamples from the project. It can combine biosamples from distinct experimental series, sampling campaigns, and inputs. In DeepOmics, the raw sequencing data and/or the data processed with a bioinformatics pipe-line can be uploaded.

To analyze amplicon sequencing data, we strongly advise to use the pipe-line developed specifically for DeepOmics. Indeed, it is convenient as the obtained outputs are directly compatible with import into DeepOmics. Moreover, this pipe-line is well adapted to microbial variants comparison in a data warehouse framework (relying on DADA2 for the clustering, which produces Amplicon Sequencing Variants, ASVs). Finally, using always the same tool will favor homogeneity across the DeepOmics data warehouse, which aims at promoting meta-analysis. However, if you prefer to use your own pipe-line, it is possible, provided that the data are formatted in the correct way for upload into DeepOmics (see format in section 7).

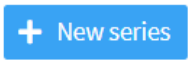
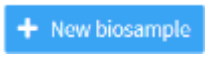
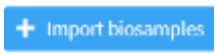
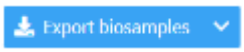
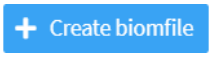
DeepOmics is intended to store and request meta-omics data analyzed beforehand with bioinformatics pipe-line. It is not oriented towards the visualization and statistical analysis of meta-omics data. For such tasks, it is possible to use Easy16S, a distinct user-friendly web application, freely available on INRAE-MIGALE bioinformatics platform.

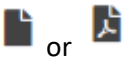
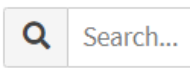
Biosamples

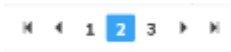
The biosamples are created within each experimental series (laboratory data) or sampling campaign (industrial data). In the tab dedicated to meta-omics analysis, all the biosamples from the project are visible (all experimental series and all sampling campaigns); it is however only possible to visualize them, not to modify them.

11.3 Common browse buttons

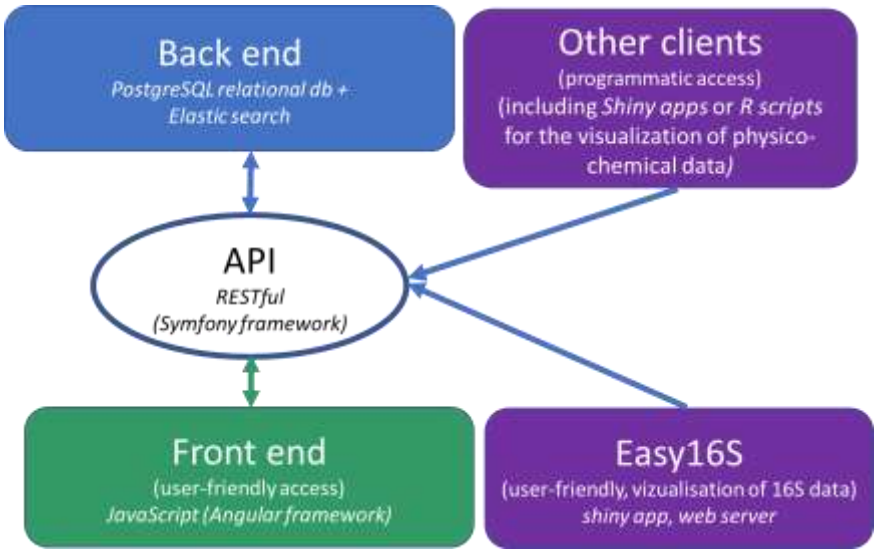
Widget	Function
	Tooltip: a message pops up when the mouse pointer hovers over this symbol
	To edit the corresponding data
	To delete the corresponding data

	To delete the corresponding item
	To delete the corresponding item
	Mandatory field
	To add a missing element
	To enter new data of the indicated type (current example: a new series)
	To enter a new biosample (individually, by filling a form on the interface)
	To import biosamples (batch mode, using the xlsx template)
	To export biosamples (xlsx file)
	To generate a biomfile from the selected biosample results
	To view the project history
	Save button

	Close button
Close	Close button
 or	Indicates a file which can be downloaded from the interface
	Indicates that the required data are complete
	Indicates that some required data are still missing
View details	Expand button
Hide	Collapse button
	Expand button
	Collapse button
	Search bar
	Check box

	Pagination bar
	Pagination bar (for Operating conditions)
	Left and right pagination buttons (for Operating conditions)

12. Informatic structure



Schematic view of DeepOmics informatics structure

DeepOmics Information System is an n-tier web application: the user interface is a single page application built with the Angular framework. It accesses the data using a RESTful API. Data are stored in a PostgreSQL relational database. Easy16S is an interactive R shiny interface based on two main R packages, shinydashboard and phyloseq. Easy16S is currently deployed on the INRAE-MIGALE server (https://migale.inrae.fr/).

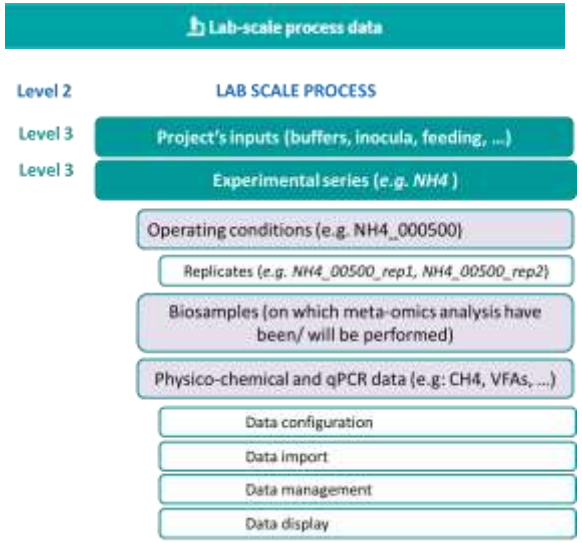
List of illustrations

Memo on DeepOmics functional structure

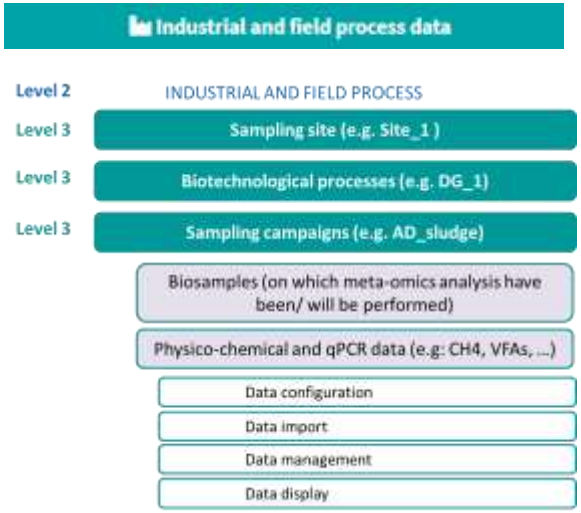
A recap on DeepOmics structure

Level 1 **Projects (e.g. 00_demo_AD_inhib_Poirier)**

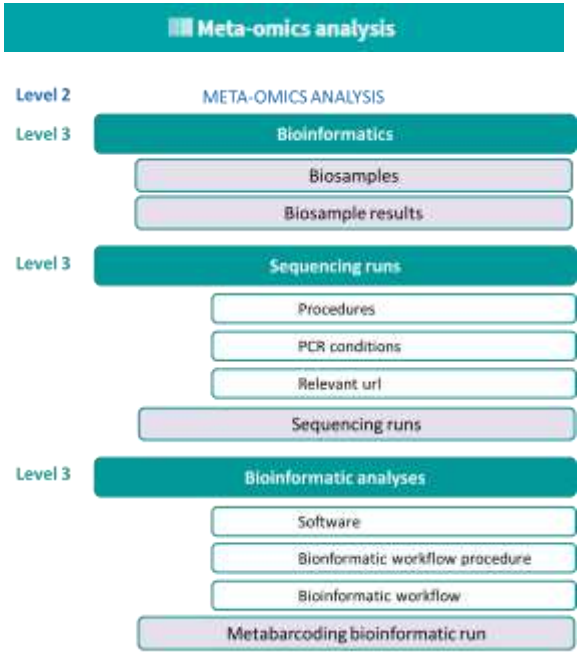
Left tab: labscale



Middle tab: industrial and field



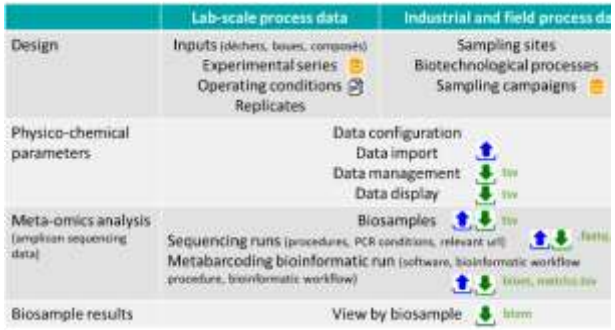
Right tab: meta-omics



Quick start / Memo on how to enter a new dataset into DeepOmics

Main steps of the procedure

Summary table



	Lab-scale process data	Industrial and field process da
Design	Inputs (déchets, boues, composés) Experimental series Operating conditions Replicates	Sampling sites Biotechnological processes Sampling campaigns
Physico-chemical parameters	Data configuration Data import Data management Data display	
Meta-omics analysis (amplicon sequencing data)	Biosamples Sequencing runs (procedures, PCR conditions, relevant url) Metabarcoding bioinformatic run (software, bioinformatic workflow procedure, bioinformatic workflow)	
Biosample results	View by biosample	

Project, reactors and processes

CREATE A PROJECT

1. Make sure you are on the project page (click on Home or Projects in the uppest menu bar)
2. Create a new Project (private or public) (*e.g. 00_demo_AD_inhib_Poirier* or *00_demo_AD_plants*)
3. Click on the created project to reach the project page

CASE 1, LAB WORK, DESCRIBE THE REACTORS

1. Example in *00_demo_AD_inhib_Poirier*
2. Make sure you are in the up-left panel: Lab-scale process data
2. Create the Inputs used in the project (buffers, substrates, sludge, chemical compounds, etc)
3. Create a new Experimental series (*e.g. NH4*)
4. Create Operating conditions (*e.g NH4_000000, ...*) and describe in the form the precise conditions (*e.g. number of compartments, input amounts, temperature, volume, etc*)
5. Create Replicates for each Operating condition (*e.g. NH4_00000_rep1, NH4_00000_rep2, ...*)

CASE 2, INDUSTRIAL PLANTS, DESCRIBE THE CAMPAIGNS

1. Example in *00_demo_AD_plants*
2. Make sure you are in the up-middle panel: Industrial and field process data
2. Create a new Sampling sites (*e.g. Site_1*, left menu bar)
3. Create a new Biotechnological process (*e.g. DG_1*, left menu bar)
4. Create Sampling campaigns (*e.g. AD_sludge*, left menu bar)

Physico-chemical data

DESCRIBE THE ANALYTICAL PARAMETERS AND UPLOAD DATA

1. Make sure you are either in an Experimental series (lab) or in a Sampling campaigns (industiral) (you can navigate by clicking on the up and left menu bars)
2. Click on Data configuration (left menu bar)
3. Select on the left the parameters corresponding to your experiment, set their unit, and validate)
4. Click on Data import (left menu bar)
5. Select on the left the desired parameters, in the middle the desired template type and on the right the desired reactors/processes
6. Generate the xlsx template
7. Fill in the xlsx template, for instance in Excel application (external to DeepOmics)
8. Import the data by selecting the filled xlsx file, at the bottom of the same DeepOmics page

CREATE GRAPH COLLECTIONS TO VISUALIZE THE DATA

1. Clic on Data display (left menu bar)
2. Clic on New graph collection
3. Chose a name and create New graphs

Amplicon sequencing data (e.g. 16S rRNA gene metabarcoding)

CREATE BIOSAMPLES

Biosamples correspond to samples for which you plan to acquire / have acquired amplicon sequencing data

WARNING: chose identical names for the biosamples, fastq files and sample names in the biomfile

1. Make sure you are either in an experimental series (lab) or in a sampling campaign (industrial) (you can navigate thanks to the up and left menu bars)
2. Click on Samples in the left menu bar
3. Click on Import samples (top right) and subsequently on Download template (xlsx)
4. Fill in the template with approriate information, e.g. in Excel application (external to DeepOmics). You can find some help on the DeepOmics page, by clicking on ?Show help to fill in template, or directly in the template, on the 2d sheet
5. Import the Biosamples by selecting the filled-in xlsx file, at the bottom of the same DeepOmics page

NB: For experimental series (lab), a biosample can also be created for Inputs. In this purpose, make sure you are in the interface Lab-scale process data (left pannel). Click on Project's input management towards the top of the left menu bar. Then, click on the desired Input (e.g. a sludge). Finally, click on +Add, on the top right of the form, next to "Biosample".

CREATE A SEQUENCING RUN

1. Make sure you are in the right panel (Meta-omics analysis)
2. Start by creating the description of your Procedures (extraction, amplification, library, sequencing, etc). It can be a document, an url or a doi. Also describe the PCR conditions (left menu bar). This step is not mandatory but it is advised. These functions are available in the left menu bar. The described procedures will be available for selection (through their name) in the template created afterwards.
3. Click on Sequencing runs (left menu bar) and add a New sequencing (top right). Fill in the form and Save.

CREATE SEQUENCING SAMPLES AND IMPORT SEQUENCING RESULTS

1. Make sure you are in a sequencing run (if required, click on Sequencing runs in the left panel and click on the desired sequencing run).
2. You should see: Manage sequencing samples. Download the xlsx template, fill it in (e.g. in Excel application, external to DeepOmics) and import the filled-in template on the same DeepOmics page.
3. Towards the bottom of the left menu bar, click on Sequencing sample metadata
4. Export the xlsx template, fill it in (e.g. in Excel application, external to DeepOmics), and import the filled-in template on the same DeepOmics page.
5. Click on Import fastq (towards the bottom of the left menu bar)
6. Export, fill-in and import the xlsx template, similar to above.
7. On the same page, add the fastq files (raw sequencing data)

NB1: the Biosamples of a Sequencing run can be selected in the whole Project (all Experimental series and all Sampling campaigns)

NB2 : when clicking on Biosamples (top of the left menu bar), you can see the list of all the Biosamples of your project

A recap on the different xlsx templates

1. Template for **physico-chemical or qPCR data** (to import the data corresponding to such monitored parameters)
2. Template for **biosamples** (samples likely to be sequenced)
3. Template for **biosamples included in a given sequencing run**
4. Template for **sequencing metadata** associated to the sequencing run (basically the same template as above, to further fill in)
5. Template for **fastq files**